Predicting the intensity JND from the loudness of tones and noise

Predicting the intensity JND from the loudness of tones and noise

Stephen T. Neely

Boys Town National Research Hospital
555 North 30th Street, Omaha, Nebraska/68131, USA

Jont B. Allen

AT&T Labs Research
600 Mountain Avenue, Murray Hill, New Jersey/07974, USA

1 Loudness

The *loudness* of a sound is its perceived intensity. A *just-noticeable difference* (JND) in intensity has a corresponding JND in loudness. We will interpret loudness L and the loudness JND ΔL as the first and second moments of a random decision variable, which we call the *single-trial loudness* $\tilde{L}(I)$. This interpretation, combined with signal detection theory, allows us to relate the loudness of tones and noise to the corresponding intensity JND.

1.1 Loudness measurements for tones and noise

The loudness data of Fletcher and Munson (1933) for a 1-kHz tone are shown as a solid line in the left panel of Fig 1. The loudness data of Miller (1947) for wide-band noise (WBN) are represented by the circles and dashed line in the left panel of Fig. 1.

We define the log-loudness slope by

$$\nu \equiv \frac{d}{d\beta}\left[10\log_{10}(L)\right] = \frac{I}{L}\frac{dL}{dI} \tag{1}$$

where $\beta = 10\log_{10}(I)$. The log-loudness slope for a 1-kHz tone and WBN, computed from the loudness data, are shown in the right panel of Fig.1. The variable ν approximately describes the local power-law dependence of loudness on intensity, i.e. $L(I) \approx I^{\nu(I)}$.

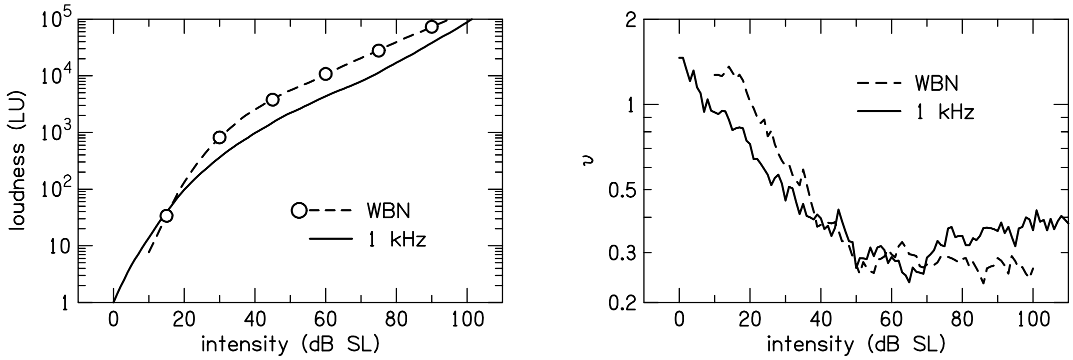

Figure 1. Loudness for tones and noise. The solid line in the left panel represents loudness data for a 1-kHz tone from Table III of Fletcher and Munson (1933), which is tabulated in 1-dB steps from -10 to 129 dB. One loudness unit (LU) equals 975 sones. The circles represent the loudness level for wide-band noise (WBN) measured by Miller (1947). The Fletcher and Munson loudness function was used to convert Miller's loudness-level data to loudness. The dashed line is a polynomial fit to the 6 data points. The right panel shows the log-loudness slope ν for the same loudness data.

1.2 Model assumptions

1. We assume that the *single-trial loudness* $\tilde{L}(I,t,f)$ is proportional to the total number of neural spikes that occur within a time-window of duration $\Im$ seconds

$$\tilde{L}(I,t,f) = \lambda \int_0^{x_L} \int_{t-\Im}^{t} \tilde{\Re}(I,t,f,x)\, dt\, dx \tag{2}$$

where λ is a constant with units of loudness and $\tilde{\Re}(I,t,f,x)$ is a random variable that describes the neural spike rate at time *t* associated with place *x* on the basilar membrane, for a tone of frequency *f* and intensity *I*. The length of the basilar membrane is x_L.

2. From signal detection theory, the relation between the loudness JND ΔL and the standard deviation of the single-trial loudness is $\Delta L = d'\sigma_L$, where d' is a constant.

3. From empirical observation, the loudness variance $\sigma_L^2 \equiv \mathrm{E}\left[\left(\tilde{\mathrm{L}} - \mathrm{L}\right)^2\right]$ is the sum of two components $\sigma_L^2 = \sigma_r^2 + \sigma_a^2$ where σ_r^2 is the variance associated with spike rate and σ_a^2 is the variance associated with something other than spike rate, perhaps spike amplitude.

4. We assume that the dominant component of the loudness variance at low levels is due to randomness in the rate of neural spike generation that causes the spike count to be Poisson distributed. The rate variance is $\sigma_r^2 = \lambda\, L r(L)$, where $r(L)$ is the variance-to-mean ratio (sometimes called the Fano factor) of the underlying spike counting process. If the spike counts are Poisson distributed, then $r(L) \equiv 1$; however, synchrony and refractoriness may cause $r(L)$ to be less than one at moderate levels (Lowen and Teich, 1996, Fig. 3).

5. We assume that the dominant component of the loudness variance at high levels is $\sigma_a^2 = L^2/a^2$. We suggest that this variance component may be due to randomness in the amplitude of neural spikes; however, we only know, based on empirical observation,

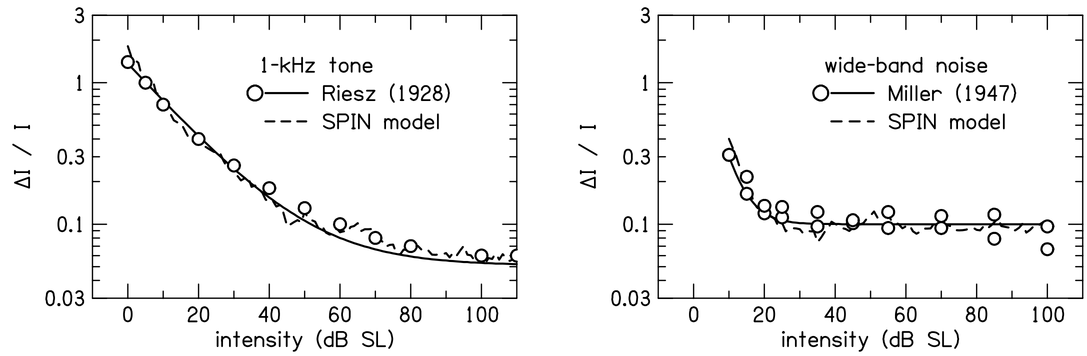

Figure 2. Intensity JND for tones and noise. The circles in the lower two panels represent the measurements of the relative intensity JND. The data for a 1-kHz tone in the lower-left panel are from Reisz (1928). The data for wide-band noise are from Miller (1947). The solid lines represent the original authors' fit to their data. The dashed lines were computed from the loudness data using Eq. (5).

that the constant a describes the maximum signal-to-noise (amplitude) ratio that is achievable by the auditory system.

6. Because JNDs are relatively small, the loudness-growth function $L(I,f)$ has a slope dL/dI that is a good local approximation to the ratio of the loudness JND to the intensity JND $\Delta L/\Delta I \approx dL/dI$.

Together, these assumptions define our model of loudness (Allen and Neely, 1997), which is an extension of the work of Hellman and Hellman (1990). Because the Poisson internal noise is only significant at lower levels (less than 40 dB SL) in this model, we refer to it as the saturated-Poisson-internal-noise or SPIN model. In the next section, we compare measurements of the loudness JND with SPIN model predictions.

2 Just-Noticeable Difference

Relative intensity JND measurements ($\Delta I/I$) are shown in Fig. 2 for a 1-kHz tone (Reisz, 1928) and for wide-band noise (Miller, 1947). From these intensity-JND data and the loudness functions shown in Fig. 1, we can derive estimates of the loudness JND for tones and noise. The loudness JND ΔL is defined in terms of the intensity JND ΔI as $\Delta L \equiv L(I+\Delta I) - L(I)$. Using this definition, we have computed $\Delta L(L)$ for tones from the data of Riesz (1928), Fletcher and Munson (1933), and for noise from the data of Miller (1947). The results are shown by the dashed lines in Fig 3.

Combining assumptions 3, 4, and 5 from section 2.2, we can express ΔL as a function of L ,

$$\Delta L = d'\left[\lambda\, L\, r(L) + \frac{L^2}{a^2}\right]^{\frac{1}{2}} \tag{3}$$

where d' , $r(L)$, and a are dimensionless and λ has units of loudness. In fitting Eq. (3) to the 1-kHz tone data, we found it necessary to let the Fano factor $r(L)$ become less than 1 at moderate levels. We did this by letting $r(L) = \left[1 + L^2/L_r^2\right]^{-\frac{1}{2}}$ for tones. No such decrease in $r(L)$ was needed to fit the wide-band noise data. The dashed lines in Fig. 3 show our fit to the derived $\Delta L(L)$ data for a 1-kHz tone and for wide-band noise. In the

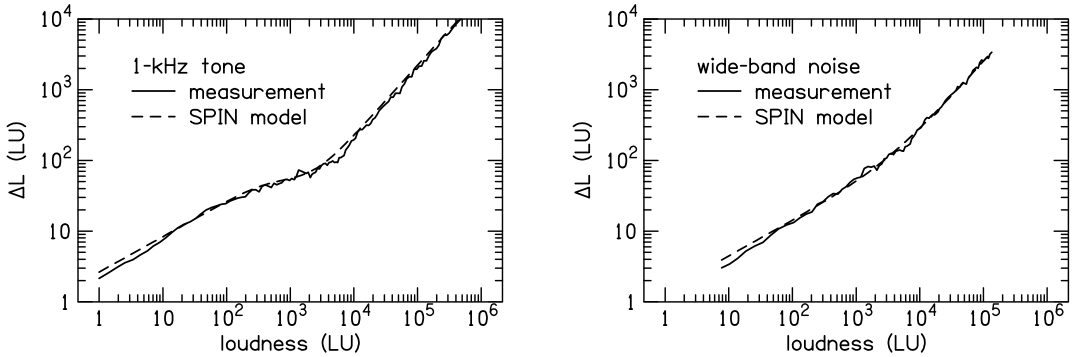

Figure 3. Loudness JND as a function of loudness. The solid lines are empirically derived from measurements of loudness and intensity JNDs using the definition of ΔL given in the text. In order to obtain continuous functions for ΔL, we used the original authors' fit to their $\Delta I / I$ data, instead of actual data values. In the left panel, the 1-kHz data are from Reisz (1928) and Fletcher and Munson (1933). In the right panel, the wide-band noise data are from Miller (1949). The dashed lines in both panels represent our fit to the data using Eq.(3) and the form of r(L) given in the text.. When we set d' =1, the parameter values were: for the tone case a=45, λ =7 LU, L_r =400 LU; and for the noise case a=40, λ =2 LU, $L_r = \infty$.

same way, the ΔL data for tones of other frequencies can also be fitted. The good agreement between the dashed lines and the solid lines provides support for the SPIN model.

Using Eq. (1) and assumption 6, we can express the intensity JND as

$$\frac{\Delta I}{I} = \frac{1}{\nu} \frac{\Delta L}{L} \tag{4}$$

According to the assumptions of the SPIN model, the loudness JND $\Delta L / L$ is determined by the statistics of the neural spike generation process. Our physical interpretation of the log-loudness slope ν is that it represents the non-linear transformation from sound pressure to neural spike rate that occurs within the cochlea. Thus, Eq. (4) appears to separate the influence of non-linear cochlear mechanics from the neural, signal-detection task. Another important interpretation of Eq. (4) is that it relates the signal-to-noise ratio of the psychophysical detection task L / σ_L with a physically measurable signal-to-noise ratio. The noise associated with the psychophysical detection task is $\sigma_L = \Delta L / d'$. If we define the loudness signal-to-noise ratio as $SNR_L = d' L / \Delta L$ and the physical correlate of this internal signal-to-noise ratio as $SNR_I = d' I / \Delta I$, then Eq. (4) tells us that $\nu = SNR_I / SNR_L$. In Fig. 2, we see that $\Delta L / L$ reaches a minimum value of about 0.2, corresponding to a maximum SNR_L of $50\, d'$ high levels.

Substituting in Eq. (4) for ΔL from assumption 2 and Eq. (3), gives

$$\frac{\Delta I}{I} = \frac{d' \sigma_L}{\nu L} = \frac{d'}{\nu} \left[\frac{\lambda\, r(L)}{L} + \frac{1}{a^2} \right]^{\frac{1}{2}}. \tag{5}$$

Eq. (5) uses the SPIN model to express the intensity JND in terms of loudness L and the log-loudness slope ν. The results of applying Eq. (5) to the loudness data are shown in Fig. 2 along with the measured intensity JND. The agreement between the SPIN model and the measured data is essentially the same as in Fig 3.

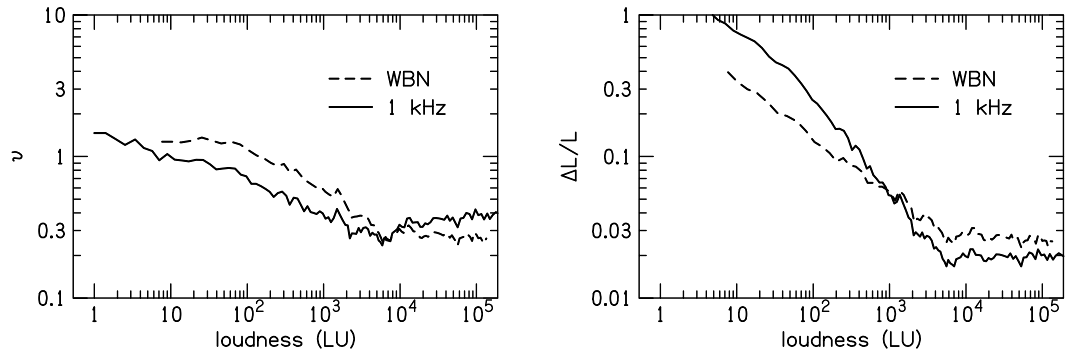

Figure 4. The log-loudness slope and the loudness JND as a function of loudness. Weber's law requires similar slopes in the left and right panels. In the right panel of Fig. 4, we see that that $\Delta L/L$ reaches a minimum value of about 0.02 at high levels; thus, according to the definition of SNR_L in the text, the maximum signal-to-noise ratio *for loudness appears to be about* $50d'$.

3 Weber's law

Weber's law states that $\Delta I/I$ is constant over a range of intensities. We see in Fig. 2 that Weber's law holds for wide-band noise above 20 dB SL, but not for a 1-kHz tone. The observation that Weber's law does not hold for tones is often called the "near-miss" to Weber's law.

In Eq. (4), we see that Weber's law is equivalent to the statement that $\Delta L/L$ is proportional to ν . We associate $\Delta L/L$ with the neural spike generation process and ν with non-linear cochlear mechanics. Weber's law requires that any variation in $\Delta L/L$ with level be matched by a corresponding variation in ν .

In order to investigate the "near-miss" for tones, $\Delta L/L$ and ν are plotted on similar scales as a function of L in Fig. 4. The similar slopes of the dashed lines in the right and left panels indicates that Weber's law holds for WBN. On the other hand, the differing slopes of the solid lines between the right and left panels indicates that Weber's law does not hold for a 1-kHz tone. At moderate levels (near 1000 LU) the solid line in the right panels differs in slope from the other 3 lines shown in Fig 4. This observation suggests that the "near miss" at moderate levels is due to the unusually steep slope of the loudness JND for tones. According to our SPIN model, the reason for this steeper slope is that the Fano factor $r(L)$ becomes less than 1 at moderate levels. This suggests that the "near miss" may be due to synchrony and/or refractoriness in the neural spike generation for tones of moderate level. This interpretation differs from, but does not necessarily exclude, the interpretation of the "near miss" as resulting from the integration of partial loudness over a range of firing rates which has been suggested by Allen and Neely, 1997.

To further investigate the different relation between $\Delta L/L$ and ν for tones and noise, $\Delta L/L$ is plotted as a function of ν in left panel of Fig. 5. The observation that the dashed line has a slope of 1 (below 20 dB SL), when plotted on log-log co-ordinates, is an indication that $\Delta L/L$ is proportional to ν and Weber's law holds for WBN. The observation that the solid line has a slope of 3 is an unexpected and intriguing result. Comparison of the solid line with the light guide line suggests that $\Delta L/L$ is approxi-

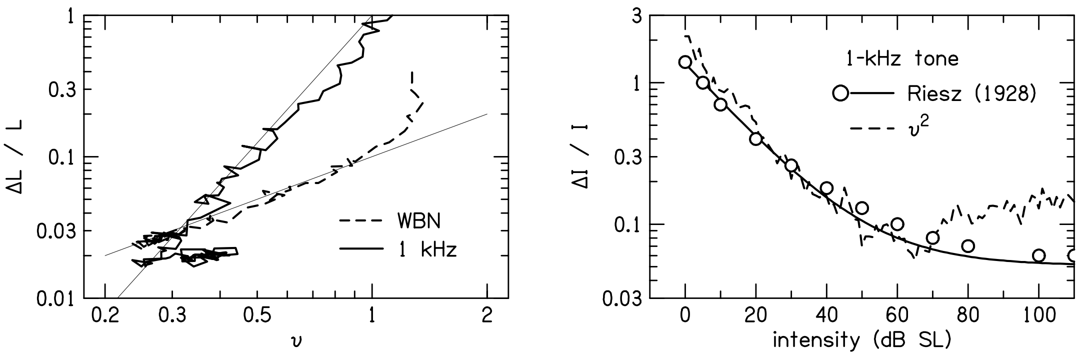

Figure 5. The loudness JND as a function of the log-loudness slope and the intensity JND as a function of intensity. A slope of 1 in the left panel indicates $\Delta I / I$ is constant. The light lines in the left panel represent 0.1ν *and* ν^3 *. The intensity JND data are compared with* ν^2 *in the right panel.*

mately equal to ν^3 for a 1-kHz tone below 60 dB SL. This relation is clearly inconsistent with Weber's law and not predicted by the SPIN model.

We can use the empirical observation that $\Delta L / L = \nu^3$ for a 1-kHz tone below 60 dB SL together with Eq. (4) to obtain a new relation between the intensity JND and the log-loudness slope $\Delta I / I = \nu^2$. The result of applying this equation to the 1-kHz loudness data is shown in the right panel of Fig. 5 along with the measured intensity JND. The agreement between $\Delta I / I$ and ν^2 is surprisingly good. Apparently, the term "near miss" underestimates the significance of the deviation from Weber's law observed for tones.

The empirical observation that $\Delta I / I = \nu^2$ places an additional constraint on the growth of loudness which complements the constraints imposed by the SPIN model. Combining this new result with Eqs. (1) and (5), leads to the conclusion that $L(I) = \left[1 + k\beta\right]^6$ (LU) for a low-level, 1-kHz tone, where $\beta = 10\log_{10}(I)$ and $k \approx 1/19$. In fact, this formula for $L(I)$ fits the Fletcher and Munson (1933) loudness data over a wider range of intensities than expected (0 to 100 dB SL) and may be a useful alternative to traditional power-law formulae for the dependence of loudness on intensity.

4 Conclusions

The agreement between the SPIN model and the measured JND data in Figs. 2 and 3 supports the interpretation that loudness and the loudness JND are the first and second moments of a random variable (called the *single-trial loudness*) associated with the neural spike generation process. At low levels, the variance-to-mean ratio is consistent with a Poisson spike generation process for both tones and noise. The observed deviation for a 1-kHz tone at moderate levels of the variance-to-mean ratio from that of a Poisson process is attributed to synchrony and/or refractoriness in the spike generation process and may be related to the so-called "near miss to Weber's law". At high levels, "internal noise" apparently limits the maximum signal-to-noise amplitude ratio achievable by the auditory system to about 50.

5 Acknowledgements

We thank Walt Jesteadt and Huanping Dai for helpful comments.

6 References

Allen, J.B. and Neely, S.T. (1997). “The relation between the intensity JND and loudness for pure tones and wide-band noise,” J. Acoust. Soc. Am. (submitted).

Fletcher, H. and Munson, W. (1933). “Loudness, its definition, measurement, and calculation,” J. Acoust. Soc. Am. **5**, 82-108.

Hellman, W. and Hellman, R. (1990). “Intensity discrimination as the driving force for loudness. Application to tones in quiet,” J. Acoust. Soc. Am. **68**, 807-813.

Lowen, S. B. and Teich, M. C. (1996). “The periodogram and Allan variance reveal fractal exponents greater than unity in auditory nerve spike trains,” J. Acoust. Soc. Am. **99**, 3585-3591.

Miller, G.A. (1947). “Sensitivity to changes in the intensity of white noise and its relation to masking and loudness,” J. Acoust. Soc. Am. **19**, 609-619.

Riesz, R. (1928). “Differential sensitivity of the ear for pure tones,” Phys. Rev. **31**, 867-875.