Extending sharpness calculation for an alternative loudness metric input

S. Hales Swift, and Kent L. Gee

Citation: The Journal of the Acoustical Society of America **142**, EL549 (2017);

View online: https://doi.org/10.1121/1.5016193

View Table of Contents: http://asa.scitation.org/toc/jas/142/6

Published by the Acoustical Society of America

Extending sharpness calculation for an alternative loudness metric input

S. Hales Swift[a] **and Kent L. Gee**
Department of Physics and Astronomy, Brigham Young University, Provo, Utah 84602, USA
hales.swift@gmail.com, kentgee@byu.edu

Abstract: Sound quality metrics help improve the psychoacoustic acceptability of devices and environments by modeling and thus enabling deliberate improvement of perceptual attributes. Sharpness as defined in DIN 45692 [(2009). Deutsches Institut für Normung, Berlin] requires inputs from Zwicker's loudness metric [ISO 532-1 (2017). International Organization for Standardization, Geneva]. This letter demonstrates that sharpness can be formulated to accept specific loudness values from Moore and Glasberg's loudness metric [ISO 532-2 (2017). International Organization for Standardization, Geneva; ANSI S3.4 (2007). American National Standards Institute, Inc., Washington, DC]. Sharpness calculations using the two loudness metrics produce similar results. This method thus enables evaluation of sharpness as a straightforward add-on to standard loudness calculations using Moore and Glasberg's metric, for which sharpness calculations were not previously available.

[QJF]
Date Received: September 7, 2017 **Date Accepted:** November 13, 2017

1. Introduction

Sound quality metrics fill an essential role in efforts to create more desirable product and environmental sounds. By modeling and predicting human responses to sounds, they enable the preliminary assessment of sound quality without having to resort immediately to a listener panel, which may prove unacceptably costly during a preliminary design process. By modeling human responses cheaply, sound quality metrics facilitate the serious consideration of sound quality as a design tool and the selection of designs optimized for psychoacoustic pleasantness. Two perceptual dimensions of sound quality—loudness and sharpness—have sufficiently successful calculation procedures in place that they have been standardized.

The less commonly known of the two, sharpness, relates to the balance of high- and low-frequency components within a sound and is computationally similar to finding the spectral center of mass of a sound. Sounds with greater high-frequency content are described by listeners as having a sharper timbre. The version of sharpness that has been standardized (DIN 45692, 2009) calls for specific loudness (loudness per frequency-like unit) inputs provided by a loudness metric based on work by Zwicker (Fastl and Zwicker, 2007) which has itself been standardized as ISO 532-1 (2017). A second loudness metric, designed by Moore and Glasberg (Moore *et al.*, 1997) has, with further development and modification (Glasberg and Moore, 2006), been embodied in an American standard (ANSI S3.4, 2007) and, with the modifications of Moore and Glasberg (2007), an international standard (ISO 532-2, 2017). However, the current standard sharpness formulation only supports inputs from the Zwicker method. This limits the range of circumstances in which sharpness calculations are available for use. This letter shows techniques for modifying the sharpness metric procedure embodied in DIN 45692 (2009) to accept specific loudness inputs from Moore and Glasberg's loudness metric so that this important sound quality can be more easily assessed in a greater variety of computational contexts.

2. Derivation

In order to implement sharpness using the outputs of the Moore and Glasberg metric (ISO 532-2, 2017; ANSI S3.4, 2007), a transformation of variables is required in order to achieve compatibility. Both loudness models considered here use a perceptually

---

[a] Author to whom correspondence should be addressed.

relevant transform of the frequency scale: the critical-band-rate scale with units of Barks for the Zwicker model and the ERBN-number scale (equivalent rectangular band number; N indicates that this is for a normally hearing individual) with units of Cams for the Moore-Glasberg model. For convenience, these scales will be described as “transformed frequency.” We begin with the formula for sharpness using input from Zwicker loudness (Fastl and Zwicker, 2007; DIN 45692, 2009) shown in Eq. (1). Here, z_1 is the frequency expressed in Bark (Fastl and Zwicker, 2007), N' is the specific loudness (loudness, in sones, per Bark), $g_1(z)$ gives the relationship between frequency transformed to Bark and corresponding sharpness (acum), and C is a calibration constant chosen to give the reference value of 1 acum for a critical-band-wide narrowband noise centered at 1000 Hz with a loudness of 4 sone$_G$,

$$S = C\frac{\int N_1'(z_1)g_1(z_1)z_1dz_1}{\underbrace{\int N_1'(z_1)dz_1}_{N}}[\text{acum}]. \tag{1}$$

The subscript “G” attached to sones indicates that it was calculated using the German standard method, i.e., Zwicker’s method. The sharpness, S, may thus be predicted as a division of integrals with the integral on the bottom giving the total loudness, N, and the integral on the top weighting the specific loudness by frequency somewhat similarly to a center of mass calculation.

In order to use inputs from the Moore and Glasberg loudness model, it is necessary to transform the critical-band-rate scale in Bark, z_1, to the ERBN-number scale in Cams, z_2. This change of frequency variables is accomplished via the usual calculus procedure for a change of variables as in

$$S = C\frac{\int \overbrace{N_1'(z_1(z_2))\frac{dz_1}{dz_2}}^{N_2'}g_1(z_1(z_2))z_1(z_2)dz_2}{\int \underbrace{N_1'(z_1(z_2))\frac{dz_1}{dz_2}}_{N_2'}dz_2}[\text{acum}], \tag{2}$$

where the substitution $z_1 = z_1(z_2)$ and $dz_1 = (dz_1/dz_2)dz_2$ has been made. The value of the bottom integral in Eq. (2) is still the loudness, and given the assumption that both metrics predict loudness accurately, the integrand in the bottom equation must be the specific loudness of the Moore and Glasberg loudness model. The assumption of equivalent loudness prediction requires some qualification because some systematic differences exist between the two loudness metrics. However, differences between sharpness calculations tend to be smaller than the differences between the loudness metrics providing inputs. These differences tend to be smaller for narrowband signals than for broadband signals and will be quantified for characteristic test cases below. Methods for mitigating systematic differences in specific loudness predictions for the purpose of calculating sharpness were explored by Swift and Gee (2017). With these caveats in mind, the quantities indicated by underbraces and overbraces in Eq. (2) can be identified as N_2', the specific loudness distribution in the second loudness calculation scheme. To the extent that the two models agree in loudness prediction and produce a “true” specific loudness prediction, then the relationship between the specific loudnesses in the two models can be expressed as

$$N_2'(z_2) = N_1'(z_1(z_2))\frac{dz_1}{dz_2}[\text{sones/Cam}]. \tag{3}$$

This realization allows us to express the sharpness as a function of the transformed frequency variable in somewhat simplified terms as

$$S = C\frac{\int N_2'(z_2)g_1(z_1(z_2))z_1(z_2)dz_2}{\int N_2'(z_2)dz_2}[\text{acum}]. \tag{4}$$

Equation (4), remarkably, is of the same form as the original sharpness formula: In the denominator is an integral over the specific loudness and in the top we have an integral over the specific loudness multiplied by the transformed frequency scaled by $g(z)$. It should be noted that this formulation requires as an intermediate step evaluating the transformed frequency in Bark, z_1, corresponding to each transformed frequency in Cam, z_2.

3. Implementation and test cases

A sharpness calculation procedure using Eq. (4) was introduced into an implementation of Moore and Glasberg's loudness metric (Moore *et al.*, 1997; Glasberg and Moore, 2006) following ANSI S3.4 (2007). This was compared with results from a previously validated implementation of Zwicker's loudness following ISO 532-1 (2017) in which the sharpness procedure of DIN 45692 (2009) was also implemented and validated. The validation sounds from DIN 45692 (2009) were used as test cases to gauge the performance of the two metrics. Additionally, 40 real-world sounds from Vettel (2009) were used as inputs to the metrics. These real-world sounds included a variety of household sounds from bouncing balls, jingling keys, ringing bells, tearing or crinkling paper, splashing water, typing, sweeping, knocking, and so forth.

The standard's validation test sounds were of two sorts: (1) Sounds that were one critical band wide with the listed center frequency in Bark, and a nominal loudness of 4 sones_G, and (2) high-pass filtered broadband sounds with a lower cutoff frequency at the listed frequency in Bark, an upper cutoff frequency of 10 000 Hz (22.4 Bark) and a nominal loudness of 4 sones_G. Nominal sharpness values for these signals are given in the sharpness standard (DIN 45692, 2009) as well as in Fastl and Zwicker (2007, Fig. 9.1).

For the validation sounds, we compared outputs from the Zwicker metric using the three $g(z)$ functions indicated in the standard—the standard weighting function, the weighting function of Aures and the weighting function of von Bismarck (Aures, 1985; von Bismarck, 1974; DIN 45692, 2009). We compared these with the implementation using the standard weighting function applied to specific loudness values from the metric of Moore and Glasberg. This allowed us to compare the relative magnitudes of errors that exist between the several embodiments of the sharpness metric. For the real-world sounds, we examine only the relative difference between the two metrics using the standard $g(z)$ function, treating the Zwicker outputs as the reference. We also compared the loudness results from the Zwicker metric and Moore and Glasberg metric to assess the magnitudes of the discrepancies between the two metric frameworks for loudness and sharpness calculations in response to the DIN 45692 (2009) test sounds. Free-field frontal incidence listening conditions were employed in all calculations of sharpness and loudness in this paper.

4. Results and discussion

The sharpness predictions for all metric implementations in response to the DIN 45692 (2009) test sounds are shown in Fig. 1. In this figure, the x axis shows the center frequency of the narrowband sounds and the lower cutoff frequency of the broadband sounds. Additionally, the sharpness values from Fastl and Zwicker's "Psychoacoustics: Facts and Models" (Fastl and Zwicker, 2007, Fig. 9.1) are plotted.

As expected, the sharpness predictions from the Zwicker-based metric using the standard $g(z)$ function (black ×'s) are well aligned with the nominal values listed in the standard for these test signals (black line). This is the case for both the narrowband and broadband test signals. Zwicker's metric also agrees well with the standard

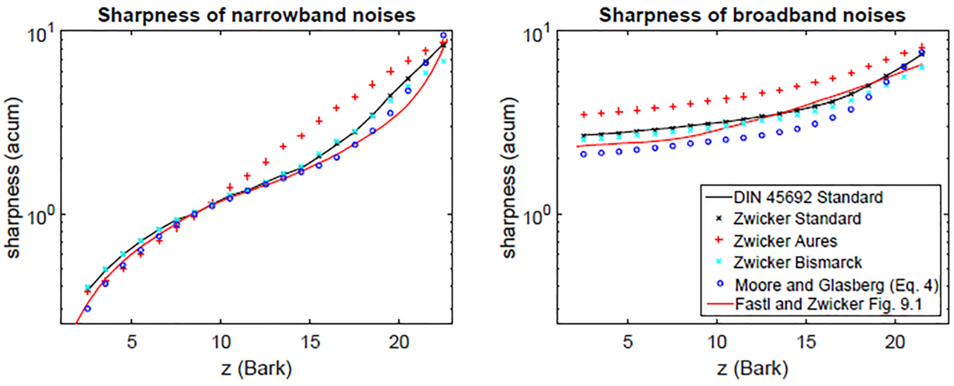

Fig. 1. (Color online) Sharpness predicted using input from Zwicker's and Moore and Glasberg's loudness metrics for the narrowband (left) and broadband (right) test sounds from DIN 45692 (2009) compared with both the standard values and Fastl and Zwicker (2007, Fig. 9.1). The x axis shows center frequency for the narrowband sounds and lower cutoff frequency for the high-pass broadband sounds.

values when the $g(z)$ function of von Bismarck is used, except at high frequencies where small discrepancies occur. The Zwicker-based metric is next used with the $g(z)$ function of Aures, which has the unique feature among the $g(z)$ functions in the standard of taking into account the loudness of the signals, consistent with relationships identified in psychoacoustic data (Aures, 1985). Greater sharpness values are predicted when Aures's $g(z)$ function is used than when the standard $g(z)$ function is used for both broadband signals across all frequencies and for narrowband signals at frequencies above about 10 Bark. This appears to be due to the differences in its general shape from the standard $g(z)$ function at mid to high frequencies. The Moore and Glasberg-based metric implemented as in Eq. (4) shows acceptable agreement with the standard values for narrowband signals and actually agrees more closely with the values from Fig. 9.1 of Fastl and Zwicker's book (Fastl and Zwicker, 2007). Indeed, for narrowband signals it was the closest to these values of any of the metrics evaluated. For broadband signals, the Moore and Glasberg-based metric predicted lower sharpness values than the other metrics while still capturing the relative trends except at high frequencies. There is a significant difference in the specific loudness pattern predicted by the two metrics between about 20 and 22 Bark (Swift *et al.*, 2017), and this drives differences in the broadband cases. We expect better agreement for broadband sounds not containing these frequencies (though this should be verified through calculation) or (as found below) when real-world sounds are considered.

In order to evaluate the agreements and disagreements of the metrics in quantitative terms, we calculated both $\varepsilon_{\max}$, maximum absolute percentage error and $\bar{\varepsilon}$, mean absolute percentage error (sometimes abbreviated as MAPE and hereafter called mean relative error), with absolute percentage error, ε, calculated as

$$\varepsilon = 100\frac{|S_k - E_k|}{S_k}, \tag{5}$$

where S_k is a standard value and E_k is some estimate of it. Both are expressed here as percentages. Metric predictions from each model were compared to the standard and Fastl and Zwicker values, which serve as reference values. These error values are listed in Table 1. For narrowband sounds, the largest differences from both reference data sets occurred when the Aures $g(z)$ function was used, with mean relative error values of $\bar{\varepsilon} = 24.8\%$ and 36.0% relative to the two reference data sets. The other sharpness implementations all showed smaller differences from the reference values. Mean relative error values for the Zwicker metric when using the standard or Bismarck $g(z)$ function and the Moore and Glasberg metric when using the standard function were all less than $\bar{\varepsilon} = 12\%$ relative to both sets of reference values.

When 40 real-life sounds of Vettel (2009) are considered, the difference between the outputs of the two sharpness metrics [each using the standard $g(z)$ function] is smaller than for the validation sounds. The Moore and Glasberg metric differed from the Zwicker metric by a $\bar{\varepsilon} = 4.7\%$ and $\varepsilon_{\max} = 14.5\%$ for this set of sounds. This supports the use of the Moore and Glasberg-based metric for real-life sounds as well as standard test sounds.

Furthermore, these differences in sharpness predicted by the two metrics are of similar magnitude to the differences between the two sets of sharpness reference values. Pedrielli *et al.* (2008) reported just-noticeable difference values of around 0.04

Table 1. Maximum relative error ($\varepsilon_{\max}$) shown as a percentage and mean absolute percentage error ($\bar{\varepsilon}$) for each sharpness model using input from Zwicker's loudness (Z) or Moore and Glasberg (MG) compared with the standard values for the test signals from DIN 45692 (2009) and the values shown in Fastl and Zwicker (2007, Fig. 9.1). The column labeled "alt. stnd." contains comparisons between the two standards.

Narrowband discrepancy	Z-standard	Z-Aures	Z-Bismarck	MG	alt. stnd.
$\varepsilon_{\max}$ Re DIN 45692	4.3%	59.3%	19.7%	20.0%	27.7%
$\bar{\varepsilon}$ Re DIN 45692	0.8%	24.8%	3.5%	9.5%	9.8%
$\varepsilon_{\max}$ Re Fastl and Zwicker	37.8%	84.5%	28.4%	25.8%	38.3%
$\bar{\varepsilon}$ Re Fastl and Zwicker	12.0%	36.0%	11.5%	5.3%	11.8%
Broadband discrepancy	Z-standard	Z-Aures	Z-Bismarck	MG	alt. stnd.
$\varepsilon_{\max}$ Re DIN 45692	1.1%	34.8%	14.9%	21.2%	14.0%
$\bar{\varepsilon}$ Re DIN 45692	0.5%	29.6%	7.0%	17.5%	8.0%
$\varepsilon_{\max}$ Re Fastl and Zwicker	15.3%	51.9%	12.3%	24.1%	16.3%
$\bar{\varepsilon}$ Re Fastl and Zwicker	8.4%	38.2%	7.5%	13.9%	8.8%

acum for a 1.42 acum signal at around 60 dB in their study on perception of sounds from earthmoving equipment, a just-noticeable difference of 2%–3%. However, the differences between the data reported in DIN 45692 (2009) and Fastl and Zwicker (2007) is around $\bar{\varepsilon} = 10\%–12\%$, suggesting that more variation than this can exist even among idealized sharpness results, e.g., both the standard and Fastl and Zwicker sharpness data are given without error bars or other indications of spread in the underlying subjective data. Given that, for both sets of test sounds, (1) the Moore and Glasberg-based metric gives results closer to the standard values than the Aures weighting (which is included as an appendix in the standard) and (2) that its differences with the standard values are comparable to the differences that exist between two well-publicized reference sharpness data sets, it follows that the accuracy of the Moore and Glasberg-based implementation in matching the Zwicker-based metric values is comparable to that of existing sharpness metrics.

These relative differences in sharpness predictions may also be usefully compared with the relative differences between the loudness predictions of the two loudness metrics. These are shown in Fig. 2 and tabulated in Table 2. The maximum differences between the two loudness schemes exceed two sones (out of nominally 4). The mean relative difference between the two metrics using the Zwicker metric as reference is $\bar{\varepsilon} = 11.9\%$ for narrowband and $\bar{\varepsilon} = 40.6\%$ for broadband sounds. These differences existing in standardized metrics seem to reflect considerable latitude of acceptable opinion about the loudness of these test sounds. Fortunately, the sharpness calculation procedure results in smaller differences in predicted values than the differences that exist between the two loudness metrics. Excluding the Aures weighting, all models agree with error of less than $\bar{\varepsilon} = 12\%$ for narrowband and $\bar{\varepsilon} = 18\%$ for broadband signals relative to both standards. The sharpness procedure is thus somewhat less sensitive to differences at earlier stages of calculation than the loudness metrics.

Systematic differences in specific loudness calculation between loudness metrics may slightly limit the comparability of predictions made using different metrics or that compare the sharpness of narrowband sounds to that of very broadband sounds. However, as most practical comparisons use a single metric to compare the sharpness of sounds under study and are between sounds that are more similar to one another than this extreme case, or that differ by a small amount, a sharpness metric based on Moore and Glasberg loudness should be able to provide predictions of sharpness that are useful in many engineering applications and can fill an important void in the availability of the sharpness metric by providing a version compatible with the American standard for loudness (ANSI S3.4, 2007).

5. Conclusions

A mathematical procedure for transforming the sharpness metric to accept specific loudness inputs from the loudness metric of Moore and Glasberg (Moore *et al.*, 1997; ISO 532-2, 2017; ANSI S3.4, 2007) was developed and implemented. The sharpness of the validation signals from DIN 45692 (2009) was predicted using both the Moore and Glasberg-based metric implementation given in Eq. (4) and the standardized Zwicker-based implementation. The latter was tested with all three weighting functions while the former used only the standard curve. The differences between the metric outputs were evaluated relative to the sharpness values specified in DIN 45692 (2009) and Fastl and Zwicker (2007). Differences between nominal predicted values for the Moore and Glasberg-based implementation appear to be in the acceptable range of values as

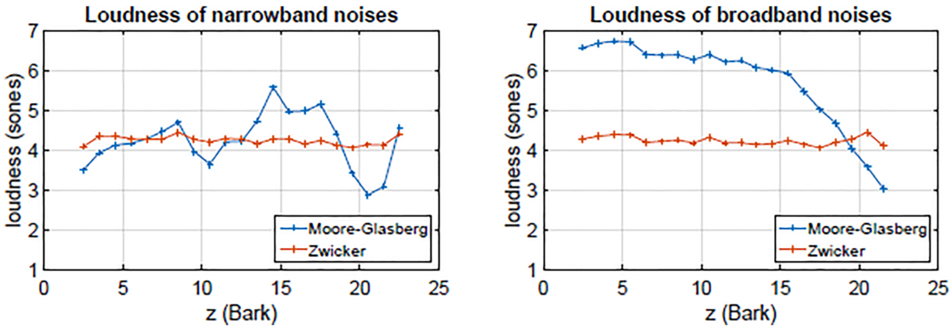

Fig. 2. (Color online) Loudness predicted by Zwicker's and by Moore and Glasberg's metrics for the narrowband (left) and broadband (right) test sounds from DIN 45692 (2009) compared to one another. The x axis shows center frequency for the narrowband sound and lower cutoff frequency for the high-pass broadband sounds.

Table 2. Maximum absolute percentage error ($\varepsilon_{\max}$) and mean absolute percentage error ($\bar{\varepsilon}$) for loudness prediction differences between the Zwicker and Moore and Glasberg metrics in response to the narrowband and broadband test sounds from DIN 45692 (2009).

Narrowband $\varepsilon_{\max}$ difference	30.4%
Broadband $\varepsilon_{\max}$ difference	53.5%
Narrowband $\bar{\varepsilon}$	11.9%
Broadband $\bar{\varepsilon}$	40.6%

determined by comparing the outputs of the Zwicker-based metric using the weightings included in the standard and the differences between the DIN 45692 and Fastl and Zwicker reference values. The two metrics were also compared using the standard weighting function for a set of 40 real-world sounds from Vettel (2009). The metric outputs were more similar for real-world sounds than for the DIN 45692 test sounds with an average relative error of $\bar{\varepsilon} = 4.7\%$. The approach taken in this paper of validating a metric relative to another metric, while sufficient to establish its usefulness in serving as a proxy for a conventional sharpness metric, should ultimately be followed up by comparisons between the new metric and psychoacoustic sharpness data.

The method outlined in this letter thus enables sharpness calculations of sufficient fidelity for many applications using existing Moore and Glasberg loudness implementations (such as those complying with ANSI S3.4, 2007) by the insertion of only a few lines of code. We hope that this will thus enable more frequent and fruitful use of sharpness calculations in the pursuit of improved sound quality.

Acknowledgments

This work was funded in part through an SBIR grant with Blue Ridge Research and Consulting through Air Force Research Laboratory. We would like to thank the reviewers of this work for their feedback and input.

References and links

ANSI (**2007**). ANSI S3.4-2007. *Procedure for the Computation of Loudness of Steady Sounds* (American National Standards Institute, Washington, DC).

Aures, W. (**1985**). “Der sensorische Wohlklang als Funktion psychoakustischer Empfindungsgrößen” (“Sensory euphony as a function of auditory sensations”), Acustica **58**(5), 282–290.

DIN (**2009**). DIN 45692-2009. *Measurement Technique for the Simulation of the Auditory Sensation of Sharpness* (Deutsches Institut für Normung, Berlin).

Fastl, H., and Zwicker, E. (**2007**). *Psychoacoustics: Facts and Models*, 3rd ed. (Springer, New York).

Glasberg, B. R., and Moore, B. C. J. (**2006**). “Prediction of absolute thresholds and equal-loudness contours using a modified loudness model,” J. Acoust. Soc. Am. **120**(2), 585–588.

ISO (**2017**). ISO 532-1-2017. *Acoustics—Methods for Calculating Loudness—Part 1: Zwicker Method* (International Organization for Standardization, Geneva).

ISO (**2017**). ISO 532-2-2017. *Acoustics—Methods for Calculating Loudness—Part 2: Moore-Glasberg Method* (International Organization for Standardization, Geneva).

Moore, B. C. J., and Glasberg, B. R. (**2007**). “Modeling binaural loudness,” J. Acoust. Soc. Am. **121**(3), 1604–1612.

Moore, B. C. J., Glasberg, B. R., and Baer, T. (**1997**). “A model for the prediction of thresholds, loudness and partial loudness,” J. Audio Eng. Soc. **45**(4), 224–240.

Pedrielli, F., Carletti, E., and Casazza, C. (**2008**). “Just noticeable differences of loudness and sharpness for earth moving machines,” J. Acoust. Soc. Am. **123**(5), 3164.

Swift, S. H., and Gee, K. L. (**2017**). “Implementing sharpness using specific loudness calculated from the ‘Procedure for the Computation of Loudness of Steady Sounds,’ ” Proc. Mtgs. Acoust. **30**(1), 030001.

Vettel, J. M. (**2009**). “Neural integration of multimodal events,” Ph.D. dissertation, Brown University, Providence, RI. http://wiki.cnbc.cmu.edu/Real_World_Events (Last viewed October 24, 2017).

von Bismarck, G. (**1974**). “Sharpness as an attribute of the timbre of steady sounds,” Acustica **30**(3), 159–172.