Audio Engineering Society

Convention Express Paper 352

Presented at the AES 158th Convention
2025 May 22–24, Warsaw, Poland

This Express Paper was selected on the basis of a submitted synopsis that has been peer-reviewed by at least two qualified anonymous reviewers. The complete manuscript was not peer reviewed. This Express Paper has been reproduced from the author's advance manuscript without editing, corrections, or consideration by the Review Board. The AES takes no responsibility for the contents. This paper is available in the AES E-Library (http://www.aes.org/e-lib), all rights reserved. Reproduction of this paper, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.

Towards a Headphone Target Curve for Spatial Audio

Alexander Mülleder[1], Nils Meyer-Kahlen[2], and Matthias Frank[1]

[1]*University of Music and Performing Arts Graz*
[2]*Acoustics Lab, Dpt. of Information and Communications Engineering, Aalto University, Espoo, Finlawend*

Correspondence should be addressed to Alexander Mülleder (alex.muelleder@gmx.at)

ABSTRACT

In order to reproduce audio over headphones as intended, it is essential to have well-defined and consistent references of how headphones should sound. With the aim of stereo reproduction in mind, the field has established a de-facto reference target curve called the Harman Target Curve to which headphone transfer functions are commonly compared. This contribution questions if the same target curve is suitable when used for the reproduction of spatial audio. First, the origins the Harman Curve are revisited; it is motivated by the frequency response of loudspeaker playback in a specific listening room. The necessary measurement procedures are described in detail. Then, the paper discusses the applicability of existing targets to spatial audio. Therein, it is possible to embed convincing spatial room information directly into the production, thereby calling into question the motivation for incorporating a listening room in the headphone target. The paper concludes with a listening experiment that compares the preference of different target curves for both spatial audio and stereo.

1 Introduction

In the process of designing and evaluating headphones, it is essential to consider a target curve. It is used to design the frequency response of headphones, or to compare the response of an existing pair with. Typically, target curves try to equalize the headphone response to a free field response, a diffuse field response, or try to mimick the response obtained during loudspeaker reproduction in a room. The latter approach, which is sometimes called semi-reflective field target, together with parametric adjustments (e.g., shelving filters) based on subjective listening experiments, was used to derive the widely adopted Harman Target Curve. For stereo content, such an approach is justifiable, and listening experiments have shown that the Harman Curve yields high preference amongst listeners [1, 2]. As a consequence, it is often used as a reference curve to compare other headphone responses against, at least in the HiFi community, e.g., [3, 4].

However, deriving such modified semi-reflective field targets naturally depends on the specific room, head and torso simulator (HATS), and preference of the listeners involved in determining the parametric adjustments.

Furthermore, the attempt to emulate loudspeaker-based stereo reproduction on headphones by equalization alone may be questionable as such; to really emulate a room, convolution with a binaural room impulse response would be required, which is neither widespread among consumers nor recommended for standard use. With the increasing availability of spatial audio content, drawing the motivation for targets from stereo reproduction in a room becomes even more debatable. Thus, the question of whether a new target should be found arises. Indeed, this would provide the opportunity to define it in a more generalizable manner, so that it is easier to employ it independently of a specific room, HATS, and preference profile.

This paper proposes a method for defining a headphone target curve for spatial audio reproduction. It uses a diffuse-field target as a basis, which takes into account the fact that spatial productions involve all directions and that convincing spatial room information can be embedded into the production itself. Moreover, using a diffuse-field target as a basis allows for measurements with any HATS, for which a diffuse-field measurement is available. To assure compatibility with stereo content, a -1 dB per octave slope is applied on top, which we show to yield a curve that is similar to the Harman Curve, except for its low bass boost. Such a slope is easy to apply and invert if a diffuse-field response is desired for specific applications.

We point out that being able to equalize a headphone to match this target does not mean that every headphone manufacturer will and should do so. We highlight that this target curve can also be used to compare other headphone transfer functions against, where what remains of their response after dividing by it can be seen as a preference curve, defining the unique characteristics of the headphones. In the following section, we explain how headphones are equalized to different target curves and highlight the practical and conceptual problems around the process. Then, in Section 3, we describe the process of obtaining the proposed target curve. In Section 4, we compare listener preference for different target curves, including the proposed one, for stereo and spatial audio playback in a listening experiment. We also show that headphone equalization can have an influence on the spatial qualities of the reproduction. Section 5 discusses the implications of the choices and shows the use of the target curve not only for equalization, but also for evaluating and comparing different headphones. Section 6 concludes the paper.

2 Headphone Targets

To determine the headphone transfer function (HpTF), for evaluation or subsequent equalization, the headphones must be measured on a human head with binaural microphones, or on a HATS, which simulates the behavior of an average human ear as closely as possible. Such ear simulators will have properties that become embedded in the measurements. In addition to the characteristics of the built-in microphones, the ear canal resonance will be included in the HpTF measurement, leading to a prominent peak between 2 and 5 kHz. However, using a measurement apparatus that does not include an ear canal is not a good solution, as the interaction between the headphone and the ear would otherwise not be captured correctly. At the same time, features of the apparatus are not part of the headphones and, therefore, must be compensated for. Ways to do so lead to different target curves.

In general, equalization with respect to a specific target curve is performed by designing a filter with a magnitude response that is equal to that of

$$H_{\mathrm{EQ}}(f) = \frac{H_{\mathrm{Target}}(f)}{H_{\mathrm{HpTF}}(f)}, \tag{1}$$

where f denotes frequency.

Three typical approaches can be distinguished when selecting the target curve. The first approach yields the so-called free-field target. Headphones meet this requirement if the reproduced sound pressure at a listener's eardrum is equal to the pressure from a single active source in anechoic conditions, using the same HATS as used for the HpTF measurement. This transfer path is characterized by the HATS's head-related transfer function (HRTF). A natural choice is to use the frontal direction, i.e., azimuth and elevation angle $\phi = 0^\circ, \theta = 0^\circ$, so that the free-field target becomes [5]

$$H^{\mathrm{FF}}_{\mathrm{Target}}(f) = |\mathrm{HRTF}(0^\circ, 0^\circ, f)|. \tag{2}$$

The second approach results in the diffuse-field target, in which the reproduced sound pressure at the eardrum is equal to that observed when the HATS is placed in a diffuse field. In theory, a diffuse-field response can be obtained by integrating squared HRTFs over the sphere [5]

$$H^{\mathrm{DF}}_{\mathrm{Target}}(f) = \sqrt{\frac{1}{4\pi}\int_{\phi}\int_{\theta} |\mathrm{HRTF}(\phi, \theta, f)|^2 \sin\theta \, d\theta \, d\phi}. \tag{3}$$

In practice, the integral is approximated by a finite set of measurement directions

$$H_{\text{Target}}^{\text{DF}}(f) \approx \sqrt{\frac{1}{L}\sum_{l} w_l |\text{HRTF}(\phi_l, \theta_l, f)|^2}, \quad (4)$$

where appropriate quadrature weights w_l need to be used depending on the measurement grid, see for example [6] for a discussion in the context of spatial audio.

The third approach is known as the semi-reflective room target, which aims to make the timbre of the headphones sound like loudspeaker reproduction in a given room. One well-known room-informed curve is the so-called Harman Target. The process of its creation is documented in detail in [1] and [2]. It was created by first equalizing loudspeakers of a given loudspeaker setup within a room to a flat target and then measuring the same response with a HATS. Afterwards, a listening experiment in which participants could adjust bass and treble ensured that the target was preferred by a panel of listeners.

3 Proposed Method for a Headphone Target for Spatial Audio

The procedure of obtaining the Harman Target bears a clear limitation. As it is based on a measurement with a specific HATS in a specific room, only HpTF measurements with the same HATS (a GRAS KEMAR) can be used to compare against that target curve. Moreover, with the increasing availability of spatial audio, it is questionable whether the motivation to match the headphone response to a good playback room holds up. One could argue that there is no ideal listening room for spatial audio, as it allows for incorporating natural sounding reverberation of any room. With the possibility of including such spatial reverb and the fact that sound sources may originate from all directions, using a diffuse-field target might appear reasonable.

Even though the Harman Target was not originally intended to be a reference curve, it has been widely adopted as a target curve by many in the community, and drastically changing it is not feasible. Therefore, a reference is needed that is not only well-defined but also compatible with existing practices. That is, using it for listening to existing material should not yield low preference. Existing studies have shown that equalization to a diffuse-field target is generally not the most preferred option [1]. As a solution, we propose using a diffuse-field target with an additional spectral slope. Effectively, the resulting curve is similar to the semireflective field Harman curve, except for the included low bass boost. This approach provides a simple, reproducible, and parametric modification that does not depend on a specific HATS and room measurement. The target can be written as

$$H_{\text{Target}}^{\text{DF-Slope}}(f) = H_{\text{Target}}^{\text{DF}}(f) 10^{\frac{\gamma \log_2(f'/f_0)}{20}}, \quad (5)$$

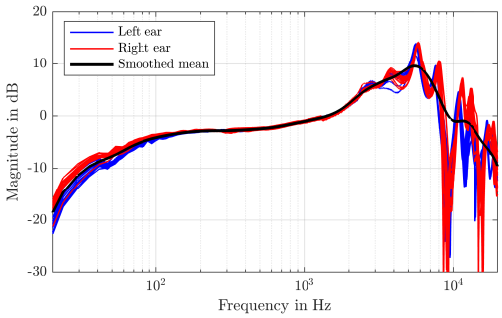
where γ is the slope in dB per octave. A slope of approximately -1 dB per octave corresponds to a reduction of -10 dB for the audible range of 20 Hz to 20 kHz, or -3.3 dB per decade. f_0 is a normalization frequency at which there is no modification. Interestingly, a -1 dB per octave slope was also considered an optimal curve for HiFi equipment in rooms in early work regarding HiFi quality [7]. Certain communities, such as those on [8], have also recently adopted this target. In the following listening experiment, we also included a steeper slope.

4 Listening experiment

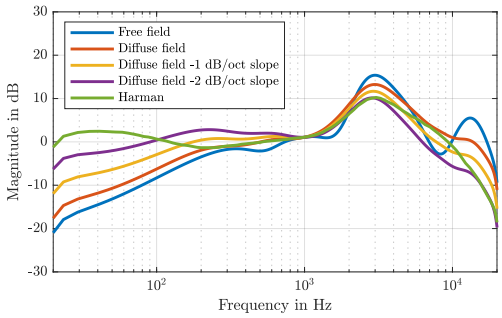
We conducted a listening experiment that compared a headphone equalized to five different headphone targets. In the first part, the targets were compared regarding timbral preference, using stereo audio material. In the second and third parts, binaural renderings of spatial audio material were used. Here, we asked for timbral preference, and preference regarding spatial features.

A Sennheiser HD800 high-quality circumaural headphone was selected for the experiment, as in [2]. A GRAS KEMAR HATS with an ear simulator was employed to design the headphone equalization filter. The experimenter placed the headphones on the HATS 15 times. Upon each placement, an HpTF was measured using an exponential sine sweep. The result can be seen in Fig. 1(a), along with the third-octave smoothed average magnitude response between the two ears, which was used for designing the equalization filters. The five tested target curves were:

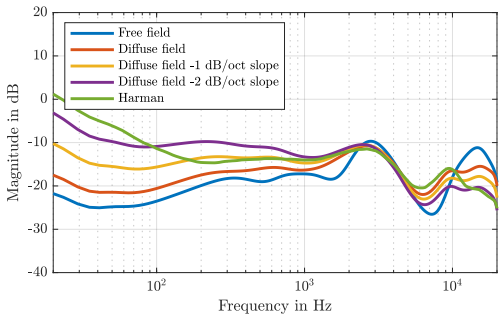
- Free field, Eq. (2),
- Diffuse field, Eq. (4),
- Diffuse field with $\gamma = -1$ dB/oct slope, Eq. (5),
- Diffuse field with $\gamma = -2$ dB/oct slope, Eq. (5),
- Harman Curve (obtained from [9]).



(a) HpTFs



(b) Target curves



(c) Equalization filters. Normalized to A-weighted level.

Fig. 1: Magnitude responses used in the process of designing the equalization filters.

The required HRTFs were obtained from measurements of the HATS conducted in the anechoic listening room "Wilska" at the Aalto Acoustics Lab, comprising 45 Genelec 8331 loudspeakers. To avoid the influence of reflections from equipment in the room, the measurements were truncated to about 5.2 ms. Fig.1(b) shows the different target curves.

The magnitude responses of the equalization filters were then designed according to Eq. 1. In addition, a fourth-order Butterworth bandpass with its passband between 20 Hz and 20 kHz was applied and minimum-phase filter responses were created by means of cepstral windowing. Fig.1(c) shows the resulting filter magnitude responses. Each filter was normalized according to the A-weighted magnitude, so that perceived loudness differences would not influence the experimental result.

4.1 Audio Material

For the first part of the experiment, which used stereo productions, excerpts from the following pieces were used:

- Steely Dan. (2000). Two against nature. On Two against nature [Album]. Giant Records/WEA,
- Jessel, L., Morrissey, J. J., Fennell, F. (Conductor), & Dallas Wind Symphony. (2015). Parade of the wooden soldiers [Arr. for wind ensemble]. On Dallas Wind Symphony sampler [Album]. Reference Recordings,
- Dave Brubeck Quartet. (1959). Take five. On Time out [Album]. Columbia Records.

The selection is similar to that in [1]. Therein, the same recording of the "Parade of the Wooden Soldiers" and a piece from the same Steely Dan album were used. For the spatial audio stimuli, there exist no common pieces yet, so we decided to use this year's winner in the Grammy category "Best Immersive Audio Album", which was the album "i/o" by Peter Gabriel, mixed by Hans-Martin Buff. Excerpts from the following three songs were selected:

- Gabriel, P. (2023). The court (H.-M. Buff, Dolby Atmos Mix). On i/o [Album]. Real World Records,
- Gabriel, P. (2023). This is home (H.-M. Buff, Dolby Atmos Mix). On i/o [Album]. Real World Records,
- Gabriel, P. (2023). Road to joy (H.-M. Buff, Dolby Atmos Mix). On i/o [Album]. Real World Records.

It is important to note that no additional acoustic space or reverb was added during the binauralization process. The binauralizations were created by first rendering the pieces to discrete channels that were then rendered using the SPARTA Binauraliser[1] [10] using its "11.x (7+4)" preset with additional LFE-channel which fullfills the standard for dolby atmos source positioning [11]. This decision was made to maintain an objective testing scenario, as well as to highlight the possibility of spatial audio integrating acoustic space into the production itself. However, it is likely that even more prominent use of spatial reverberation is possible than in the selected songs.

4.2 Procedure

The listening test was conducted using a MUSHRA app [12], which also allows for presenting a reference-free, MUSHRA-like test, as the one conducted here. A total of 15 participants took part in the test (13 male and two female). The three parts (timbral preference using stereo, and timbral and spatial preference using spatial audio) were presented in randomized order. The task was to rate the five different headphone equalization curves for preference while listening to the audio material. The questions had the following wording: "Please rank the stimuli regarding your timbral preference." and "Please rank the stimuli regarding your spatial preference (localization, externalization, reverb)," respectively.

A similar experiment was conducted in [13], comparing target curves for stereo and spatial material. However, their study only included the normal diffuse field and Harman Target curves, with spatial audio material created using stereo content and BRIRs.

4.3 Results

The results of the listening experiment are shown in Fig. 2. In the timbral preference for stereo playback, depicted in Fig. 2(a), the Harman Target was preferred over all other targets (Bonferroni-Holm-corrected Wilcoxon signed-rank test, $p \leq 0.013$), except for the diffuse-field curve with a 1 dB per octave slope ($p \approx 1$). The preference of the Harman Target was expected for the chosen songs, as it is the only curve among those tested that is actually tuned to the acoustics of a room. The diffuse-field curve with a 1 dB per octave slope was clearly preferred over the raw diffuse-field curve ($p \ll 0.001$). The free-field curve received the lowest ratings ($p \ll 0.001$).

The timbral preference of each curve was very similar when listening to spatial audio. Also for spatial preference, the diffuse field with a 1 dB per octave slope and the Harman Target curve received the highest ratings. Interestingly, the raw diffuse-field response was rated similarly ($p \approx 1$). As expected, the free-field curve as well as the diffuse-field curve with 2 dB slope added, are rated the worst here ($p \ll 0.001$).

5 Discussion

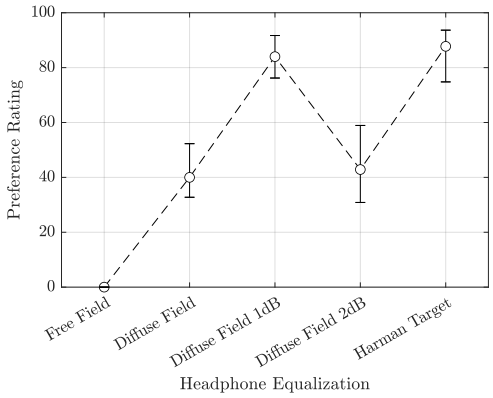
The results confirm the high preference for the Harman curve found in other experiments [1, 2, 13]. In general, using stereo content or binaural renderings of spatial audio content does not change the preference much. Interestingly, and perhaps counterintuitively, spatial qualities are influenced by the equalization curve; when asking for spatial preference, the raw diffuse-field equalization received higher scores than when asking for timbral quality.

Most importantly, however, the diffuse-field target with a 1 dB slope yields preference ratings similar to the Harman curve for both timbral and spatial preference and stereo and spatial audio material. While at the same time, unlike the Harman curve, which is derived from a specific room setup, the diffuse-field target is easy to replicate with different HATS. This makes it a promising candidate for a target curve.

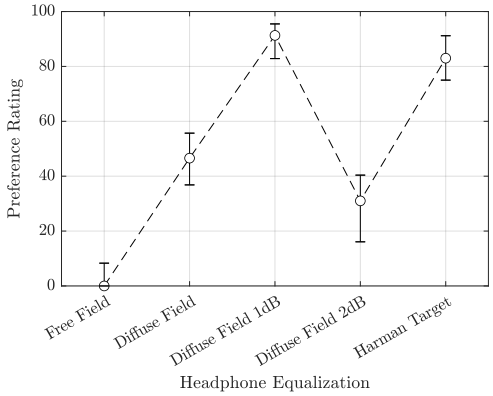
5.1 Reference and Preference

Being able to equalize a headphone to match a target does not mean that every headphone manufacturer will or should do so. A target curve like the proposed diffuse field with 1 dB per octave slope can be seen as a possible *reference* to compare other headphones against. After subtracting it (on a dB scale) from another headphone curve, the remaining curve could be seen as the unique characteristics of the headphones. Fig. 3 compares the Harman curve to the diffuse field with a 1 dB per octave slope curve in this way. After subtracting the two, what remains of the Harman curve

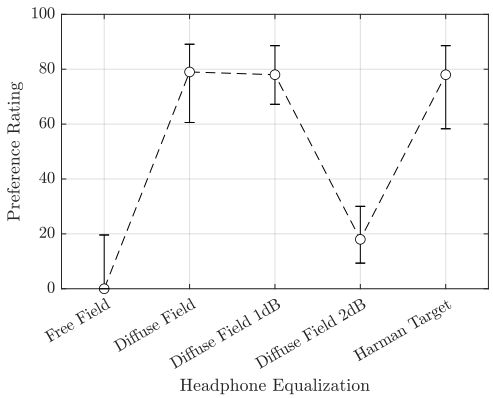
[1] freely available as part of the SPARTA plug-in suite `https://leomccormack.github.io/sparta-site/`



(a) Timbral preference, stereo content

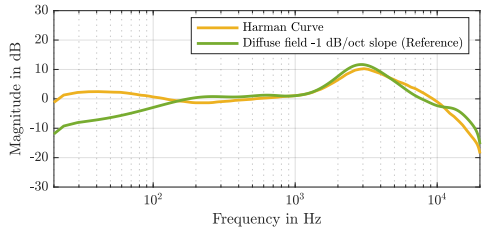


(b) Timbral Preference, Spatial content

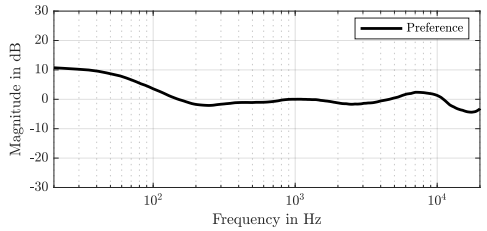


(c) Spatial preference, spatial content

Fig. 2: Results of the listening experiment: Median values and 95% confidence intervals.



(a) Harman and diffuse field with -1 dB per octave as reference



(b) Remaining "Preference" profile.

Fig. 3: When considering the diffuse field with -1 db/oct slope as a reference, the difference to a given curve, e.g., Harman, could be considered a preference curve.

is mainly the low bass boost. This can be seen as its unique preference profile.

Concerning such preference profiles, in spatial audio, a fully defined playback chain is theoretically possible when the acoustic space is integrated into the production. Thus, there is no technical reason to include preference. In stereo, however, there is always the undefined but necessary listening room and it is also introduced in the headphones domain when attempting to imitate a stereo listening room. Therefore, additional preference curves like the one implied by the Harman Target are easily justifiable in stereo audio playback. At the same time, the experiment conducted here did not show a clear benefit from it.

5.2 Low bass in the employed diffuse-field response

When comparing the diffuse-field curve used in this experiment to other diffuse-field measurements of the KEMAR HATS, it becomes apparent that our measurement has less low bass (< 100 Hz). This is probably due to the roll-off of the 8331 loudspeakers used for

measurements of the HRTFs and the truncation performed on the responses. In [9], the response at low frequencies is set to a flat line, avoiding the problem. Using such an approach, the obtained filters in the experiment would have slightly higher low bass, making the diffuse field with 1 dB slope even more similar to the Harman Curve. As both methods yield similar results already, it is unlikely that the outcomes would change much. However, this points to the fact that being able to use any ear simulator is only useful if high-quality diffuse field measurements are available for it.

6 Conclusion

The authors propose a method for resolving the conflict between the two playback philosophies, stereo and spatial audio, and suggest a headphone reference target curve that accommodates spatial audio reproduction: A diffuse-field curve with an additional spectral slope of -1 dB per octave. Using this curve as a reference would rid headphone analysis of the dependence on the specific room and the specific preferences encompassed in other semi-reflective field targets like the commonly used Harman Curve.

References

[1] Olive, S., Welti, T., and McMullin, E., "Listener preferences for different headphone target response curves," in *AES 134th Convention*, 8867, 2013.

[2] Olive, S., Welti, T., and McMullin, E., "Listener preferences for in-room loudspeaker and headphone target responses," in *AES 135th Convention*, 8994, 2013.

[3] "Audio Science Review, Headphone Amplifier Reviews & Discussion," Accessed: 2025-04-27, audiosciencereview.com/forum.

[4] "The Hangout," Accessed: 2025-04-27, graph.hangout.audio/headphones/?share=Harman_AE_OE_2018_Target,SR-X_Mk2.

[5] Møller, H., Jensen, C. B., Hammershøi, D., and Sørensen, M. F., "Design Criteria for Headphones," *J. Audio Eng. Soc*, 43(4), pp. 218–232, 1995.

[6] Zotter, F., *Analysis and Synthesis of Sound-Radiation with Spherical Arrays*, PhD Thesis, Universität für Musik und Darstellende Kunst Graz, 2009.

[7] Møller, H., "Hi-fi tests-with 1/3 octave pink weighted random noise," in *AES 47th Convention*, A-5, 1974.

[8] "Headphones.com," Accessed: 2025-04-27, https://headphones.com.

[9] "AutoEq Target Curves," Accessed: 2025-04-27, github.com/jaakkopasanen/AutoEq/tree/master/targets.

[10] McCormack, L. and Politis, A., "SPARTA & COMPASS: Real-time implementations of linear and parametric spatial audio reproduction and processing methods," in *AES International Conference on Immersive and Interactive Audio*, York, UK, 2019.

[11] "Dolby Atmos Home Theater Speaker Guide," Accessed: 2025-04-27, https://www.dolby.com/about/support/guide/speaker-setup-guides/7.1.4-overhead-speaker-setup-guide.

[12] "MUSHRA," Accessed: 2025-04-27, https://git.iem.at/rudrich/mushra.

[13] Engel, I., Alon, D. L., Scheumann, K., Crukley, J., and Mehra, R., "on the Differences in Preferred Headphone Response for Spatial and Stereo Content," *Journal of Audio Engineering Society*, (271-283), 2022.