Audio Engineering Society

Convention Paper

Presented at the 130th Convention
2011 May 13–16 London, UK

Perceptually Robust Headphone Equalization for Binaural Reproduction

Bruno Masiero[1], Janina Fels[1]

[1] *Institute of Technical Acoustics, RWTH Aachen University, Germany*

Correspondence should be addressed to Bruno Masiero (bma@akustik.rwth-aachen.de)

ABSTRACT

Headphones must always be adequately equalized when used for reproducing binaural signals if they are to deliver high perceptual plausibility. However, the transfer function between headphones and ear drums (HpTF) varies quite heavily with the headphone fitting for high frequencies, thus even small displacements of the headphone after equalization will lead to irregularities in the resulting frequency response. Keeping in mind that irregularities in the form of peaks are more disturbing than equivalent valleys, a new method for designing headphone equalization filters is proposed where not the average but an upper variance limit of many measured HpTFs is inverted. Such a filter yields perceptually robust equalization since the equalized frequency response will, with high chance, differ from the ideal response only by the presence of valleys in the high frequency range.

1. INTRODUCTION

The realism of binaural reproduction through headphones significantly increases if the headphones are adequately equalized. Nevertheless, the correct equalization of headphones is still an open research topic as the equalization strongly depends on the individual coupling between headphones and the listeners' ear.

It can be verified that open-type headphones work as a volume cavity system up to frequencies neighboring 4 kHz. Above this frequency region, standing waves (or equivalently, modes) start to build up inside the cavity and thus the resulting pressure at the listeners' eardrums becomes strongly dependent on the headphone fitting [1, pg. 84]. The *Thevenin* model of the human external ear described by Møller in his seminal work on binaural technology [2] is then no longer valid for this frequency range. This effect makes a direct equalization of the headphone, where usually a microphone is briefly placed in the entrance of the blocked ear canal and later removed, nearly impossible since one cannot remove the microphone without temporarily displacing the headphone.

Also the fact that the headphone transfer function (HpTF) used for equalization is measured with the blocked ear canal has an influence on the equalization quality, as in the listening situation a different impedance match occurs, altering the pressure levels at the entrance of the ear canal and consequently on the ear drums [3].

Bücklein has shown through speech intelligibility tests that human listeners are more sensitive to spectrum irregularities in the form of peaks than to equivalent valleys [4]. Assuming that this behavior extends also to spatial perception, we describe in this paper a headphone equalization method that avoids the occurence of resonance peaks in the entrance of the ear canal. Other types of equalization filters also inspired by the work of Bücklein were tested by Lindau and Brinkmann [5].

In this paper we first discuss the influence of headphone fitting on the measured HpTF, both for an artificial head as well as for individual listeners. We then present a new headphone equalization technique that is perceptually robust to variations in headphone fitting and conclude discussing the obtained results.

2. HEADPHONE FITTING

The variance of HpTF with the headphone fitting has already been extensively investigated [2, 6, 7]. It has also been shown that the spectral differences caused by distinct headphone fittings are perceivable by listeners [8]. Thus, perfect equalization of headphone transfer functions will increase the localization quality of presented binaural signals.

As discussed by Hammershøi, the pursuit of a single headphone equalization filter, valid for every user, makes sense only in applications where sound localization is not an important criterion, since the variation of HpTF across subjects is high and this will lead to spectral colorations that mainly affect the perception of elevation but also the externalization of sounds [9]. On the other hand, when the listeners are allowed to place the headphone as it is most comfortable for them, the variation in measured HpTF reduces considerably, suggesting that an individual

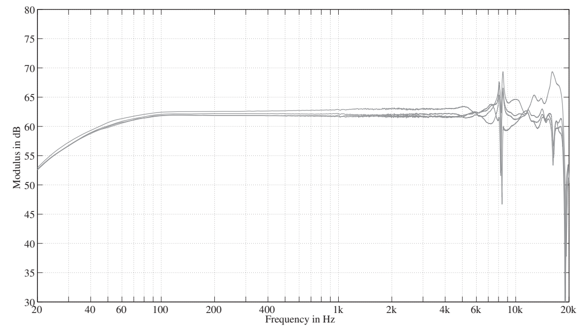


Figure 1: Frequency response of a headphone equalized with a filter designed from the average of four HpTF and no equalization below 100 Hz.

headphone filter can be designed. Furthermore, Lindau and Brinkmann argued that for the reproduction of signals generated by binaural synthesis, the headphone equalization filters with better perceptual property is the one originated from the HpTF of the same (dummy) head used for the recordings of the head-related transfer function (HRTF) data base used for the synthesis [5]. They do, however, acknowledge that for individually measured HRTF individual headphone equalization should be the best choice.

Thus, the following question is posed: is it sufficient to use a single HpTF, measured after the user has placed the headphone at the most comfortable fit, to generate an equalization filter? Or should the average of several HpTF measurements be used? The answer: a headphone equalized with a filter generated from the inverse of either a single measurement or the average of several measured HpTF will most certainly have an equalized response containing peak irregularities, as exemplified in Fig. 1 for the open-type Sennheiser HD600 headphone.[1]

It seems reasonable to take into account how the HpTF varies for each user in the design of the equalization filter, avoiding the presence of such peak irregularities in the equalized HpTF. To do so, we first need to analyze how an individual HpTF varies with the headphone fit. A qualitative analysis of this variation is presented below. Two open-type

[1] All HpTF plots have the y-axis displayed in dB relative to 1 Pa/V and only results of the left ear are displayed.

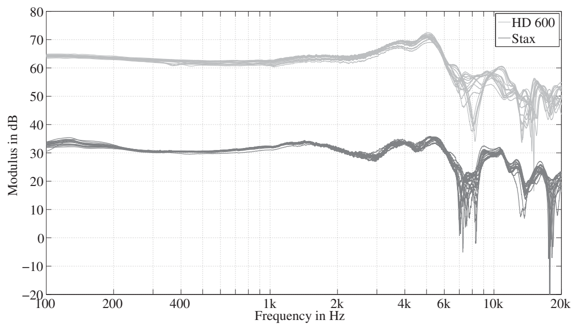


Figure 2: HpTF of two open-type headphones measured with a dummy head. 18 fittings are shown for Sennheiser HD 600 (light gray) and Stax SR Lambda (dark gray, values shifted -30 dB for clarity).

headphones were used for the measurements: the dynamic headphone Sennheiser HD600 and the electrostatic headphone Stax SR Lambda.

2.1. Artificial Head

It is obviously much easier to measure the HpTF on a dummy head, since microphones are already fixed on the entrance of the ear canal. But since it is a passive listener, the variance of the measured HpTF is expected to be higher than in the case when the listener is allowed to place the headphone at its most comfortable fit.

As shown in Fig. 2, the HpTF has little variance up to about 4 kHz. In this region just a constant level variation can be noted, caused by different leakage with the fitting, as commented by Toole [6]. In this frequency range the headphone works as an acoustic cavity. For higher frequencies we can observe resonances that vary with the headphone fitting and the geometry of the listeners' ears. This occurs because at this frequency range standing waves start to build up inside the headphone cavity and also on the external ear structure [1], meaning that in this region the HpTF behavior is highly individual.

2.2. Individual Fitting

We also carried out a series of individual HpTF measurements with 15 listeners. The same behavior observed in the measurements with the dummy head can be seen in the individual measurements, as presented in Fig. 4. In lower frequencies, only level differences are present while at higher frequencies

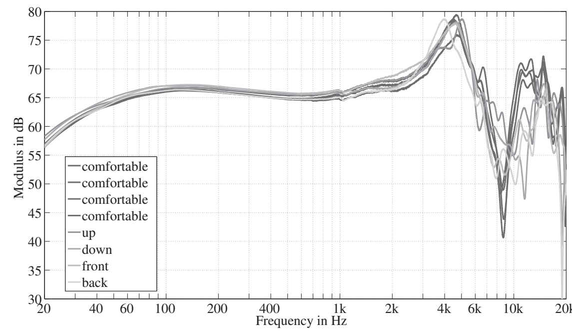


Figure 3: HpTF of a single subject measured four times with comfortable fit and at four extreme positons with the Sennheiser HD600 headphone.

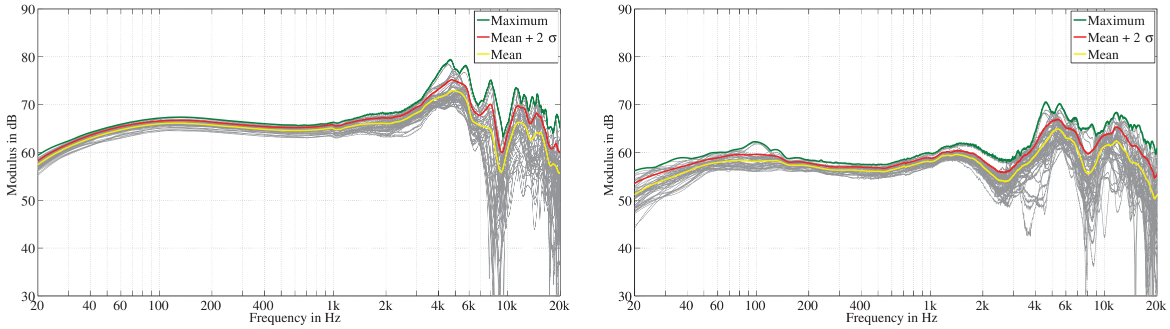
we observe the presence of very individual resonance structures.

Each subject was asked to place the headphone four times with the most comfortable fit and four times with extreme fits, i.e. with ears positioned as much as possible to the front, back, top and bottom of the headphone interior. This should give an overview of the extremes of how the individual HpTF varies, as exemplarily shown for one subject on Fig. 3. No trends could be observed at the four measured extreme positions. Nonetheless, by having confirmed the low measurement variability when the listener is allowed to fit the headphone, just the variability of these measurements should be considered when designing the equalization filter.

3. HEADPHONE EQUALIZATION

The perceptually robust equalization is based on the assumption that notch irregularities on the equalized response are not as disturbing as peaks. To achieve this goal, we measure the HpTF several times, always completely removing the headphones between measurements, and then calculate the upper amplitude limit of the HpTFs. If sufficiently many positions are covered, all local notches that would result in peaks in the equalization filter should be ignored, ensuring that, if the equalized HpTF contains irregularities, these irregularities will most probably be in the form of valleys and not peaks.

The most straightforward way to obtain the upper limit of the frequency response is to take, for every



(a) (b)

Figure 4: Individual HpTF measured at 15 subjects (4 repetitions). Subjects should place the headphone at its most comfortable fit. Left speaker of (a) Sennheiser HD 600 and (b) Stax SR Lambda.

frequency, the maximum of the absolute value of all measurements. Taking the maximum yields a curve with a locally discontinuous derivative, that should then be removed with a smoothing function. Yet, this method is very sensitive to outliers, so a statistical approach might be more appropriate. If we assume that the measured points are normally distributed and calculate for each frequency the mean value μ and the standard deviation σ, the curve obtained with $\mu + 2\sigma$ will be above the measured HpTFs with over 95 % chance. This method has two main advantages: it results in a relatively smooth curve and it is more robust to outliers. Also, to reduce the influence of irregularities with high quality factor Q, all HpTFs are smoothed with a 1/6-octave moving average function.

Both methods work with the absolute value of the HpTF giving no information about the behavior of the phase. This means that artificial phase information must be generated. We generated a minimum phase spectrum by the means of the Hilbert transform, thus producing a causal equalization filter. Other approaches, as averaging the group delay or unwrapped phase did not provide substantially different time results. Furthermore, listening tests conducted by Lindau and Brinkmann showed that listeners could not distinguish between headphone equalization filters with either unconstrained phase or minimum phase [5].

As with any other equalization filter, care must be taken below and above the low and high roll-off frequencies, respectively, as frequency correction at these regions would lead to very large gains that would produce undesired nonlinearity in the response. If the headphone is to be equalized in the frequency range from 20 Hz to 100 Hz we end up with a very long FIR equalization filter. Since these equalization filters are aimed for use in real time virtual reality systems, it is important that they are kept short in order to avoid extra latency. As the low frequency range does not contribute to localization and the HpTF variation to individual fitting in this frequency range is very low, this frequency region was left unchanged. This is done by fitting a constant line below the first maximum of the HpTF.

Regarding the filter gain, ideally, the equalization filter should not alter the loudness of the reproduced signals; but loudness measurement is dependent on the type of signal being played. For broadband signals, if the overall sound pressure level is kept constant, then there should be negligible variation on the loudness values. So the inverse HpTF is weighted with the inverse of its root mean square value.

4. RESULTS

It is clear that individual equalization should always be preferred. But there might be cases where individual HpTF measurements are not practicable. In such cases Brinkmann and Lindau suggest using a filter generated from the HpTF measured with the same "head" used for the binaural measurement/synthesis [10]. Still, they admit that their result might have been strongly influenced by their test setup. Another approach would be to design a

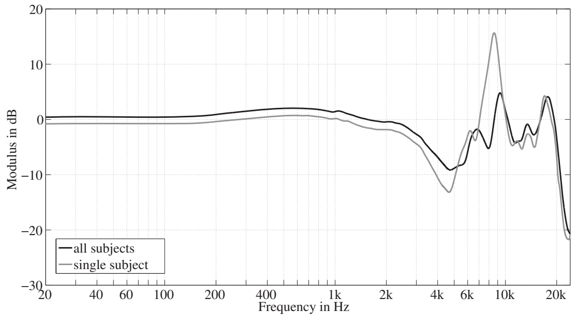


Figure 5: Equalization filter for the Sennheiser HD600 headphone obtained by inverting $\mu+2\sigma$ from 15 subjects (black) and of a single subject (gray).

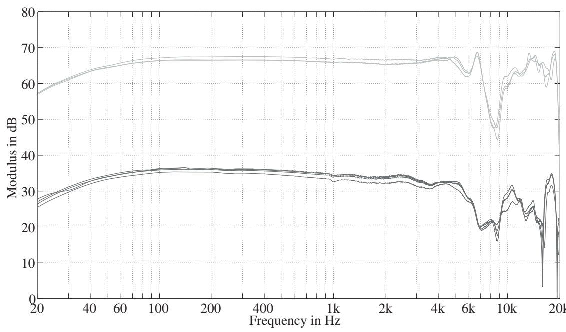


Figure 6: Equalization response for two subjects using the black filter from Fig. 5, also for the Sennheiser HD600 headphone. The dark gray curve is shifted -30 dB for clarity.

filter as the inverse of $\mu + 2\sigma$ from an ensemble of measured HpTF. Fig. 5 shows an example of such a filter compared to an individual equalization filter, both generated with the technique presented in the previous section. An example of the equalized response of such a filter can be seen on Fig. 6. We note that for higher frequencies quite large frequency dips are present, which might result in audible artifacts. The effectiveness of this filter must still be tested in an appropriate listening test.

If an individual measurement of the HpTF is possible, then this equalization method should be favored. It is clear from Fig. 7 that the equalized responses obtained in this manner contains much less irregularities and the irregularities that do occur are

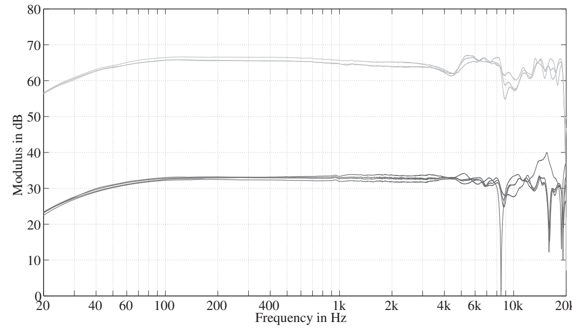


Figure 7: Equalization responce for two subjecs using a individual filter. The light gray curves were generated with the filter shown in Fig. 5. Dark gray curve is shifted -30 dB for clarity.

almost always frequency valleys. However, outliers might still occur, as depicted in one of the dark gray curves of Fig. 7, where a peak irregularity is present. It is important to mention that the HpTFs used in the filter generation are not the same as the ones used for the equalized response calculation in the examples of Fig. 6-7.

One should keep in mind that the real statistical variance of the HpTF measurements is unknown and is substituted by a sample variance taken from a limited number of measurements. Certainly the quality of the variance estimation will increase with the number of measurements. Ferranti et. al. suggest that eight measurements suffices to safely replace the exact variance by the sample variance for a 95 % confidence bound [11].

4.1. Impedance Match

So far, all measurements and corresponding equalization filters were made at the blocked entrance of the ear canal, not considering how the impedance mismatch will influence the final equalization. The impedance mismatch occurs because the HpTFs are measured with a blocked ear canal whereas under normal use the entrance of the ear canal is not obstructed.

The effect of the impedance mismatch was analyzed using a dummy head equipped with an ear simulator, i.e. with an artificial ear canal. Ideally, binaural reproduction wants to correctly reproduce, at the listeners' ear drums, the signal that would be present there if the listener was located in the virtual scene. So we assume a very simple scene with one point source in free-field to analyse how the pressure spectrum at the ear simulator changes because of the impedance mismatch.

We first measure the transfer function from source to the ear simulator H_{open} and then the transfer function from source to the blocked ear canal entrance H_{closed}. Then we measure the HpTF to the closed ear canal H^p_{closed} and to the ear simulator H^p_{open}. If the headphone is equalized with $1/H^p_{closed}$ then we have a flat frequency response at the entrance of the blocked ear canal. If there were no impedance mismatch caused by the presence of the earplug, the pressure spectrum at the ear simulator generated by the equalized headphone when multiplied by H_{closed} should be the same as H_{open}. That is,

$$H^p_{open} \cdot H_{closed}/H^p_{closed} = H_{open}, \tag{1}$$

$$\frac{H^p_{open}/H^p_{closed}}{H_{open}/H_{closed}} \equiv 1. \tag{2}$$

Møller named (2) as “pressure division ratio”. This shows how the headphone coupling, caused by the different radiation impedance “seen” by the ear canal, influences the pressure at the ear drum. If the equality is observed, then the headphone is termed a “free-field equivalent coupling” (FEC) headphone.

We verify that both headphones discussed in this paper have a PDR that varies relatively little (± 2 dB) and could be said to nearly meet the FEC criterion, as shown on Fig. 8. This means that up to about 10 kHz the influence of impedance mismatch can be neglected for these headphones.

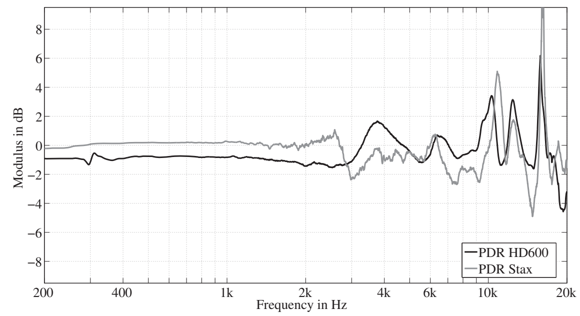


Figure 8: Pressure division ratio for headphones Stax SR Lambda and Sennheiser HD600. Average of eight HpTF measurements made with the help of a mounting devide to reduce high frequency variability and HRTF measured for five different directions.

5. CONCLUSION

Headphone Transfer Functions (HpTFs) were measured with a dummy head and individual subjects, confirming that for the low and middle frequency range only small level variations are present while for the high frequency range very individual resonance patterns are found. Low measurement variability is achieved if the subject is allowed to fit the headphone at the most comfortable position. This reinforces the fact that individual equalization should be used when possible.

A perceptually robust equalization filter design is proposed, inverting not the average but the upper limit of several measured HpTFs. This leads to an equalized response whose potential irregularities will be in the form of dips and not peaks. This is preferable because the human hearing system is more sensitive to peak irregularities that to equivalent valley irregularities. The upper limit is calculated from the average plus two times the standard deviation from the amplitude of the measured HpTFs. Since phase information is lost at this process, minimum phase is used. To keep the FIR filter short, avoiding latency at the reproduction chain, the headphone is not equalized for frequencies below 100 Hz.

At the time of writing only informal listening tests have been carried out, nevertheless, three experienced listeners reported a significant increase in scene realism when a synthesized binaural signal was played through a headphone equalized with the presented method in direct comparison to the same signal played with no headphone equalization.

6. ACKNOWLEDGMENT

Bruno Masiero is a scholarship holder from the National Council of Scientific and Technological Development; CNPq - Brazil.

7. REFERENCES

[1] S. Schmidt, “Finite Element Simulation of External Ear Sound Fields for the Optimization of Eardrum-Related Measurements,” PhD, Ruhr-Universität Bochum, 2009.

[2] H. Møller, "Fundamentals of binaural technology," *Applied Acoustics*, vol. 36, no. 3-4, pp. 171–218, 1992.

[3] H. Møller, D. Hammershøi, C. B. Jensen, and M. F. Sø rensen, "Transfer Characteristics of Headphones Measured on Human Ears," *Journal of the Audio Engineering Society*, vol. 43, no. 4, pp. 203–217, 1995.

[4] R. Bucklein, "The audibility of frequency response irregularities," *J. Audio Eng. Soc*, vol. 29, no. 3, pp. 126–131, 1981.

[5] A. Lindau and F. Brinkmann, "Perceptual evaluation of individual headphone compensation in binaural synthesis based on non-individual recordings," in *3rd ISCA/DEGA Tutorial and Research Workshop on Perceptual Quality of Systems*, Dresden, Germany, 2010, pp. 137–142.

[6] F. E. Toole, "The Acoustics and Psychoacoustics of Headphones," in *AES 2nd International Conference*, Anaheim, USA, 1984.

[7] C. G. M. Cruzado, "Influence of the Acoustic Impedance of the Headphone on Psychoacoustic Effects," Master Thesis, Institute of Technical Acoustics, RWTH Aachen University, 2002.

[8] M. Paquier and V. Koehl, "Audibility of headphone positioning variability," in *Audio Engineering Society Convention 128*, 2010.

[9] D. Hammershøi and H. Møller, *Binaural Technique - Basic Methods for Recording, Synthesis and Reproduction*. Springer-Verlag, 2005, p. 379.

[10] F. Brinkmann and A. Lindau, "On the effect of individual headphone compensation in binaural synthesis Measuring individual HPTFs auditory modeling of inversion effects," in *DAGA - Fortschritt der Akustik*, Berlin, Germany, 2010.

[11] F. Ferranti, Y. Rolain, L. Knockaert, and T. Dhaene, "Variance weighted vector fitting for noisy frequency responses," *Microwave and Wireless Components Letters, IEEE*, vol. 20, no. 4, pp. 187–189, 2010.