Implementing sharpness using specific loudness calculated from the "Procedure for the Computation of Loudness of Steady Sounds"

S. Hales Swift and , and Kent L. Gee

Citation: Proc. Mtgs. Acoust. **30**, 030001 (2017); doi: 10.1121/2.0000542
View online: http://dx.doi.org/10.1121/2.0000542
View Table of Contents: http://asa.scitation.org/toc/pma/30/1
Published by the Acoustical Society of America

Proceedings of Meetings on Acoustics

Volume 30 http://acousticalsociety.org/

Acoustics `17 Boston

173rd Meeting of Acoustical Society of America and 8th Forum Acusticum
Boston, Massachusetts
25-29 June 2017

Engineering Acoustics: Paper 5pNSa9

Implementing sharpness using specific loudness calculated from the "Procedure for the Computation of Loudness of Steady Sounds"

S. Hales Swift and **Kent L. Gee**
Physics and Astronomy, Brigham Young University, Provo, UT, 84602; hales.swift@gmail.com; kentgee@byu.edu

Sharpness as defined in DIN 45692 and in Fastl and Zwicker [2007] is a measure of the scaled center of frequency of a sound, i.e., sounds with relatively greater high-frequency content have a greater sharpness than those with proportionally greater low-frequency content. The sharpness percept also influences the overall "pleasantness" of a sound and is thus important to sound acceptability. Measures of sharpness have been predominantly constructed to work with Zwicker's loudness model [DIN 45631/ ISO 532B], with less attention to the standardized model given in "Procedure for the Computation of Loudness of Steady Sounds" [ANSI S3.4-2007]. In this paper, a mathematical method for adapting the standardized sharpness measure for use with the ANSI S3.4-2007 loudness standard (and other loudness metrics producing a specific loudness distribution) is discussed. Benchmarking results for the resultant ANSI-based metric implementations and potential limitations of this approach are also addressed.

Published by the Acoustical Society of America

 [DOI: 10.1121/2.0000542]

1. INTRODUCTION

A significant goal of research in psychoacoustics and sound quality is to understand and develop means to accurately predict human responses to sound stimuli. As researchers have sought insight into the variety of human responses to sound, they have identified a number of distinct perceptual dimensions. These "percepts" gain added importance when their presence or absence significantly affects a listeners comfort or well-being, and especially when their presence intrudes upon or interferes with a desired state or activity.

Fastl and Zwicker,[1] following Aures,[2] identify a number of perceptual dimensions that can be used to predict the relative sensory pleasantness (or unpleasantness) of a sound. These percepts include loudness, sharpness, roughness, fluctuation and tonalness. Of these percepts, two—loudness and sharpness—have standardized calculation procedures, with a third—roughness—being considered for possible standardization.[3] These standards have typically utilized some form of Zwicker's loudness metric[4] as an essential input to the algorithms or formulae used in their calculation.

Zwicker's loudness metric is one of two main methods of calculating loudness that have been made into national or international standards. Its stationary form was standardized as DIN 45631:1991[6] and in its time-varying form in DIN 45631/A1.[7] An international standard is also available, ISO 532-1:2017,[4] delineating its calculation. In the United States, ANSI S3.4-2007[8] standardized a stationary loudness method based on Moore and Glasberg's model and it has also received international standardization as ISO 532-2:2017.[5] A method for extending this model to predict the loudness of simple time-varying sounds was published by Glasberg and Moore in 2002.[9]

Although Zwicker's loudness metric is used as the basis for the standardized model of sharpness, there is some advantage to be gained in being able to calculate it accurately outside of the Zwicker's loudness platform, e.g., as an appendage to a metric implementing ANSI S3.4-2007[8] or the time-varying loudness metric related to it.[9] Both the American and European standards have strengths and weaknesses. The denser auditory filter spacing in ANSI S3.4-2007, for example, helps avoid the degree of ripple seen in the calculation of loudness and sharpness when using the Zwicker metric for tones while Zwicker's metric is more computationally efficient.

In this paper, we derive and demonstrate an approach to making sharpness predictions using the outputs from ANSI S3.4-2007 or other similarly constructed loudness metrics. This will ideally enable sharpness predictions in contexts where they have not previously been available. Several ANSI-based sharpness implementations encompassing differing assumption sets are developed and tested against the validation data from DIN 45692[10] and the sharpness data reported in Fastl and Zwicker.[1] Limitations of each assumption set and implementation are also addressed.

2. DERIVATION

Sharpness, as standardized in DIN 45692,[10] is designed for compatibility with Zwicker loudness. However, other similar loudness metrics could be used to provide the inputs to a sharpness calculation. If two loudness metrics each give accurate predictions of loudness for all sounds and, as a step in the prediction process, give an accurate representation of specific loudness (generalized as loudness per frequency-like unit) then a transformation exists between the two metrics so that any metrics that can be calculated with the outputs of one could be calculated using the outputs of the other by appropriately transforming the variables. In order to implement sharpness, then, using the outputs of ANSI S3.4-2007, a transformation of variables is required. We begin with the formula for sharpness using input from Zwicker loudness.[1,10] C_1 is a calibration constant, N' is the specific loudness and z_1 is the psychoacoustic frequency in Bark

$$S = C_1(1000Hz, 60dB)\frac{\int N_1'(z_1)g(z_1)z_1 dz_1}{\int N_1'(z_1)dz_1}\text{[acum]}. \tag{1}$$

We apply the expected calculus procedure for a change of variables with $z_1 = z_1(z_2)$,

$$S = C_1(1000Hz, 60dB)\frac{\int \overbrace{N_1'(z_1(z_2))\frac{dz_1}{dz_2}}^{N_2'} g(z_1(z_2))z_1(z_2)dz_2}{\int \underbrace{N_1'(z_1(z_2))\frac{dz_1}{dz_2}}_{N_2'} dz_2}\text{[acum]}. \tag{2}$$

Given that the value of the bottom integral is the loudness, and given the assumption that both metrics predict loudness accurately, then the integrand in the bottom equation must be the specific loudness of the second loudness model. The quantity indicated in underbraces and overbraces can thus be identified as N_2', the specific loudness distribution in the second loudness calculation scheme. To the extent that both 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

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

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

$$S = C_1(1000Hz, 60dB)\frac{\int N_2'(z_2)g(z_1(z_2))z_1(z_2)dz_2}{\int N_2'(z_2)dz_2}\text{[acum]}. \tag{4}$$

This is, remarkably, of the same form as the original sharpness formula, with the only substantial change being that $g(z)$ and z must be evaluated using the original frequency-like variable (z_1). This means that, for any loudness metric that includes a true specific loudness distribution as an intermediate step the calculation of sharpness can be implemented with relative ease.

ANSI S3.4-2007 contains a step in which the specific loudness distribution is calculated in each 1/10th equivalent rectangular bandwidth (ERB, functionally similar to the Bark scale of psychoacoustic frequency in Zwicker's method).[8] The frequency-like variable in the ANSI standard (ERB) can be related directly to the center frequency. This formula, adapted from ANSI S3.4-2007, gives the relationship between ERB number and center frequency

$$f_c = \frac{10^{\frac{ERB\#}{21.366}} - 1}{0.004368}. \tag{5}$$

The center frequency can then be converted to critical band rate (Bark) using[1]

$$z = 13\tan^{-1}(0.00076 f_c) + 3.5\tan^{-1}\left(\left(\frac{f_c}{7500}\right)^2\right). \tag{6}$$

The expression for the critical band rate can then serve as input to the weighting functions. There are three weighting functions given in the German standard:[10]

1. The standard weighting curve
2. The weighting curve of Aures, which is level dependent
3. The weighting curve of von Bismarck.

In what seems a fortuitous development, these can be conveniently abbreviated as g, g_A and g_B. The standard expression for the weighting function is given by

$$g(z) = \begin{cases} 1 & z \leq 15.8\text{Bark} \\ 0.15e^{0.42(\frac{z}{Bark}-15.8)} + 0.85 & z > 15.8\text{Bark} \end{cases}, \tag{7}$$

while that of Aures is given by

$$g_A(z) = 0.78\frac{e^{\frac{0.171z}{Bark}}}{\frac{z}{Bark}}\frac{N/sone}{\ln\left(0.05\frac{N}{sone}+1\right)}, \tag{8}$$

and the weighting function of Von Bismarck is given by

$$g_B(z) = \begin{cases} 1 & z \leq 15\text{Bark} \\ 0.2e^{0.308(\frac{z}{Bark}-15)} + 0.8 & z > 15\text{Bark} \end{cases}. \tag{9}$$

3. LIMITATIONS

The approach that is pursued in this paper is, of course, only as good as its underlying assumptions. Violation of these assumptions will limit the degree of agreement between the original and translated metric. Several varieties of differences exist between the ISO 532 and ANSI S3.4-2007. Their loudness outputs are not precisely the same and thus their predicted specific loudness values must differ by some amount. To what extent does this pose a problem?

For narrowband signals, one can expect good agreement in predicted sharpness even if the two methods return discrepant total loudness provided that the distributions are both centered similarly when evaluated in Bark. This will occur provided that the two methods give similar predictions for the upward and downward spread of excitation and thus yield similarly centered specific loudness distributions. This is especially the case for lower frequencies where the sharpness weighting function $g(z)$ is constant. However, at higher frequencies, even relatively small differences in the upward spread of excitation (or masking) could be amplified by the exponentially increasing $g(z)$ and lead to significant discrepancies. Modifying $g(z)$ can account for these differences. $g(z)$ is fundamentally a scaling between psychoacoustic frequency and sharpness growth and so it can reasonably be altered in order to compensate for differences between the two models that marginally distort the centering in psychoacoustic frequency of the specific loudness distribution. A non-Zwicker loudness metric with a customized $g(z)$ should thus be able to meet the narrowband response requirements of DIN 45692 even if its total loudness predictions differed from Zwicker.

Differences in global loudness predictions might pose a greater difficulty if Aures's sharpness weighting is used. Because Aures sharpness includes the loudness as a dependent variable in the calculation of the weighting function, similar predictions using Aures's method rely more heavily on the validity of the equal global loudness prediction assumption than the other sharpness calculation methods. The general trends and relative sharpness relationships predicted by a metric using non-Zwicker inputs would still be expected to behave correctly, but the difference in loudness predictions input into Aures's weighting function could result in differing absolute values.

For broadband signals, or generally for any signal that has widely spaced components, the assumption of equal loudness prediction for all signals is more important. If this assumption were violated, and the two methods had opposite differing behavior for, e.g., high- and low-frequency components, then it is likely that the global sharpness estimate would differ between the two models due to the differing specific loudness distributions evoked by high- and low-frequency components in the two schemes. An approach to addressing this contingency is discussed in Sec. 4.3.

An additional limitation is posed by the high-frequency cutoff behavior of the two methods. Both have limited modeling above 13-15 kHz. However, because the hearing of individuals is so varied above these frequencies in response to age, noise exposure and inter-individual differences, predictions in this frequency range are of limited applicability.

4. IMPLEMENTATION

A. IMPLEMENTATION VIA DIRECT TRANSLATION

A sharpness metric was implemented using inputs from ANSI S3.4-2007 as described in the derivation in Sec. 2, by directly implementing Eq. (4) with the original $g(z)$ function indicated in the DIN 45692 standard. This implementation was tested using the signals given in DIN 45692 and the associated values given in that standard as shown in Fig. 1. The results track quite well for the narrowband signals until the g function begins to exceed 1 (which occurs at 15.8 Bark). For the broadband signals the slope is right for those signals with lower cutoff frequencies (f_{lc}) below about 16.5 Bark, but the sharpness is underpredicted relative to the standard values for all but the highest lower cutoff frequencies where the sharpness grows more quickly than the standard. The degree to which this implementation captures the trends may be sufficient for some engineering purposes. Many applications require relative rather than absolute values to guide judgments. Additionally, in absolute terms, the differences are relatively small. There are, nevertheless, noticeable differences, and this suggests that the ability of the most basic sharpness translation approach to replicate the standard broadband results is limited.

B. IMPLEMENTATION WITH CUSTOMIZED $g(z)$ FUNCTION

A second implementation considered the possibility that the two models differed in the degree of upward spread of excitation from the center frequency that they would predict. If the two models responded to the same sound by producing specific loudness distributions that were centered differently in psychoacoustic frequency (z or $ERB\#$ properly related to one another) then this could lead to a different predicted sharpness. This would lead to only small differences while $g(z)$ is one for lower values of z, but would lead to larger difficulties when $g(z)$ begins to rapidly increase in value. In the regime of rapid increase, minor differences in centering of the specific loudness distribution could result in outsize differences in predicted numerical values. This would suggest that there is a modest but potentially important interaction between the ideal shape of the

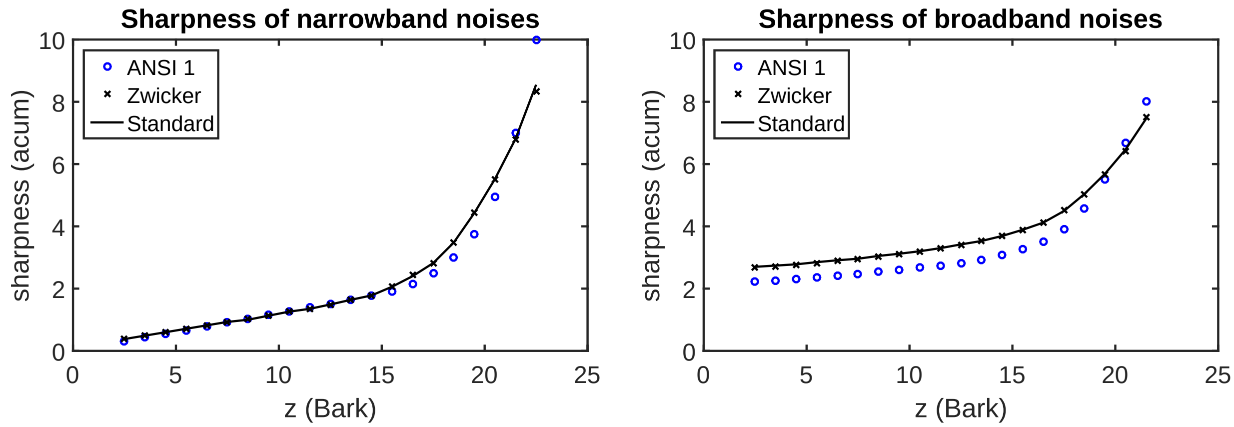

Figure 1: Sharpness of narrowband test sounds (left) and broadband test sounds (right) from DIN 45692 as dictated by the standard and as predicted by the Zwicker-based implementation of the sharpness metric and as predicted by the first (direct) ANSI-based implementations of the sharpness metric.

$g(z)$ function and the method being used to calculate the specific loudness distribution. In order to obtain such an optimized $g(z)$ function, the specific loudness patterns associated with the narrowband standard stimuli were calculated and the $g(z)$ function was iteratively modified in order to reduce the discrepancies.

This is done in several steps:

1. Standard narrowband validation sound sharpness is predicted by the unmodified metric

2. These sharpness values are divided by the standard values to get a ratio correction

3. Correction values are interpolated and used to modify the $g(z)$ function. The modified function is used to recalculate the sharpness and returned to step 2 iteratively.

This significantly improved the agreement for the narrowband sounds as can be seen in Fig. 2 and, indeed, the metric is able to pass the DIN 45692 requirement for the narrowband tests cases. This change in the weighting also led—although not a goal of the optimization—to reductions in the discrepancies between the metric predictions and the standard values for the broadband signals. Notwithstanding these improvements, the agreement was not sufficient to meet the requirements of the standard for broadband signals. The improvement is understandable because the adjustments to the $g(z)$ function led to increases in weighting in the range from around 15-20 Bark, which helps counteract the under-predictions of the most basic ANSI-based metric for broadband sounds with lower cutoff frequencies beneath this range. The continued discrepancy, on the other hand, points to the likelihood of differences in the loudness ascribed to low- and high-frequency components. In support of this hypothesis, a later attempt to optimize the g-function in order to pass both narrowband and broadband tests was unsuccessful and suggested that this approach was mathematically untenable: the balance of errors could be shifted between narrowband and broadband test signals to some extent, but they could not be mutually reduced sufficiently to pass the standard tests. Notwithstanding this limitation in satisfying the standard values, the version of the

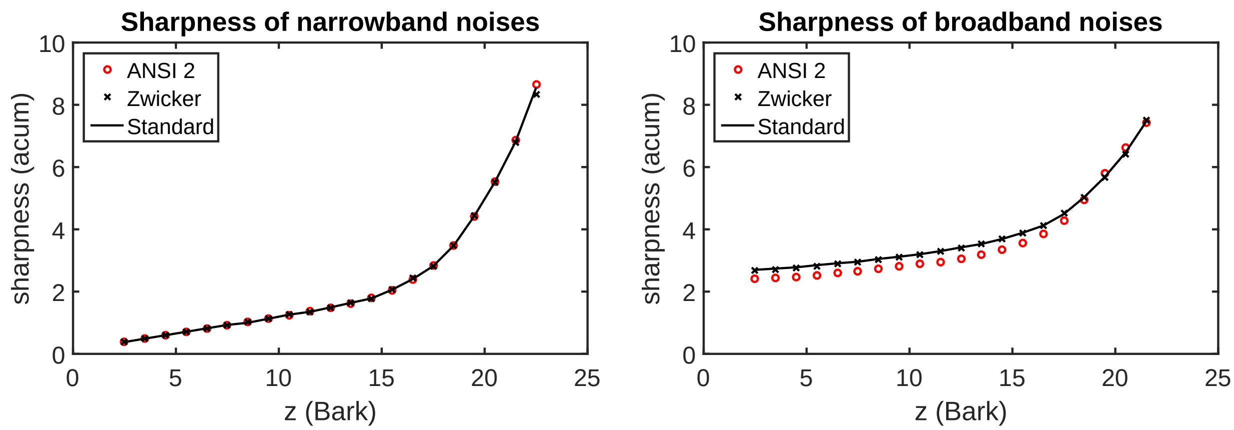

Figure 2: Sharpness of narrowband test sounds (left) and broadband test sounds (right) from DIN 45692 as dictated by the standard and as predicted by the Zwicker-based implementation of the sharpness metric and as predicted by the second (modified $g(z)$-function) ANSI-based implementations of the sharpness metric.

metric with the modified $g(z)$ function is accurate enough that it is likely to be useful for making sharpness predictions for many engineering applications. It also enjoys the additional benefit of the more densely spaced ERB filters in the ANSI loudness metric. In Fig. 3, the standard values of narrowband noises are plotted as a reference together with the ANSI and Zwicker metric responses to tones at 60 phon (ANSI) across a wide range of frequencies. The Zwicker-based metric, because of its use of a third-octave band filter-based calculation strategy, has significantly more ripple in its predictions than the ANSI-based metric, which uses some 389 filters in its excitation pattern calculations leading to the same densely-spaced number of specific loudness entries and relatively smooth consequent predictions.

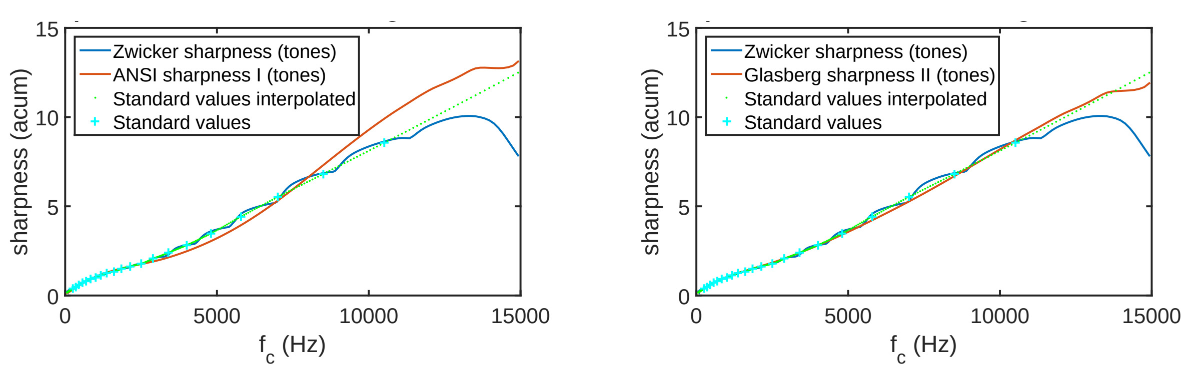

Figure 3: The response of the Zwicker and ANSI-based sharpness implementations to tones at each 1/10th ERB number increments together with the qualification values for narrowband noises from DIN 45692 and spline-interpolated values of the qualification values. Extrapolated sharpness values based on the standard are also shown. On the left is the simplest ANSI sharpness implementation, and on the right is the implementation with the modified g-function.

C. REVISED SHARPNESS METRIC WITH SPECIFIC LOUDNESS RE-BALANCING

To investigate whether systematic differences between the ANSI and Zwicker loudness predictions for components in different frequency regimes were responsible for the differences seen in the sharpness predictions, we made a preliminary examination of the loudness predicted for the test sounds. Fig. 4 shows significant differences in the loudness predicted by the ANSI and Zwicker methods for nominally 4 sone$_G$ test sounds. The ANSI S3.4-2007 method predicts loudness values that at times deviate by more than 25% from the nominal values and those predicted by the Zwicker metric. For broadband sounds, the ANSI metric predicts much larger loudness values for sounds with lower cutoff frequencies below 18 Bark. Additionally, an iterative procedure was

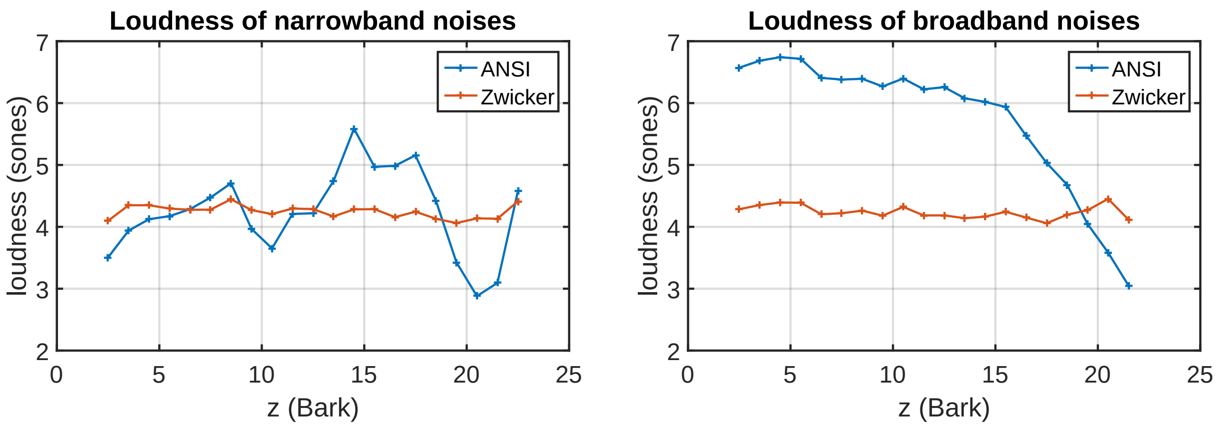

Figure 4: The calculated loudness in sones for the narrowband and broadband test signals in DIN 45692 using ANSI S3.4-2007 and Zwicker loudness (ISO 532-1:2017, stationary).

used to find 60 phon equal loudness contours for each method as shown in Fig. 5. Here, again, significant differences are seen at frequencies below 400 Hz as well as between 6000 and 10000 Hz, and the magnitudes of these differences seem adequate to cause the discrepancies in sharpness calculation that have been identified. As a starting point to dealing with systematic differences, the effects of the violated assumption of equal loudness predictions between the metrics must be taken into account. In order to do this, a second weighting function, g_s is introduced as seen in Eq. (11), which takes into account violations of the assumption that the two metrics will produce functionally equivalent expressions of transformed specific loudness or loudness density

$$N_2'(z_2) = N_1'(z_1(z_2))\frac{dz_1}{dz_2}\text{[sones/frequency-like unit]}. \tag{10}$$

The expression g_s can then be thought of as the typical ratio difference between the specific loudness predicted by the two metrics for medium level sounds in each frequency band after accounting for the differing frequency-spacing scales. Unlike the $g(z)$ function used in the regular implementation of the metric, g_s enters into the integrals in both the numerator and the denominator of the sharpness expression

$$S = C_1(1000Hz, 60dB)\frac{\int N_2'(z_2)g(z_1(z_2))g_s(z_1(z_2))z_1(z_2)dz_2}{\int N_2'(z_2)g_s(z_1(z_2))dz_2}\text{[acum]}. \tag{11}$$

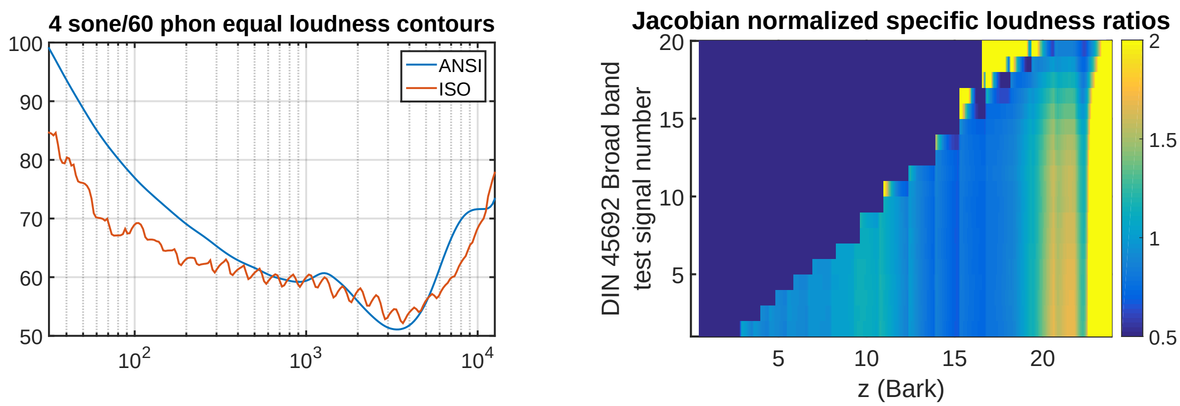

Figure 5: (Left) The 60-phon or 4-sone equal loudness contour for tones produced using ANSI S3.4-2007 and ISO 532. (Right) Specific loudness predicted by ISO 532 metric divided by specific loudness predicted by ANSI S3.4-2007 after adjusting for frequency bin width differences.

One result of this is that it has relatively little impact on narrowband sounds because the specific loudness distribution tends to fall in a more compact region and thus sees smaller variation in g_s than would a broadband sound. Thus, for narrowband sounds, both the top and the bottom are multiplied by an equal amount with small variation leading to approximately the unmodified result. For broadband signals, on the other hand, high and low frequencies are modified relatively independently because g_s tends to vary more over a wider frequency range. In order to determine g_s, the specific loudness pattern due to each of the broadband sounds was calculated using both the Zwicker and the ANSI S3.4-2007 procedures. After a Jacobian transformation to account for differing frequency spacings as in eq. (10), the two specific loudness distributions (N') were divided (ISO 532-1:2017 by ANSI S3.4-2007) and the results plotted in Fig. 5. The specific loudness ratio for a particular Bark value is seen to be approximately equal across most of the test sounds.

Accordingly, the ratios were plotted in Fig. 6 with a particular focus on the first 13 test signals as seen in the right panel of this figure. These first 13 signals were used to create the re-weighting function g_s, indicated as black x's in the right panel as a function of z in Bark. Using the newly determined g_s function makes it possible to, after optimizing the regular $g(z)$ function for the narrowband data, satisfy the requirements in DIN 45692 for narrowband and broadband signals. In Fig. 7, the 5% relative difference and the 0.05 acum absolute difference limit values are shown for each of the narrowband and broadband validation signals. The standard requires each test sound to fit within at least one of the two thresholds (absolute or relative difference).

The resultant fit can be seen in Fig. 8. There is very good agreement between the sharpness predictions of the rebalanced ANSI-based metric and the standard values. Although it has not been tested more broadly, this approach meets the requirements for responses to the standard test cases and is likely sufficient to ensure that the resultant metric is able to fill many engineering purposes and serve in many of the same capacities that a normally constructed Zwicker-based sharpness metric might serve, while also enjoying the benefit of less ripple for sinusoidal signals. At the least, we would expect very similar results for narrowband signals and broadband signals of middling level around 4 sones$_G$. Provided that the growth of loudness (and specific loudness) with level is similar between the two models, similar results would be expected more generally.

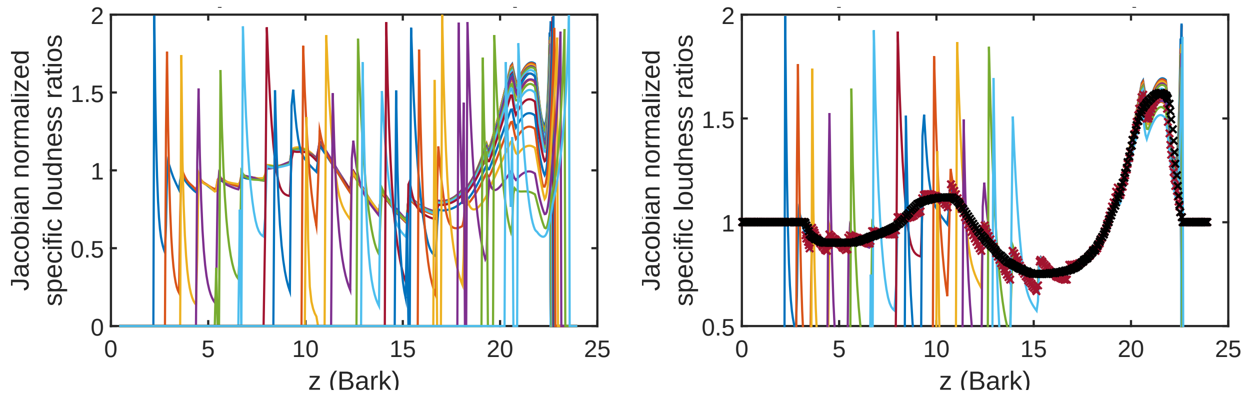

Figure 6: (Left) Jacobian adjusted N' of ISO 532 metric divided by N' of ANSI S3.4-2007 for all broadband test signals. (Right) Same as left but only the first 13 broadband test signals. (Black x's) Ratio correction accounting for N' prediction differences between the two methods.

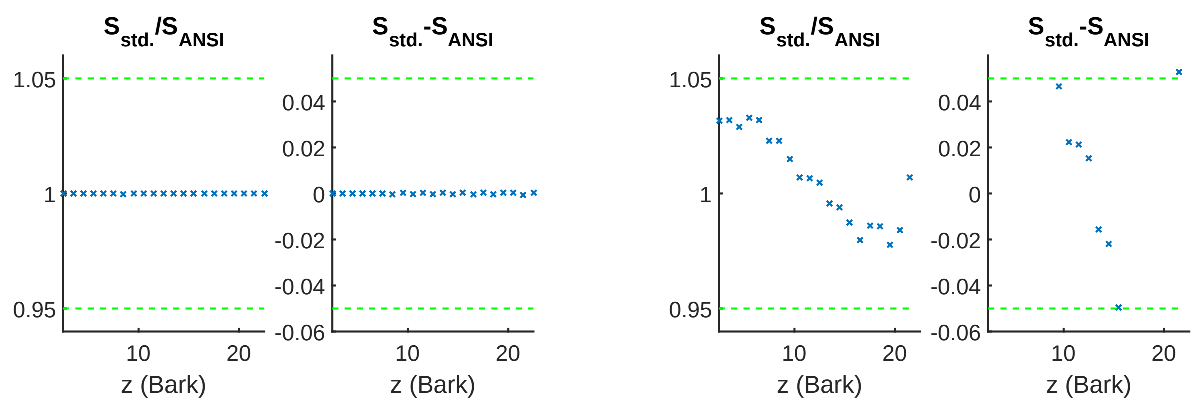

Figure 7: Error of the ANSI-based metric relative to the DIN 45692 narrowband test sounds (left) and broadband test sounds (right).

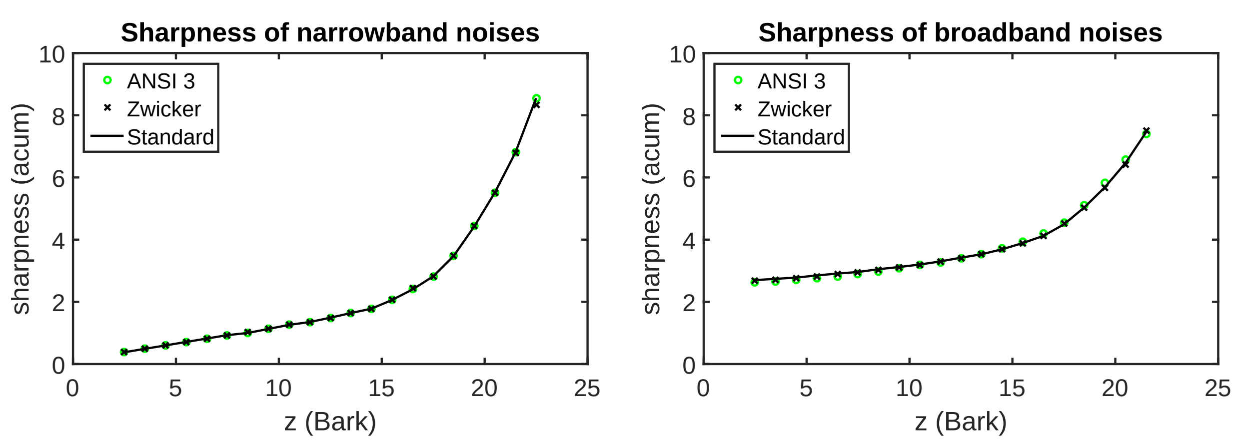

Figure 8: Sharpness of narrowband (left) and broadband (right) test sounds from DIN 45692 as dictated by the standard and as predicted by the Zwicker-based implementation of the sharpness metric and as predicted by the third ANSI-based implementations of the sharpness metric.

5. STANDARD VALIDATION OR COMPARISON WITH PSYCHOACOUSTIC DATA?

We have shown that an implementation of sharpness that includes the modification in Eq. (11) can meet the signal test response requirements given in the DIN 45692 sharpness standard. However, while clearly based on psychoacoustic data, the standard values appear to be chosen to assure that a correct implementation of the method given in the standard has been accomplished. As such, they match the outputs of the Zwicker metric to those signals closely. Other data, such as those found in Fastl and Zwicker Fig. 9.1, which shows the sharpness of narrowband noises at 60 phon, can also be profitably compared in order to evaluate the performance of the various ANSI-based sharpness implementations relative to a Zwicker-based implementation of sharpness. The sharpness of all the narrowband test sounds is shown in Fig. 9. Consistent with the presentation in Fastl and Zwicker, log-y linear-z axes are used, which helps clarify the trends, especially for lower values of z where the sharpness is small. Examining this figure, it is clear that the unmodified implementation of sharpness using inputs from the ANSI metric matches the sharpness at low frequencies better than the Zwicker implementation or any of the modified ANSI-based implementations. Furthermore, above around 15 Bark the unmodified ANSI-based metric again outperforms all of the other implementations tested. It is worth bearing in mind that this is the case despite the fact that the g-function is originally designed for Zwicker compatibility and may not be optimal for ANSI-based inputs; i.e., if the two metrics make marginally different predictions regarding the upward spread of masking then this might reasonably dictate minor modifications to the g-function. In order to quantify the degree of agreement between the several models, the standard, and the data of Fastl and Zwicker, we report correlation and RMS error values for each. For narrowband signals, the unmodified ANSI-based metric provides the best performance for matching the data reported in Fastl and Zwicker,[1] but it is also the worst for the broadband test sounds. The third ANSI-based metric best matches the data reported in Fastl and Zwicker for the broadband sounds in terms of RMS error but similarly also shows the worst match for their narrowband data. The

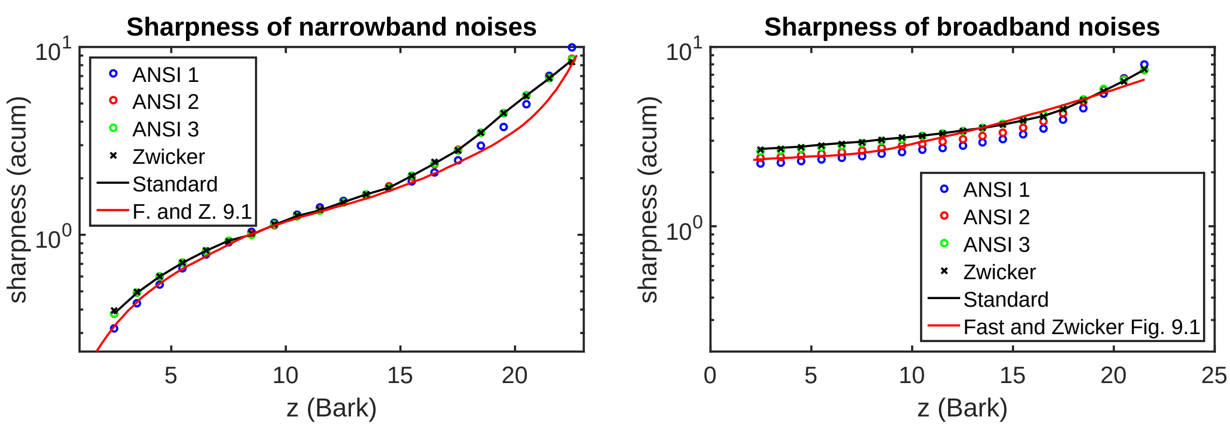

Figure 9: Response of each of the metric implementations to the narrowband (left) and broadband (right) test sounds from DIN 45692 compared with both the standard values and Fastl and Zwicker Fig. 9.1,[1] which reports subjective sharpness values.

Table 1: Correlation coefficient values for the ANSI- and Zwicker- based metrics' predictions of the sharpness of narrowband noises (left) and broadband noises (right) relative to the standard values and the values reported in Fastl and Zwicker.[1] RMS values in each case are evaluated for the first 20 narrowband and broadband signals. Minimum RMS values are underlined.

	Narrowband test sounds				Broadband test sounds			
Metric	St.ρ	Fastlρ	St. RMS	Fastl RMS	St.ρ	Fastlρ	St. RMS	Fastl RMS
ANSI 1	0.9872	0.9968	0.3310	<u>0.3556</u>	0.9945	0.9492	0.6270	0.6151
ANSI 2	1.0000	0.9865	0.1135	0.4775	0.9986	0.9727	0.4193	0.3833
ANSI 3	1.0000	0.9856	<u>0.0002</u>	0.5798	0.9991	0.9836	0.0705	<u>0.3093</u>
Zwicker	0.9999	0.9832	0.0136	0.5756	0.9999	0.9769	<u>0.0222</u>	0.3436

discrepancies seen in these sharpness predictions tend to be of lower magnitude than the differences seen in the loudness predictions between the two models. The RMS errors for all of the models relative to the Fastl and Zwicker data tends to be around half an acum. The RMS errors seen in the ANSI-based metrics relative to the standard values are of about the same magnitude as the RMS errors seen in the Zwicker-based metric relative to the Fastl and Zwicker values.[1] This confirms that all of these implementations will give sharpness predictions with inaccuracy limited to the approximate size of the differences between published datasets, suggesting that any of these metric implementations will give predictions of sharpness sufficiently accurate for many sound quality engineering applications.

6. CONCLUSION

A translated form of Zwicker sharpness was successfully implemented using input variables provided by the "Procedure for the Computation of Loudness of Steady Sounds" ANSI S3.4-2007. A similar approach should be valid for any loudness metric that includes as an intermediate step an accurate calculation of the specific loudness distribution.

Methods to account for the possibility of systematic differences in the upward spread of excitation and specific loudness as well as systematic differences in loudness predictions between metrics are also developed and discussed. Adapting the $g(z)$ function to account for possible differences between the upward spread of excitation and loudness predicted by the ANSI- and Zwicker-based metrics enabled this implementation to pass narrowband test sound requirements of DIN 45692. Achieving results in keeping with this standard for broadband noises required adding a second weighting function which took into account differences in the predicted specific loudness distributions between the two methods.

When data from Fastl and Zwicker was used as the comparand, the unmodified ANSI-based metric best matched their reported subjective sharpness values for the narrowband noises. Predictions for broadband noises showed greater differences but still showed decent agreement.

The accuracy seen in the ANSI-based implementations seems to be sufficient for use in many engineering contexts, enabling the calculation of sharpness in contexts where it has not previously been an easily available option. The ANSI-based metrics have the additional benefit of less ripple due to their closer filter spacing. Future work should seek to produce an ANSI-based metric that is calibrated to a more general set of psychoacoustic data in order to optimally predict human perception of this psychoacoustic quality.

ACKNOWLEDGMENTS

This work was funded in part through an SBIR grant with Blue Ridge Research and Consulting through Air Force Research Laboratories.

REFERENCES

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

[2] W. Aures, "Der sensorische wohlklang als funktion psychoakustischer empfindungsgrssen (The sensory euphony as a function of auditory sensations)", *Acustica* **58**, 282-290 (1985).

[3] R. Sottek and Klaus Genuit, "Perception of roughness of time-variant sounds", *Proc. Meet. Acoust.* **19**, 050195???? (2013).

[4] "Acoustics—Methods for calculating loudness—Part 1: Zwicker method" ISO 532-1:2017 (International Organization for Standardization, 2017).

[5] "Acoustics—Methods for calculating loudness—Part 2: Moore-Glasberg method" ISO 532-2:2017 (International Organization for Standardization, 2017).

[6] "Berechnung des Lautstarkepegels and der Lautheit aus dem Geruschspektrum. Verfahren nach E. Zwicker (Calculation of loudness level and loudness from the sound spectrum - Zwicker method)" DIN 45631 (Deutsches Institut für Normung, 1990).

[7] "Berechnung des Lautstarkepegels and der Lautheit aus dem Geruschspektrum. Verfahren nach E. Zwicker - nderung 1: Berechnung der Lautheit zeitvarianter Gerusche (Calculation of loudness level and loudness from the sound spectrum - Zwicker method - Amendment 1: Calculation of the loudness of time-variant sound)" DIN 45631/A1 (Deutsches Institut für Normung, 2010).

[8] "Proceedure for the Computation of Loudness of Steady Sounds" ANSI S3.4-2007 (American National Standards Institute, Inc., 2007).

[9] B. R. Glasberg and B. C. J. Moore, "A model of loudness applicable to time-varying sound", *J. Audio Eng. Soc.* **50**, 331-342 (2002).

[10] "Measurement technique for the simulation of the auditory sensation of sharpness" DIN 45692 (Deutsches Institut für Normung, 2009).