

Headphone Essentials 5: Wrapping your head around the whole flat/ neutral/Harman headphone thing

© Dale Cotton, 2020. All rights reserved. Version 1.0.

You may re-circulate this document. You may not claim authorship or copyright to it.

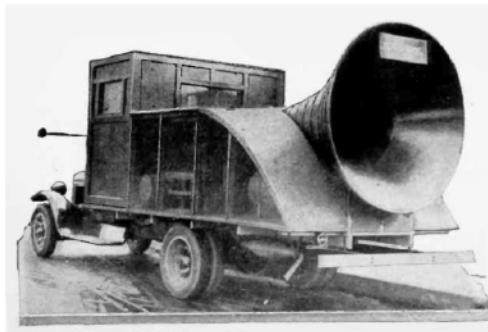
This document focuses on the frequency response or tuning of over-ear and on-ear headphones. In-ears have a related but different frequency response regime. This document is a summary and synthesis of available research from one headphone enthusiast to another. My aim is to present the material in a readable, non-technical manner. Thus inevitably, I gloss over many complexities. This analysis and interpretation of the available research is purely my own. I provide references to easily available source material so you can form your own opinions.

If you would like to start with more foundational background information on the nature of sound, see my tutorials *Basics of Musical Sound* and *Basics of Headphone Sound*, plus *The Skinny on Headphone Frequency Response Graphs* at <http://daystarvisions.com/Music/index.html>.

The search for accuracy



If you go to the store to buy a light bulb, chances are you're looking for one that produces white light. Red, purple, blue, green, etc. make for fun and games. But for boring old practicality it's hard to beat white. White light is a mixture of all other hues except black. White light is also the colour of sunlight during the greater part of the day. Illuminating things with daylight white light maximally reveals the colours of objects from apples to zebras. This happens because our eyes and brains evolved to make it so.



The whole flat/neutral/Harman headphone thing

Roughly the same situation exists in audio reproduction, except here we're dealing with sound instead of light (and reception instead of transmission). Testing consistently demonstrates that loudspeakers with a flat frequency response when in an echo-free environment are still the most accurate option for sound reproduction in a reverberant environment, such as a room or concert hall. Flat in this context means that the sound coming out of the speaker at all frequencies or pitches from deepest bass to shrillest treble is the same loudness as the electrical signal going in, like so:

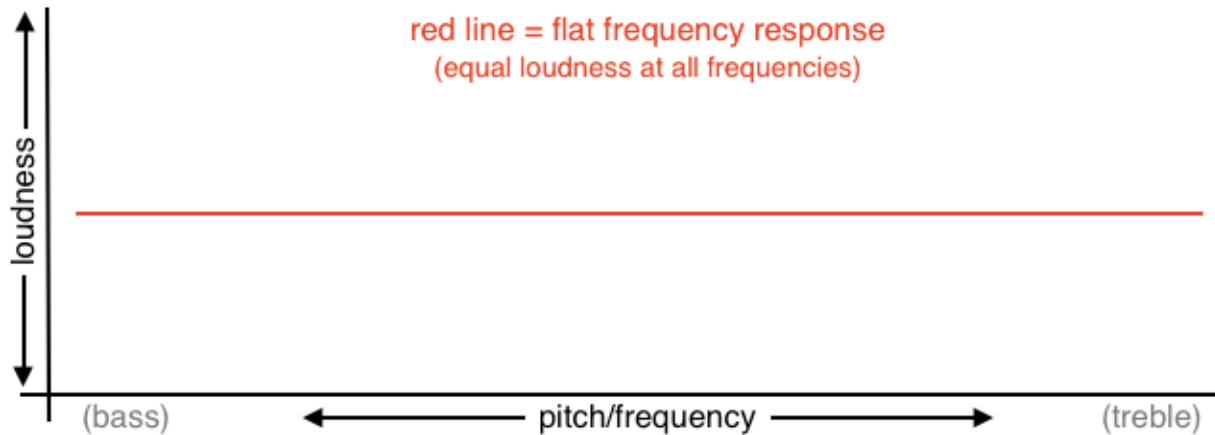


Fig. 1: flat frequency response (treble = beyond soprano)

Note: To get a better feel what this means you may want to go to <https://www.szynalski.com/tone-generator/>. Make sure the volume on your computer is initially set to a fairly low level, then press the space bar once to turn sound on, then move the slider fairly slowly and evenly left and right. If the sound changes in pitch but not in loudness you're hearing a flat response. If it gets louder and softer as you move the slider, that's a non-flat response. If possible, try this with more than one headphone or speaker to see how they vary (remembering to start at a low loudness level with each one). As is typical, the speaker built into my laptop doesn't produce any sound at all to the left of 100 Hz.

There is, however, a complication. When loudspeakers are used in a room, their sound output reverberates off the room's walls and other surfaces. The result somewhat emphasizes lower frequencies over higher frequencies. This is called *room gain*. It is this modified version of flat that our ears expect. There is no one version of this, but the general idea is:

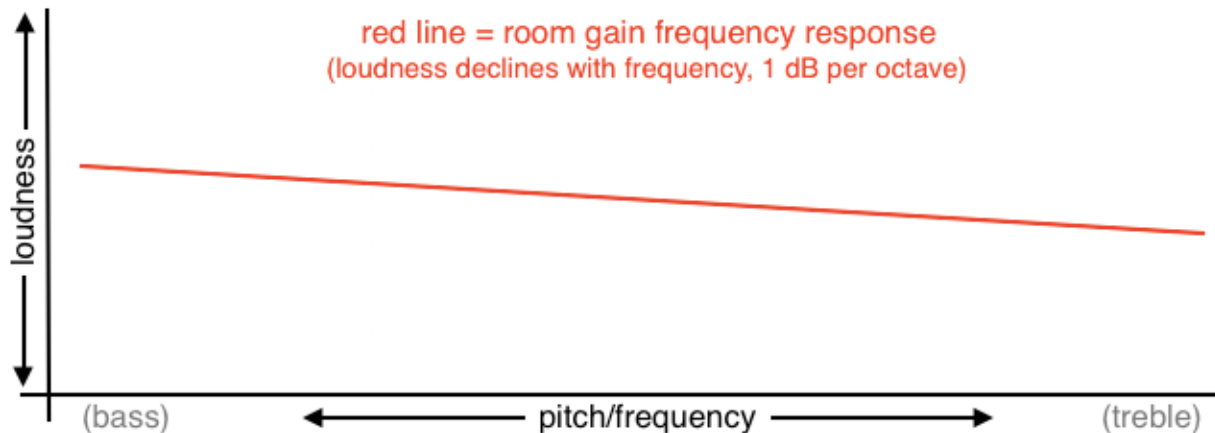


Fig. 2: room gain frequency response

The whole flat/neutral/Harman headphone thing

Headphones totally destroy this simple relation between electrical signal going in and the loudness of the sound coming out. We listen to loudspeakers by means of the sound pressure waves they generate in the air between the loudspeakers and our ears. Our brains were designed by millions of years of evolution to optimally interpret sound waves in a relatively large space, whether that be a room, a cave, a concert hall or the great outdoors. Plus, the particular shape of our outer ears and ear canals was designed to selectively amplify higher frequencies compared to lower/deeper frequencies. So if we measure the loudness of different frequencies at the eardrum instead of outside the ear, the results look something like this:

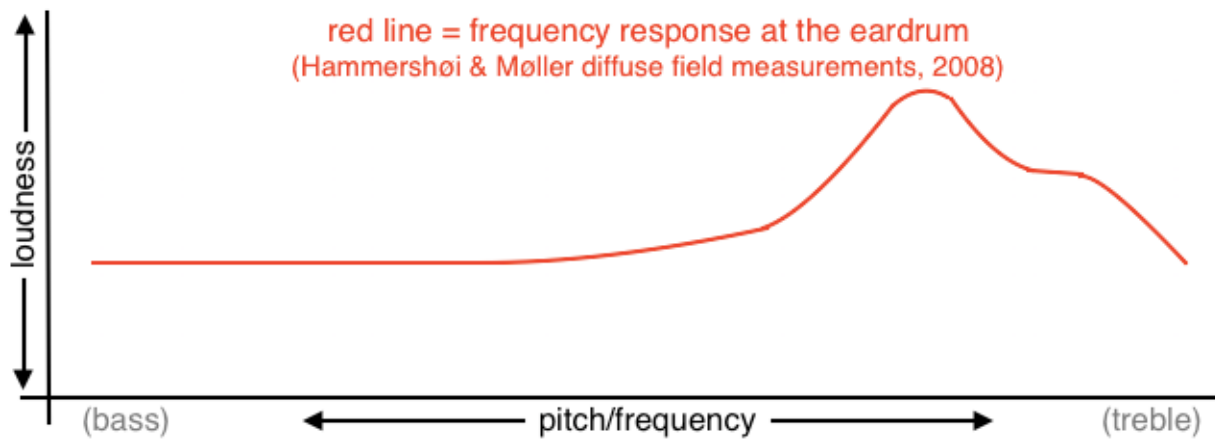


Fig. 3: eardrum frequency response (in an evenly reverberant room)

Note that this is what our ears do to a flat-line signal. What we actually hear after our brains process this signal is subjectively right back to a flat line again. In other words, our brains make use of the boost of the higher frequencies to detect faint sounds. But it does so without bringing that boost to our attention.



Unlike loudspeakers, headphones beam the sound waves directly into our ears. This by-passes most of the natural high frequency amplification (and additionally eliminates many of the cues our brains interpret for directionality). So a headphone that produced sound like Fig. 1 or Fig. 2 would sound drastically weaker in the high frequencies because the Fig. 3 boost is missing and our brains were not designed to compensate.

The whole flat/neutral/Harman headphone thing

The upshot is that headphones have to be deliberately designed to boost high frequencies to compensate for the problem they create by being coupled directly to the ear. So the frequency response of a pair of headphones should look more like Fig. 3 than Fig. 1. But not quite. At least for recorded music listening we also need to factor in a room gain element such as shown in Fig. 2. When we merge Fig. 2 and Fig. 3 we get:

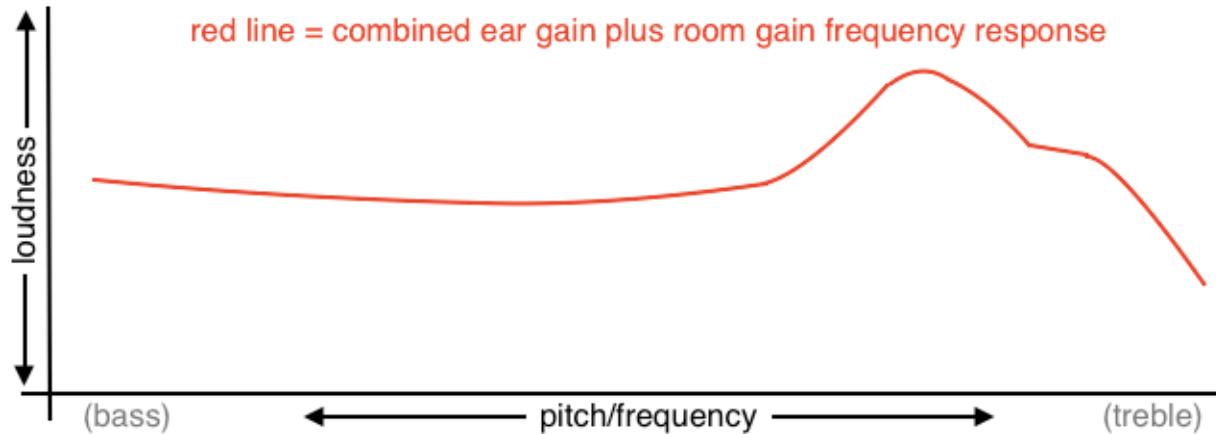


Fig. 4: eardrum + room gain (-1 dB per octave)
(compare Warren TenBrook: <https://www.innerfidelity.com/content/acoustic-basis-harman-listener-target-curve>)

Fig. 4 may not look all that different from Fig. 3. But the peak of Fig. 4 is crucially lower than that of Fig. 3. A headphone that transmits sound as shown in Fig. 3 would sound uncomfortably piercing in the high frequencies at otherwise comfortable loudness levels.



Apparently, the headphone company Sennheiser conducted similar research. Their venerable HD 600 model closely follows the Fig. 4 curve, except for some fall-off in the deep bass that is difficult to overcome with a standard speaker-cone type driver:

The whole flat/neutral/Harman headphone thing

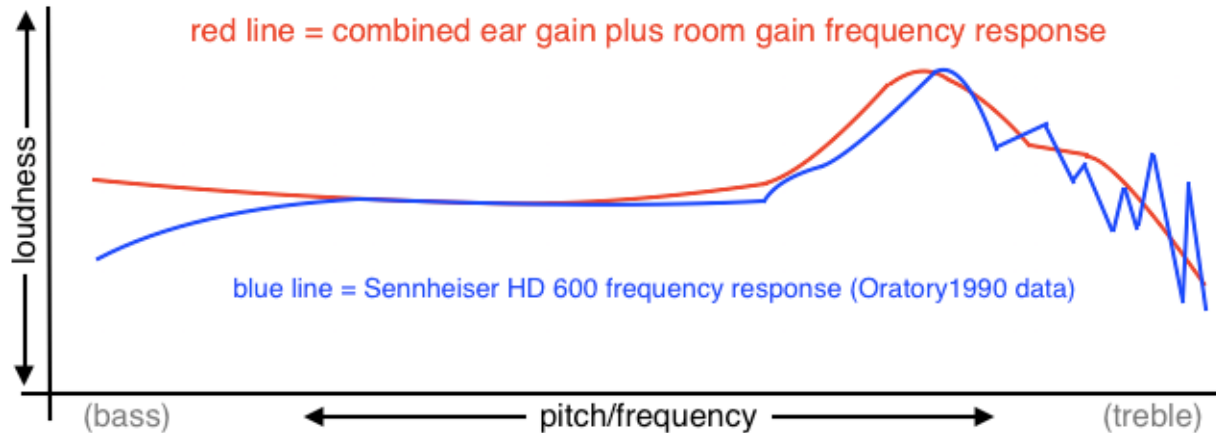


Fig. 4b: Sennheiser HD 600 compared to eardrum plus room gain
(source Oratory1990: https://www.reddit.com/r/oratory1990/wiki/index/list_of_presets)

In its product manual Sennheiser simply refers to the HD 600 as being tuned to a diffuse field response (Fig. 3). We see from Fig. 4b that in fact Sennheiser used a diffuse field response modified to correct for room gain. A true, unmodified, diffuse field response would resemble Fig. 3 and would have a much more prominent ear amplification peak. (The AKG K701 is a better example of a true diffuse field tuned headphone.)

So Fig. 4 shows us an equivalent of daylight white light for headphones. This is a practical tuning, or frequency response, one that more-or-less accurately reproduces the loudness aspect of recorded sound. However, there are still some further factors to consider. (Are we having fun yet?)

One factor I've noticed is that people who listen at high loudness levels are more likely to prefer a reduced ear amplification peak and higher frequencies as being less piercing. This also depends on the type of music involved since some instruments — for example, electric guitars, synth, and violins — that tend to have a lot of high frequencies in their sounds.



Another factor that impinges on Fig. 4 is the great variety of shapes and sizes people's ears come in (<https://link.springer.com/article/10.1186/s41935-019-0111-0>). Again, for loudspeaker listening this isn't an issue. Our brains know how to compensate for our own ear anatomy. But for headphone listening it very definitely is an issue. Fig. 4 may be close to accurate for a person with a hypothetical average human outer ear and ear canal, listening at some standard loudness. It's unlikely to be completely accurate for you. And the differences modify the shape of the ear amplification hump in Figures 3 and 4. (Technically, this is called the *head-related*

The whole flat/neutral/Harman headphone thing

transfer function or HRTF.) The differences also wreak havoc on the high frequency cues our brains use to localize the source of a given sound.

The upshot of all this is that, if you happen to want or need to hear an accurate frequency response from a pair of headphones, even Fig. 4 will only take you part way. There simply is no single frequency response graph that will be accurate for even the majority of individuals. Fig. 4 was:

1. based on Fig. 2, which *approximates* one of many possible room gain slopes
2. based on Fig. 3, which *assumes* equal reverberation from all directions (diffuse field)
3. based on a *reasonable* loudness level
4. based on a theoretical *average* human ear anatomy.

But peel yet another layer off the onion and yet another question emerges — does any of this matter? I noted above that your brain is designed to automatically compensate for the amplification resulting from the particular shape of your ears under normal circumstances. It is also designed to compensate for the often wildly varying acoustics of one room versus another and for any enclosed space versus the outdoors. Given a chance, your brain will also compensate for a certain amount of departure of a given headphone from the frequency response that would be accurate for your ear shape and listening loudness. Your brain creates a model of how your headphone departs from accurate at each frequency. It then uses that model to correct for those inaccuracies.

There seem to be limits to this, particularly at higher frequencies, on how much inaccuracy your brain can correct for. And switching from one model headphone to a different model at least temporarily defeats this ability. But given a chance your brain will learn to compensate for the inaccuracies of a reasonably accurate headphone. Audio enthusiasts call this familiar phenomenon *brain burn-in*.

Given all the above, something like Fig. 4 can be used as a reference point when shopping for a headphone when the goal is accurate recorded music listening. Your brain will overcome some departure from Fig. 4. Too much departure from Fig. 4 may well result in a persistent frequency response annoyance, such as boominess, muddiness, sibilance or shrillness.

The search for the fun factor



Fig. 4 would be a good target for headphone designers to shoot for if the goal is to please the perhaps 3 or 4 percent of potential customers who want accuracy and who listen at moderate loudnesses. Presumably, the other 96 or 97 percent are looking for something more like this:

The whole flat/neutral/Harman headphone thing

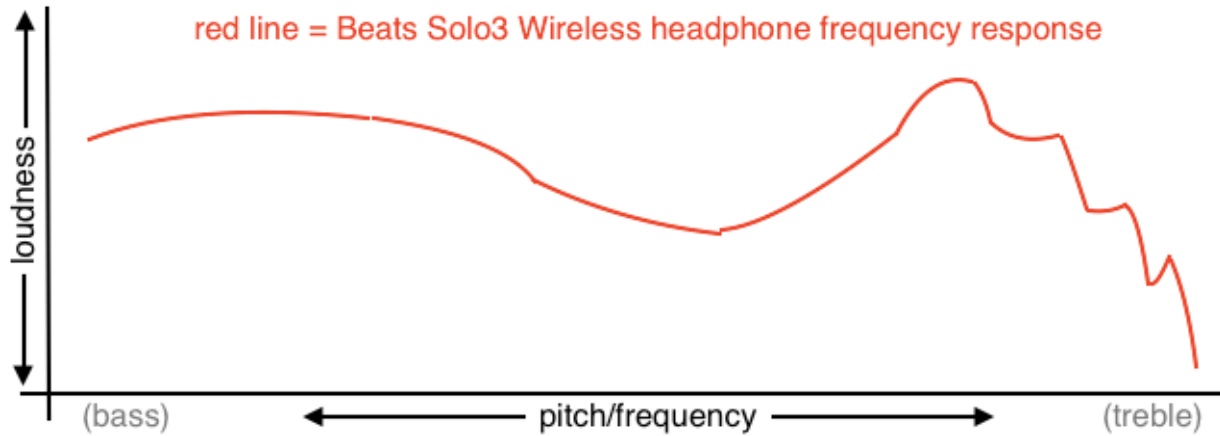


Fig. 6: example consumer preference curve
(Source Rtings.com: <https://www.rtings.com/headphones/1-3-1/graph#413/3182>)

And many headphones actually measure more or less like Fig. 6. This is what consumer demand dictates. Instant party time. Great for high energy music. The problem is that it fails miserably when listening to something like a jazz performance or a love ballad, where voice and instrument quality actually matters:

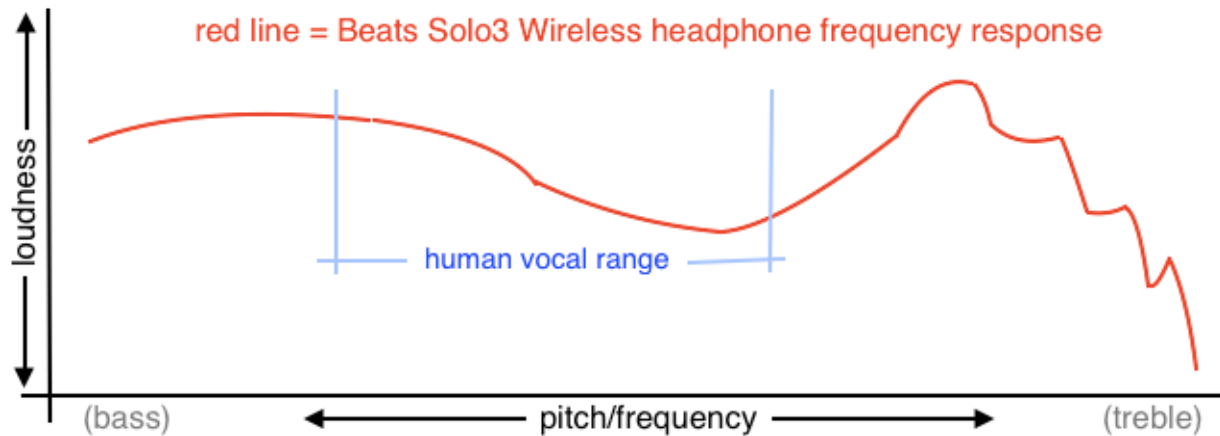


Fig. 6b: vocal range

Compare the portion of the red line in Fig. 6b between the two blue lines with the same portion of Fig. 4, which is very nearly flat. It's generally still possible to recognize a singer's voice even when distorted as extremely as shown in Fig. 6b. It's also possible to recognize yourself



The whole flat/neutral/Harman headphone thing

stretched and pinched in a fun house mirror. Whether that's what you want to look like is another matter. (Of course, I exaggerate. It's really more of a muffled sound, like listening to someone at a distance through a wall of mattresses.)

Sean Olive, Todd Welt and colleagues at Harman International (not pictured above) set themselves the challenge of finding a single frequency response curve for headphones that would be most enjoyable to the greatest percentage of people ... but, importantly, remain enjoyable while listening to a wide variety of music. Their result looks something like this:

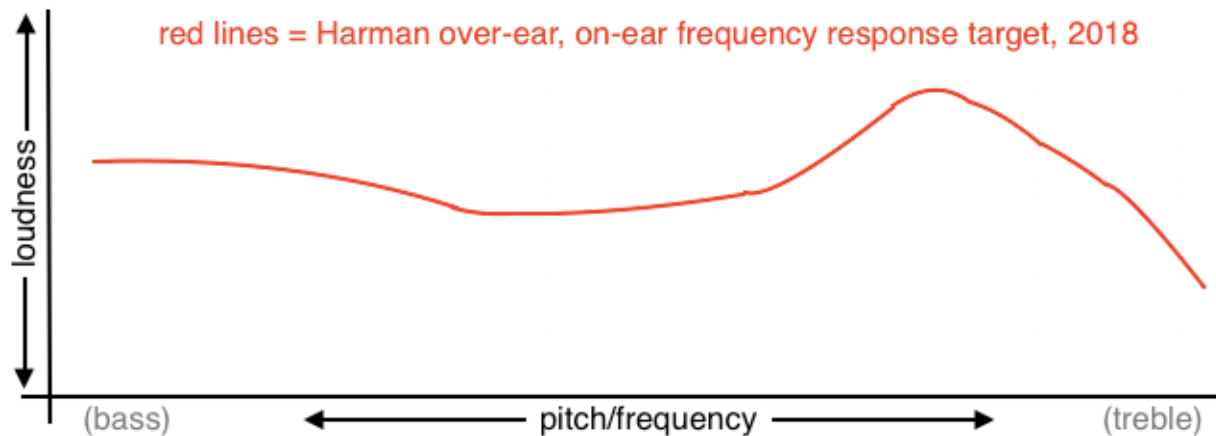


Fig. 7: Harman target

(Source: <https://www.innerfidelity.com/content/harman-tweaks-its-headphone-target-response>)

As a nod to consumer preference there is definitely a bass elevation. But it gently comes to an end where the bass range does — before the mid/vocal range. The ear amplification peak is somewhat lower than in Fig. 4 to accommodate loud listening levels. This curve was achieved by having about 250 people of different ages, sex, nationality and auditory training freely modify the the bass and treble response of a headphone while listening to a variety of music samples.



However, it's not that everyone came up with the same curve. Instead there was a range of variance something like this:

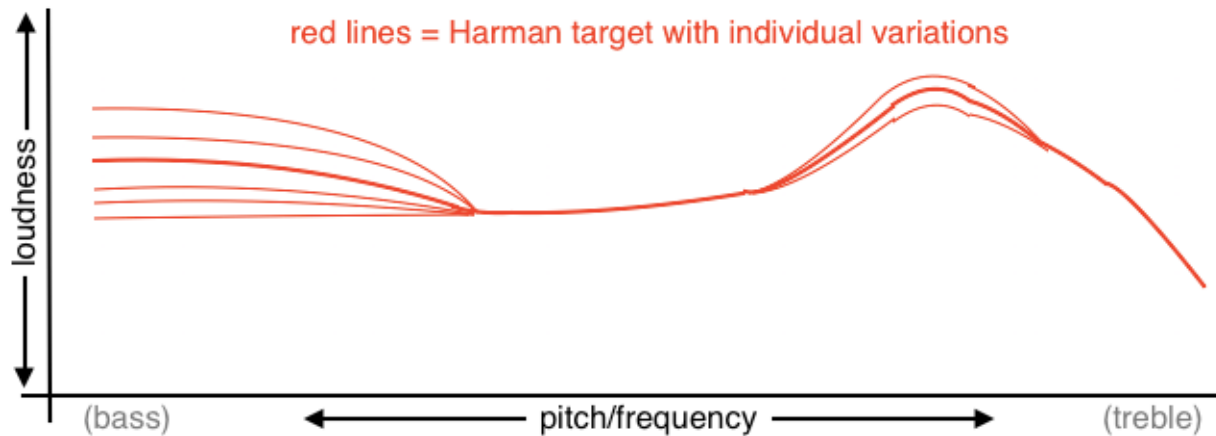


Fig. 7b: Harman target plus some individual variations
(Source Oratory1990: <https://imgur.com/AwZEg5R>, Sean Olive: <https://seanolive.blogspot.com/2015/11/factors-that-influence-listeners.html>)

(Yes, at least one individual actually preferred a slight bass reduction.) If the neutrality curve of Fig. 4 is a daylight white light bulb, the Harman target curve in Fig. 7 is a warm white light bulb. In other words, a compromise between strict practicality and the dictates of the human auditory pleasure response.

Some take-aways

In the world of 2-channel loudspeakers, listening to an accurate loudspeaker is easily described by its flat-line frequency response when room effects are removed. But those very room effects are the great challenge for 2-channel. Under ideal circumstances they can be described as a simple down-sloping line.

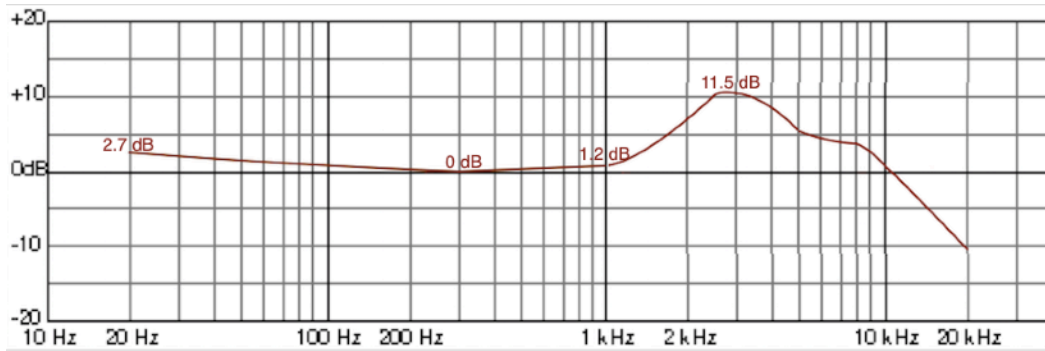
In contrast, in the world of over-ear and on-ear headphones, room effects are virtually eliminated. The challenge is re-creating a natural ear amplification. The combination of this re-created average ear amplification curve plus a room gain downslope is a reasonable approximation of the frequency response of an accurate loudspeaker system in a hypothetical ideal room. That said, there can be no one truly correct/accurate headphone tuning due to loudness and ear-shape issues. Even so, brain adaptation can and often does compensate for these issues.

In any case, accuracy does not drive headphone sales. A headphone with market appeal typically has a highly boosted bass response that extends well into the mid-range frequencies, distorting vocal and instrumental sounds. The Harman frequency response curve is a carefully researched compromise between strict accuracy and consumer preference. But by its very nature it is unlikely to be the best compromise for any single consumer.

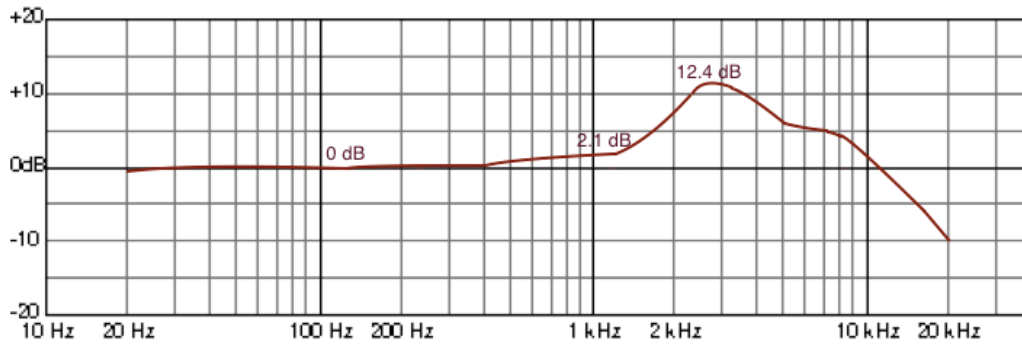
Geek stuff: obscure technical material

For those with a more technical bent, here is the eardrum + 1 dB/octave room gain graph plotted to the common logarithmic scales and labelled:

The whole flat/neutral/Harman headphone thing



An earlier study by the acoustics technology firm Brüel & Kjær showed some roll-off in the bass. When combined with the diffuse field curve we would get:



This may well be closer to what Sennheiser was targeting with the HD 600.

Here are the 1/3-octave decibel numbers I used for plotting these graphs:

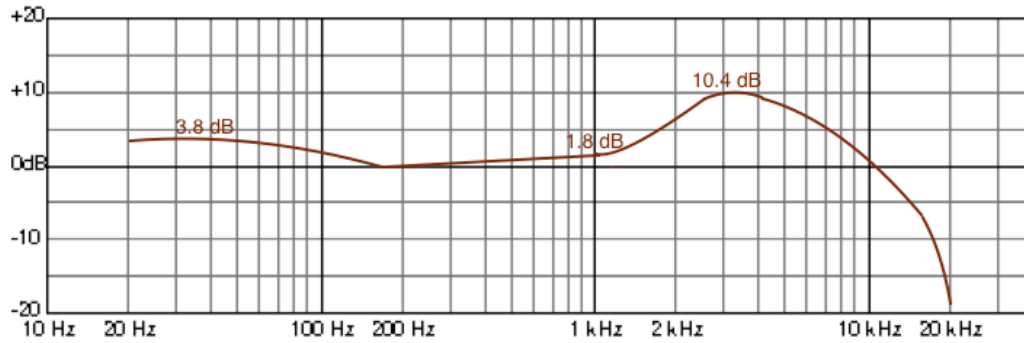
1/3 octaves in Hz	20	25	31.5	40	50	63.5	80	100	125	160	200	250	315	400	500
Diffuse field ear	0	0	0	0	0	0	0	0	0	0.5	0.8	1	1.2	1.5	2
Brüel & Kjær spkr	2.5	2.7	2.9	3	3	3	3	2.9	2.8	2.7	2.5	2.2	2	1.8	1.6
DF + BK	-0.3	-0.0	0.1	0.2	0.2	0.2	0.2	0.1	0	0.4	0.5	0.4	0.4	0.5	0.8
-1dB/octave slope	5	4.7	4.5	4.1	3.8	3.5	3.2	2.8	2.6	2.1	1.9	1.5	1.1	0.8	0.5
DF + -1dB/octave	2.7	2.4	2.2	1.8	1.5	1.2	0.9	0.5	0.3	0.3	0.4	0.2	0	0	0.2

630	800	1K	1.25K	1.6K	2K	2.5K	3.15K	4K	5K	6.3K	8K	10K	12K	16K	20K
3	3.5	4	5	8	11	15.5	15.5	13	10.5	10	9.5	7	4	0	-3.6
1.5	1.1	0.9	0.6	0.3	0	-0.3	-0.5	-1.1	-1.2	-1.6	-2	-2.2	-2.6	-3.1	-3.6
1.7	1.8	2.1	2.8	5.5	8.2	12.4	12.2	9.1	6.5	5.6	4.7	2	-1.4	-5.9	-10
0.2	-0.8	-0.5	-0.3	-1.2	-1.5	-1.7	-2.1	-2.4	-2.7	-3	-3.4	-3.8	-4.2	-4.7	-5
0.9	0.4	1.2	2.4	4.5	7.2	11.5	11.1	8.3	5.5	4.7	3.8	0.9	-2.5	-7	-10.9

The first two rows are estimates based on visual inspection of the published graphs and have no greater authority than that.

Finally, here is my rendition of the 2018 Harman target curve to the same scale for comparison:

The whole flat/neutral/Harman headphone thing



Some source material and further reading:

5. D. Hammershøi and H. Møller, "Determination of Noise Emission from Sound Sources Close to the Ears", *Acta Acustica*, Vol.94 No.1 (January 2008)
6. Brüel & Kjær, "Relevant loudspeaker tests in studios in Hi-Fi dealers' demo rooms in the home etc." (1974), <https://www.bksv.com/media/doc/17-197.pdf>
7. Wikipedia, "Head-related transfer function", https://en.wikipedia.org/wiki/Head-related_transfer_function
8. Tyll Herstens, "Headphone Measurements Explained - Frequency Response Part One", <https://www.innerfidelity.com/content/headphone-measurements-explained-frequency-response-part-one>
9. Oratory1990, <https://www.reddit.com/r/oratory1990/wiki/index>