Loudness Compensation

Aleksi Myöhänen
Aalto University
Master's Programme CCIS / AAT

lauri.myohanen@aalto.fi

Abstract

Loudness compensation is used to correct for a non-linearity in human hearing that causes a change in the spectral balance of sounds when the listening level is altered. The non-linearity is described by equal loudness level contours that show the frequency and level dependence of the sensitivity of human hearing. The non-linearity is more prominent in low frequencies, where lowering of the listening level causes perceivable lack of bass.

Three approaches to loudness compensation are reviewed. The first method sees a closed-form expression derived for the compensation based on an approximation of the original loudness level. The second method approximates the required compensation as a linear function and the third uses a sensitivity function derived from the equal loudness level contours as a target for compensation filters.

The compensation may be applied in the frequency domain as gains for frequency bins or in the time domain with FIR or IIR filters. IIR filters are the preferred choice due to their low computational cost and filter order. High-fidelity compensation can be achieved using first order IIR shelving filters.

Keywords — Loudness Compensation, Equal Loudness Level Contours, Equalization

1 Introduction

Human hearing is not equally sensitive to all frequencies of sound. In addition to the frequency dependency, the sensitivity of hearing is also level dependent in a non-linear fashion. This effect is also the greatest at low frequencies, where the lowering of the level causes a larger decrease in the loudness relative to the middle and high frequencies. Thus, audio content played at a different level from the original recording or mastering level will have its spectral balance altered. There is no standards for level but it is estimated that 82 dB is a close approximation of typical mastering level [7]. This creates the need for loudness compensation since audio content such as music is often listened to at a lower level than the original mastering level. Without compensation the intended spectral balance would not match the balance experienced by the listener.

The loss of bass when listening to music at low levels can cause people to increase the volume in order to enhance the experience. However, the increased volume may lead to damage to the hearing during prolonged exposures. Therefore, loudness compensation can also help to prevent hearing loss as satisfying bass can also be achieved at safe levels. Typical applications of loudness compensation include loudspeaker systems and headphone audio.

2 Loudness

Loudness is a psychoacoustic quantity used to describe the subjective magnitude of an auditory event. Loudness is connected to the firing rate, or the number of nerve impulses per second sent by neurons along the auditory nerve [9]. As the sound level is increased the number of nerve impulses increases up to a certain level after which the firing rate stays constant.

The range of sound pressure levels (SPL) that human hearing is able to handle extends from the threshold of hearing at 0 dB to the threshold of pain at 130 dB [9]. As the decibel is a ratio of values, the 0 dB SPL corresponds to a reference pressure p_0 of 20 μPa. However, SPL is an objective measure and does not directly indicate how loud a sound is perceived.

Loudness as a psychoacoustic quantity is defined as the attribute of sound, which can be used to order sounds in a scale from quiet to loud [9]. The unit for loudness is sone, which is related to the sound pressure level (SPL) in such a manner that a 1kHz tone at 40 dB SPL has a loudness of 1 sone. However, loudness does not depend solely on the SPL of a sound event. It also depends on the spectral content and noises with a wide spectrum are perceived louder than narrowband noises even though they have the same SPL [9].

Another quantity for loudness is loudness level. The unit of loudness level unit is Phon, which is defined so that the SPL and loudness level of the reference tone are equal [9]. In other words, a 1 kHz tone at 40 dB SPL has a loudness level of 40 Phon. Loudness level in Phon is often preferred over loudness in sones when dealing with technical applications [9].

2.1 Equal Loudness Level Contours

Tones that have the same SPL but differ in frequency may not be perceived as equally loud. This phenomenon was first studied by Fletcher and Munson in 1933 [2]. In their study, they developed the first equal loudness level contours (ELLC). The contours define the required SPLs of different tones so that they are perceived as an equally loud. They determined the contours in listening tests, where the participants were asked to adjust the loudness level of a test tone to be equal to that of a 1 kHz reference tone. The 1 kHz reference was chosen because it is easily defined and made calculations simpler. The most recent and accurate data on equal loudness level contours is presented in ISO226:2003 standard [6].

The SPL L_f of a tone of frequency f with relation to its loudness level is given by equation [6]

$$L_f = \frac{10}{a_f} \log_{10} A_f - Lu + 94. \tag{1}$$

The coefficient A_f is defined as

$$A_f = \frac{4.47}{10^3}(10^{0.025L_n - 1.15}) + (0.4 \times 10^{\frac{T_f + L_u}{10} - 9})^{a_f}, \tag{2}$$

where T_f is the threshold of hearing represented in dB SPL, L_n is the loudness level in Phon, a_f is frequency dependent parameter, and L_u is the magnitude of the frequency response in dB. Figure 1 shows the ELLC calculated according to the equations and data from ISO226:2003.

The equations show that the contours are dependent on the listening level in addition to frequency. The level-dependent differences between the contours can be easily shown by normalizing them with relation to their 1 kHz component. From the normalized curves in Figure 2, it can be seen that a 100 Hz tone has to have a higher SPL as the 1000 Hz reference tone in order to achieve an equal loudness level. For high loudness levels the contours are more flat and SPL difference between an equally loud 100 Hz tone and a 1000 Hz tone is smaller than for lower levels. Therefore, the required compensation is dependent on both frequency and the original loudness level of the audio material.

It should be noted that loudness compensation systems based on older versions of ELLC may not produce the desired effect. Compensation systems based on the original ELLC by Fletcher and Munson typically call for more bass boost than is necessary [4]. Figure 3 shows the ELLC according to the data in both the currently confirmed ISO226:2003 standard and the old, withdrawn ISO226:1987 [5] standard as the red solid lines and black dashed lines, respectively. It can be seen that significant differences up to 10 dB exist between the contours. Thus, systems that use the old data will produce significantly different compensation. Therefore, the most recent and accurate data should be used in the design of a loudness compensation system in order to provide the best compensation possible.

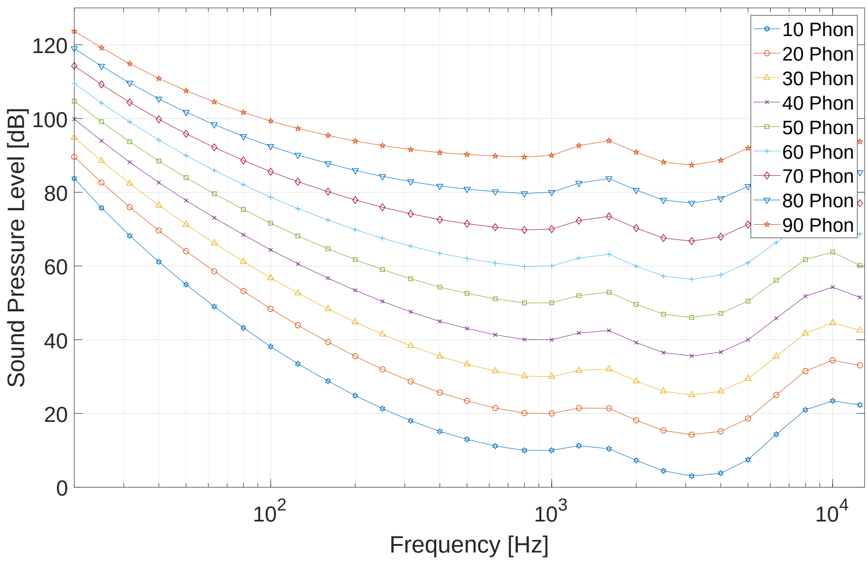

Figure 1: ELLC according to the ISO226:2003 data.

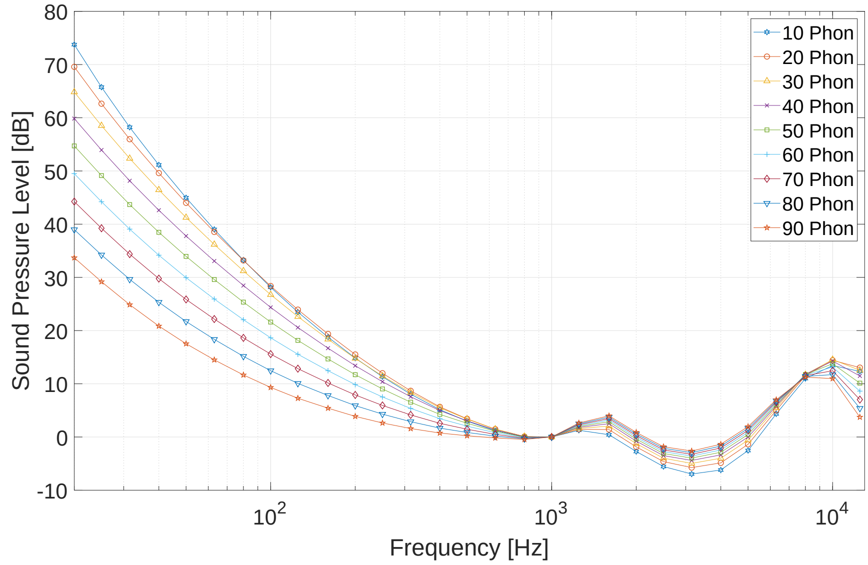

Figure 2: Normalized ELLC according to the ISO226:2003 data. Adapted from [1].

3 Loudness Compensation

Loudness compensation systems are used to correct the spectral content of a audio track played back at a different level from, where it was originally mixed. Therefore, an ideal loudness compensation system would preserve the spectral characteristics of given audio regardless of the playback level. This section presents some of the recent approaches to loudness compensation.

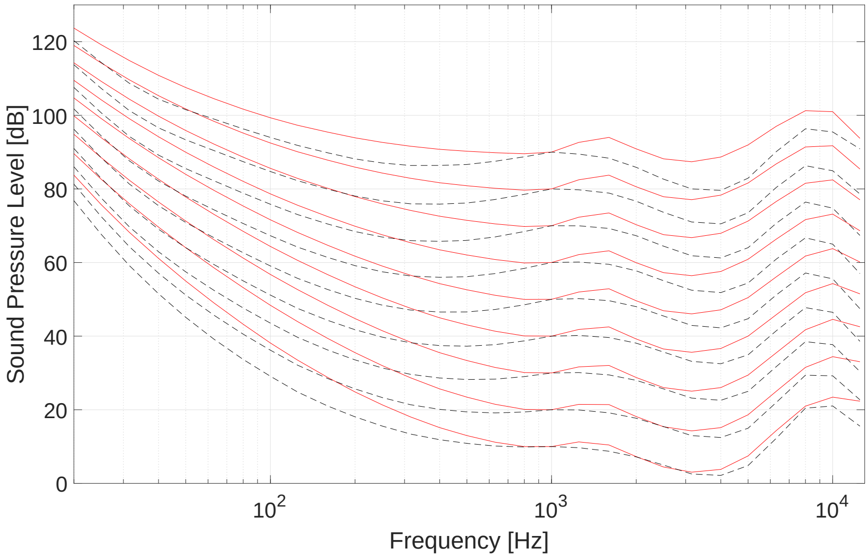

Figure 3: ELLC according to the data in the ISO226:2003 (red solid lines) and ISO226:1987 (black dashed lines). Contours representing the same loudness level can be found by inspecting the 1 kHz component.

3.1 Ideal Compensation

In [8] Prasad describes the ideal approach to loudness compensation. In an ideal system each frame of audio needs to be transformed to a filter bank domain representing the critical bands. The SPL of the audio system can the be calculated from the PCM energy. When the SPL is known, a lookup table can be used to determine the loudness level of the audio frame. Compensation calculated in this ideal manner varies spectrally and temporally. Therefore, the compensation needs to be smoothed in order to prevent artifacts from occurring the content. Then the appropriate compensation can be applied.

As the ideal compensation is calculated for each critical band and frame of audio separately, it is computationally complex and resource intensive. This constraints the implementation of loudness compensation on devices that have enough resources to handle the complexity.

3.2 Practical Approaches

Prasad [8] has proposed a compensation method, where the compensation is approximated as a function of the critical band and user applied attenuation. The motivation behind such approximation is the fact that the compensation does not greatly vary for a wide range of original loudness levels as illustrated in Figure 4. Thus, the approach proposed by Prasad reduces computational complexity by eliminating the dependency on the original loudness level of the audio. In this approach, the compensation is calculated only when the attenuation applied by the user is changed. Therefore, it is not required to calculate it for every frame of audio individually. Thus, the computational complexity of the system is significantly reduced.

Prasad removed the effect of original loudness level by averaging over the range of loudness levels from the threshold of hearing up to 90 Phon, which is the upper limit for the data in the ISO226 standard. He then derived a closed-form expression for his proposed compensation as given in equation

$$D(f,\Delta) = \Delta - \frac{10k_2\Delta}{a(f)} - \frac{Q_1(f) - Q_2(f,\Delta)}{R(f,\Delta)}, \tag{3}$$

where Δ is the user attenuation, $Q_1(f), Q_2(f,\Delta)$ and $R(f,\Delta)$ are variables related to the upper and lower limit of the averaging in Phon and the user defined attenuation.

The proposed approximation introduces an error, which Prasad quantified as the standard deviation from the ideal compensation. The reduction in complexity achieved by using this approach is around 65 % with the cost of an error of 1-1.5 Phon, when compared to the ideal compensation.

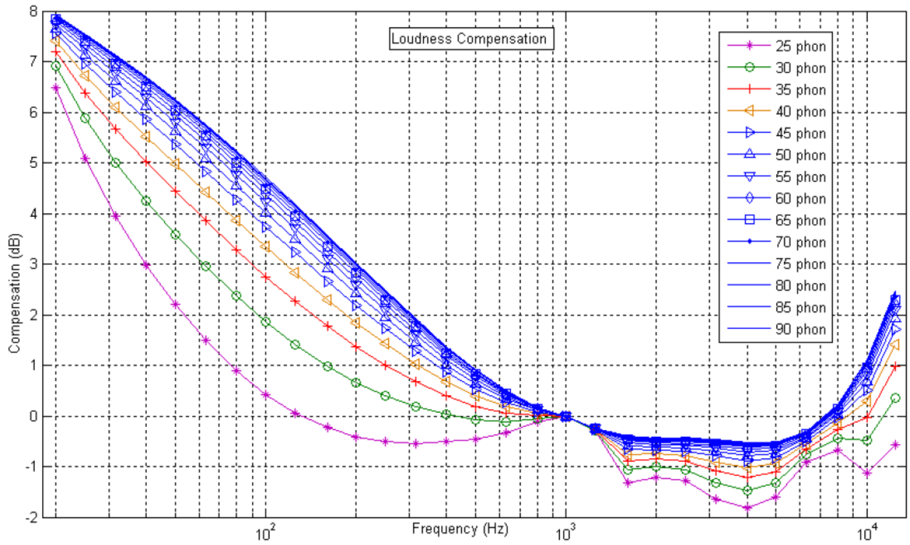

Figure 4: Ideal loudness compensation for user applied attenuation of 15 Phon. From [8].

Hawker and Wang [3] have proposed another method of loudness compensation for uncalibrated listening systems. They approximate the ELLC with a frequency dependent linear function, which represents a proportional relation between changes in loudness and SPL. In order to, achieve this representation they determined a frequency-dependent constant of proportionality by fitting a linear function to the

ISO226 data using the method of least squares. The relation between change in SPL and change in loudness level is then expressed as

$$\Delta SPL(f) = \Delta Phon(f)k(f), \tag{4}$$

where $k(f)$ is the frequency-dependent constant, which represents the change in SPL required to cause a change of 1 Phon in loudness level.

The required change in SPL to perform loudness compensation given the attenuation is then expressed as

$$\Delta SPL(f) = (1 - k(f))\Delta P, \tag{5}$$

where ΔP is the attenuation to be compensated. The absolute level of the audio is not considered in the calculation and only depends on the required change in loudness. Figure 5 shows the deviation of this linear model from the ELLC. The error of the model is less than 1 dB for most original loudness levels with the peak error being 3.27 dB for the 10-Phon curve.

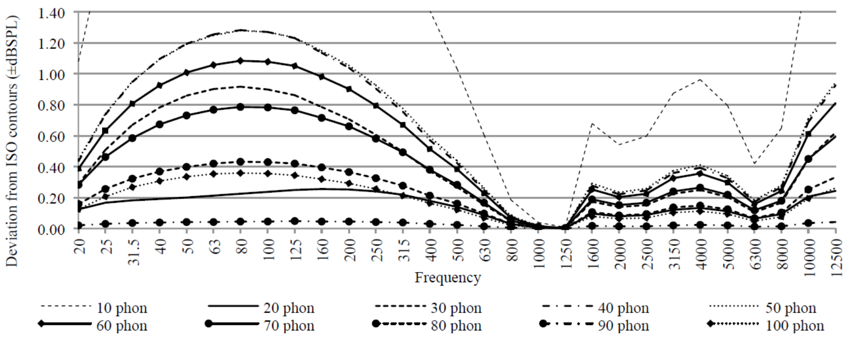

Figure 5: The error of the linear model with relation to ELLC. The 10-Phon line has a maximum error of 3.27 dB. From [3].

Fierro et al. [1] have proposed an adaptive method to loudness compensation. First they derived a sensitivity function by normalizing the ELLC relative to their value at 1 kHz. The sensitivity function is given by equation

$$S(f, L_n) = -L_p(f, L_n) + L_p(1000, L_n). \tag{6}$$

This sensitivity is the inverse of normalized curves shown in Figure 2.

A difference curve between the sensitivity function of the mastering level L_M and the listening level L_L can then calculated as

$$\Delta L_p(f, L_M, L_L) = L_p(f, L_M) - L_p(f, L_L) - N_f, \tag{7}$$

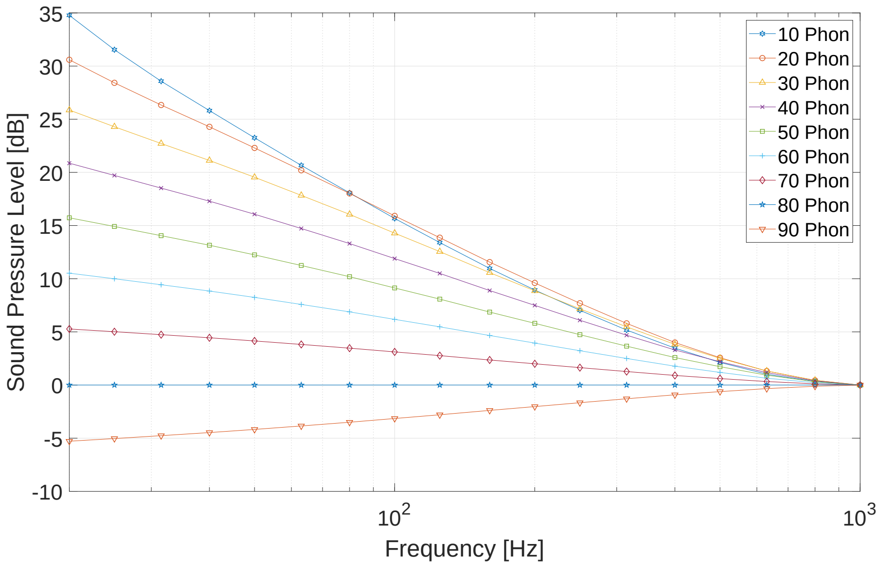

Figure 6: Compensation targets (trace-guides) for various listening levels when the mastering level is 80 dB SPL. Adapted from [1].

where N_f is normalization factor from the sensitivity function defined as

$$N_f = L_p(1000, L_M) - L_p(1000, L_L). \tag{8}$$

The inverse of the difference curve can the be used as a target for the compensation filter. Fierro et al. termed these as trace-guides. This term will be used to refer to the targets in the rest of this paper as well.

Figure 6 shows the trace-guides for audio mastered at 80 dB SPL and played back at various levels. It can be seen that lower the playback level the more bass amplification is required to maintain the same balance. However, if the playback level is higher than the mastering level a bass attenuation is required instead. This can be the case in some situations such as concerts.

3.3 Compensation Techniques

The compensation methods described above can be implemented using various approaches. Prasad used a frequency domain approach, where FFT algorithm is performed on each frame of audio and the compensation gains are applied to the corresponding frequencies. The audio frame is then transformed back to time domain using IFFT.

Hawker and Wang implemented their compensation with a FIR filter with a length of 1024. Using high order filters is computationally expensive and increases processing delay. They estimated their latency as 1.5 times the filter order corresponding to 1536 samples.

Fierro et al. utilized IIR filters, which can be implemented with lower cost than FIR filters. More precisely, they used first order IIR shelving filters. Shelving filters are used to boost or attenuate either low or high frequencies and are commonly used as bass and treble control. In contrast, to low-pass and high-pass filters they do not cut the frequencies outside their operational band. A shelving filter is described by the gain G and the cross-over frequency $\omega_c = 2\pi f_c/f_s$. The transfer function of a first order low shelving filter can be expressed as

$$H(z) = \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1}}, \tag{9}$$

where

$$\begin{cases} b_0 = \frac{G\,tan(\omega_c/2)+\sqrt{G}}{tan(\omega_c/2)+\sqrt{G}} \\ b_1 = \frac{G\,tan(\omega_c/2)-\sqrt{G}}{tan(\omega_c/2)+\sqrt{G}} \\ a_1 = \frac{tan(\omega_c/2)-\sqrt{G}}{tan(\omega_c/2)+\sqrt{G}} \end{cases}$$

Fierro et al. achieved their best result with a cascade of two first order shelving filters. The deviation from the trace-guides had a peak of 0.81 dB. A single shelving filter had a maximum deviation of 1.05 dB. Since a single filter implementation is desirable due to reduced complexity and computation time, a bias term was introduced to the gain. This improved the performance of the single filter and reduced the maximum deviation to 0.91 dB, which is less than 1 dB. Thus, high-fidelity was achieved with the single filter implementation as well. This method has low computational cost and enables real-time implementations.

The shelving filters can be calculated in advance for different listening levels and stored into memory. The appropriate filter can then be adaptively chosen depending on the listening level. If the mastering level is not known the calculation of the correct filter is not possible. However, as mastering levels are typically around 80-85 dB SPL, the trace-guides for the filters will be similar in this limited range [1]. Therefore, by assuming L_M in that range, accurate compensation can be performed for most audio material.

Loudness compensation may also be performed using dynamic range compression (DRC). DRC is used to either reducing the volume of loud sounds or amplifying quiet sounds. In both cases, the dynamic range of the signal is reduced, hence the name. DRC can be used to increase the loudness of an audio signal ensuring that the quietest

4 Conclusions

The need for loudness compensation arises from a non-linearity in human hearing. This non-linearity is described by the ELLC, which show that to achieve equal loudness tones of different frequencies require different SPLs. Importantly, low frequency tones need to have higher SPL compared to high frequency tones to produce the same loudness. The required difference in SPL is also dependent on the level of the tone in addition to frequency. Thus, audio played back at a lower level than at which the recording was made has a different spectral balance.

The perceived reduction in low frequencies at low levels is significant, especially for music listening. Without compensation listeners are tempted to increase the listening level in order to enhance the bass, which can cause damage to the hearing during prolonged exposure. Applying loudness compensation to music can provide a satisfying bass at lower levels thus avoiding the need to increase the listening level. Loudness compensation is not restricted to headphone usage only and can also be used in loudspeakers to control the bass.

Perceptually accurate loudness compensation is based on the ELLC. Prasad formulated a closed-form expression for loudness compensation through an approximation of the original loudness level. Hawker and Wang approximated the compensation as a linear relationship between required change in SPL and desired change in loudness level. Fierro et al. derived a sensitivity function, which they used to calculate difference curves between different listening levels. They then used inverses of the difference curves as trace-guides for the compensation filters. The approaches to loudness compensation can thus be different. However, all reviewed methods have in common the usage of ELLC data.

Loudness compensation can be implemented in frequency domain as in [8], where the gains can be applied to appropriate frequency bins. This required each audio frame to be transformed to the frequency domain and back to time domain using FFT and IFFT. Time domain filtering can also be used to perform the compensation using either FIR or IIR filters. A problem with FIR filters is that a high order is often required to achieve the desired result. Hawker and Wang implemented loudness compensation using an FIR filter, which had a length 1024 samples. The high order increases computational complexity and introduces a significant delay in the processing. IIR filters are more computationally efficient than FIR filters and do not require as high filter orders. Fierro et al. used IIR shelving filters to perform the loudness compensation. Their design achieved high-fidelity ($\pm 1\, dB$) using only a single first order shelving filter.

References

[1] Fierro, L., Rämö, J., and Välimäki, V. Adaptive loudness compensation in music listening. In *Proceedings of the 16th Sound and Music Computing Conference SMC* (2019), pp. 1–8.

[2] Fletcher, H., and Munson, W. A. Loudness, its definition, measurement and calculation. *Bell System Technical Journal 12*, 4 (1933), 377–430.

[3] Hawker, O., and Wang, Y. A method of equal loudness compensation for uncalibrated listening systems. In *Audio Engineering Society Convention 139* (Oct 2015).

[4] Holman, T., and Kampmann, F. Loudness compensation: Use and abuse. *Journal of the Audio Engineering Society 26*, 7/8 (1978), 526–536.

[5] ISO, B. 226: 1987: Acoustics–normal equal-loudness-level contours. *International Organization for Standardization* (1987).

[6] ISO, B. 226: 2003: Acoustics–normal equal-loudness-level contours. *International Organization for Standardization 63* (2003).

[7] Nielsen, S. A loudness function for analog and digital sound systems based on equal loudness level contours. In *Audio Engineering Society Convention 140* (2016), Audio Engineering Society.

[8] Prasad, P. D. A low complexity approach for loudness compensation. In *Audio Engineering Society Convention 129* (Nov 2010).

[9] Pulkki, V., and Karjalainen, M. *Communication Acoustics: An Introduction to Speech, Audio and Psychoacoustics.* John Wiley & Sons, 2015.