Headphone Essentials 5:

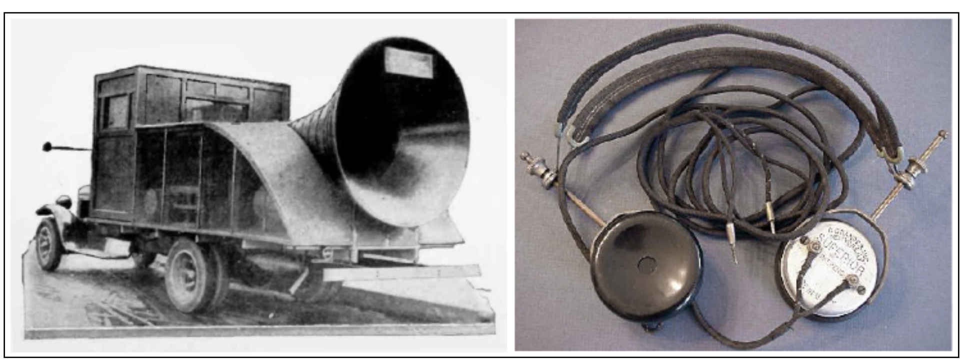
Wrapping your head around the whole flat/neutral/Harman thing

Note: this document is part of a instructional series. If you would like to start with more foundational information on the nature of sound and related material, see *Basics of Musical Sound*, *Basics of Headphone Sound* and *The Skinny on Headphone Frequency Response Graphs* at Headphone Essentials.

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 and was written by a headphone enthusiast for other headphone enthusiasts. 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.

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.



Left: horn loudspeaker truck, 1929, first commercial headphone, Brandes Superior Matched Tone, circa 1919-21. Photo credits: left: C. Sterling Gleason, right: John Davidson, License

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:

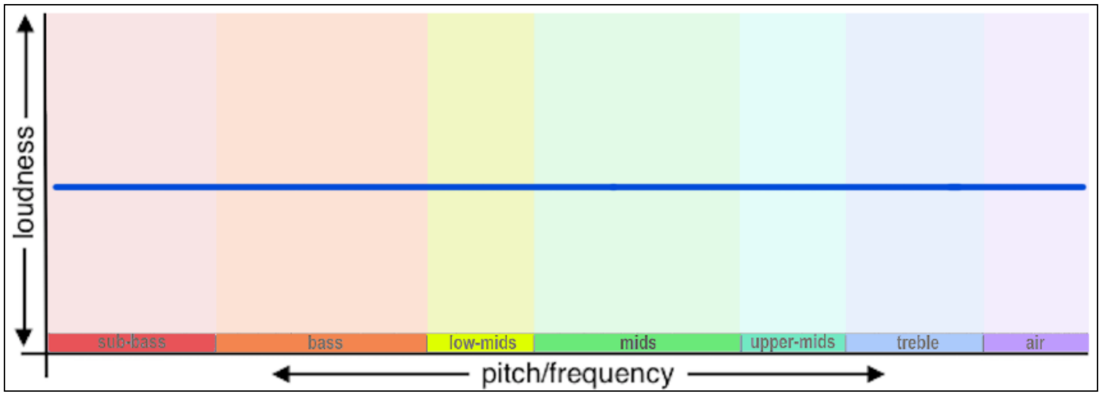


Fig. 1: flat frequency response (treble = beyond soprano)

Note: In Fig. 1 I've use the colour spectrum to suggest the spectrum of sound from bass to treble. To get a better feel what this means you may want to go to the Online 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.

Fig. 1 is how an ideal loudspeaker should measure. But by the time the sound reaches our ears that same straight-line flat measures like this:

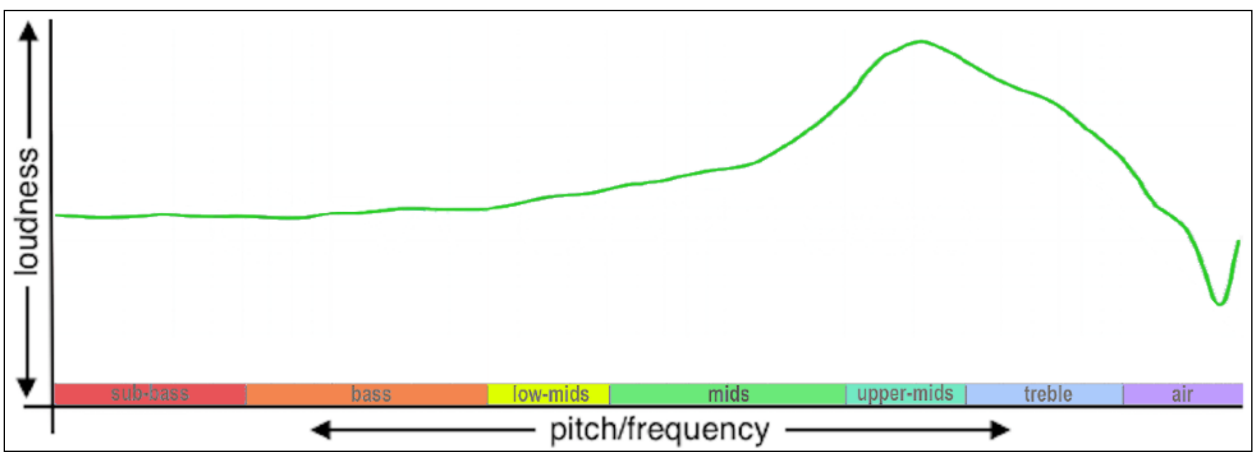


Fig. 2: flat speakers measured at the eardrum (source: Harman International)

The huge rise in loudness in the higher frequencies on the right is due to a combination of the shapes of our upper bodies, heads, outer ears and ear canals selectively amplifying those frequencies. The shape of our ears was designed by evolution to boost our hearing in an area that was presumably critical for survival. But our brains also evolved to process that enhanced sound in a way that retains the extra information but reduces its perceived loudness back to the Fig. 1 flat line again before we have conscious experience of it. This signal processing magic happens behind the scenes and in real time.

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 when listening to recorded music. There is no one version of this — even mix and mastering studio acoustics are notoriously all over the map — but the general idea is:

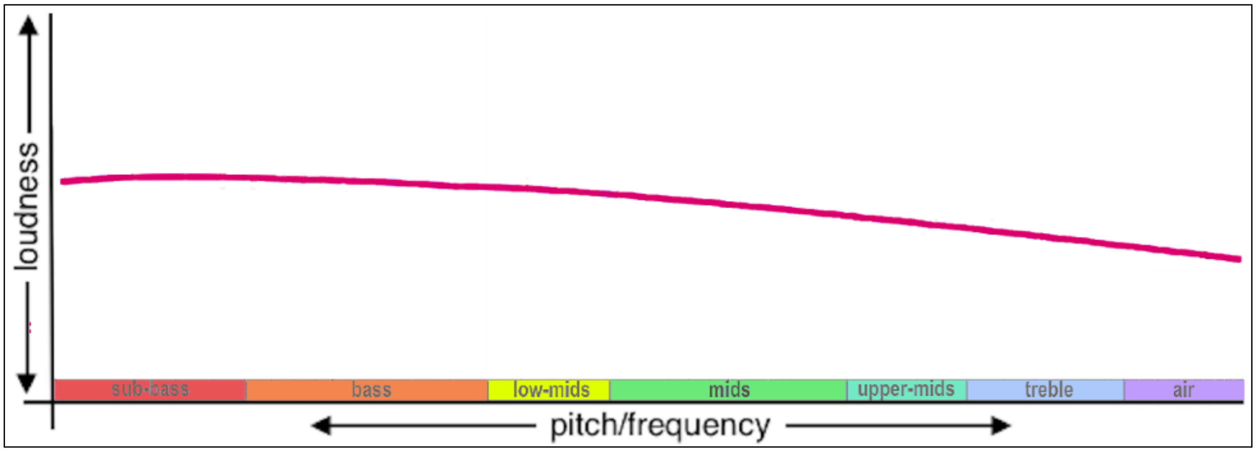


Fig. 3: room gain frequency response (B&K Optimum HiFi)

The curve in Fig. 3 is the result of a study by acoustic research firm Bruel & Kjaer of what an ideal room response should be. This is a balance between the acoustics of concert halls with that of smaller venues. But of course any given room in any given house will have wildly varying acoustics. Loudspeaker enthusiasts spend lots of time and money to try to tame the most egregious faults of their listening rooms. Since recordings are at least traditionally engineered to sound properly in a typical room of a house with typical stereo speakers in it, the sound measured at the eardrum will be a fusion of the two inputs — speaker output plus room gain:

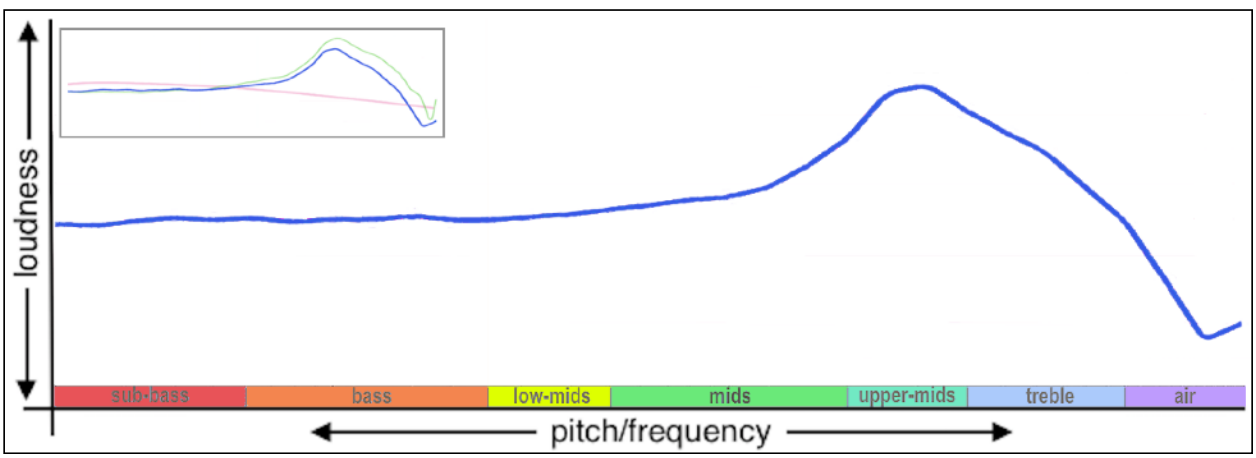


Fig. 4: Combined speaker + room gain (OHN-9)

The result is a kinder, gentler version of the unmodified speaker's output. The highs are significantly tamed over those of the loudspeaker on its own (11½ decibels, down from 15½ decibels). Let's make sure we have all this clear before moving on: Fig. 4 is what happens when an ideally ruler-flat-measuring loudspeaker produces sound in an ideally reverberant room (measuring like Fig. 3). Then the resulting sound is picked up by a statistically average pair of human eardrums situated inside statistically average anatomically correct human ears.

We care about this for the simple reason that headphones don't work like loudspeakers in a room. Speakers in a room create sound pressure waves that fill the air in the room before entering the ear from a complex variety of directions. Headphones are little speakers in a miniature “room” that shove sound straight at the ear canal. This is called *direct coupling*. It eliminates all or most of the selective high frequency amplification of the human ear and has essentially zero room gain of its own. So for a headphone to reproduce a recording the headphone needs to artificially add in the high frequency boost shown in Fig. 4. There are many tools in the headphone engineer’s toolkit to achieve this, but the results are always imperfect.

The Sennheiser HD 600 has become legendary for its natural-sounding playback since its release just before the 21st century began. If we compare the HD 600’s measured frequency response with the curve in Fig. 4 we get:

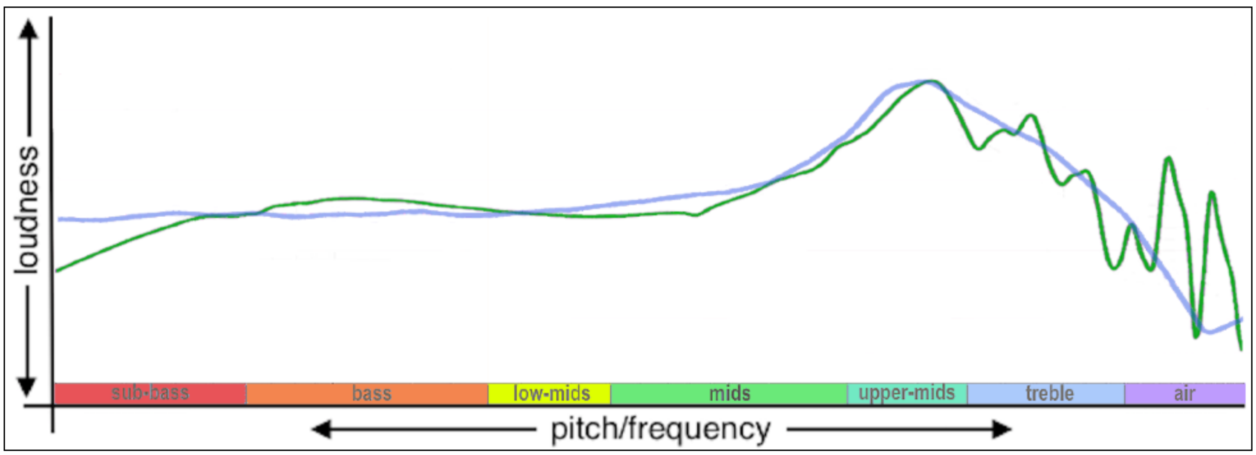


Fig. 5: Sennheiser HD 600 (green) compared to the Fig. 4 target response curve (pale blue)
(source: based on Oratory1990: reddit.com)

The loudness drop on the far left, followed by the slight overshoot in the bass, is the result of the engineering team having to cope with the same challenge that cone loudspeaker engineers always face. In a stereo speaker pair each box typically houses two or three speaker cones of different sizes to cover the entire frequency range. In an open-back headphone a single speaker cone is used in each ear cup. But the challenges of keeping enough loudness in the lowest bass frequencies as well as sculpting the ear gain rise in the higher frequencies are huge. Few headphones come this close. Another factor is that frequency response is only one of several auditory goals the designers are trying to optimize. Spatial recreation (sound stage and imaging), fine detail retention, dynamic range and dynamic impact and minimal distortion are also critical factors.

In its product manual Sennheiser simply refers to the HD 600 as being tuned to a **diffuse field response**. *Diffuse field response* is the technical term for the sound measurement at the eardrum when the sound has equal loudness at all frequencies and at all directions. A typical diffuse field curve looks similar to, but not identical to the loudspeaker curve in Fig. 2. Importantly, the ear gain rise in the high frequencies is every bit as high as in Fig. 2. We see from Fig. 5 that, given Sennheiser used a diffuse field response, they also modified it to correct for room gain as we did in Fig. 4. A true, unmodified, diffuse field response would mean that the green line in Fig. 5 would reach as high as the blue line in Fig. 2.

For another example, here is the (original) Focal Clear:

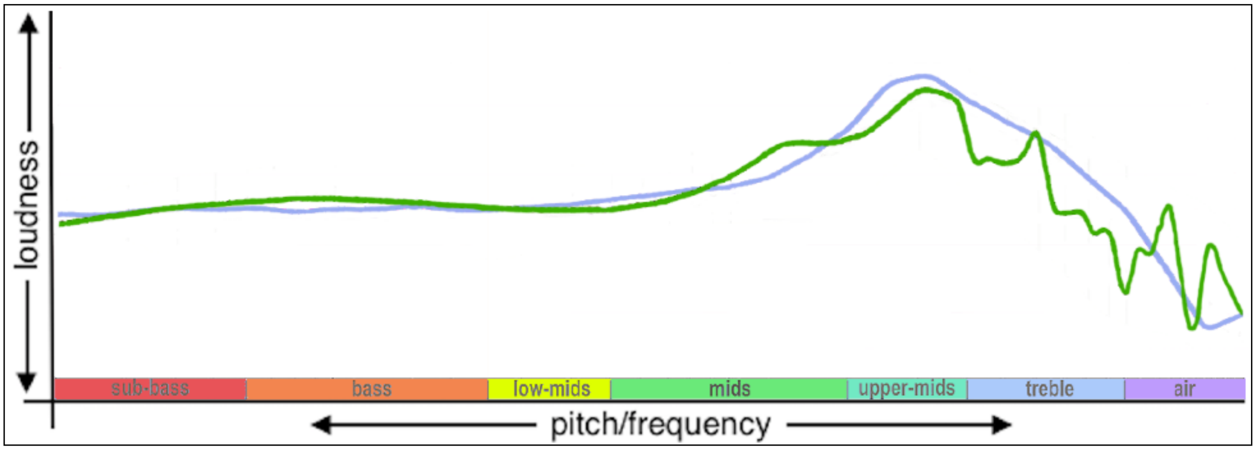


Fig. 6: Focal Clear (green) compared to Fig. 4 target response curve (pale blue)

The Clear departs a bit more from the neutral target curve, particularly in the treble range. We'll dig more deeply into this high frequency divergence in the Alternate approaches section below.

We could go on making such comparisons, but the result would be the same. The more a headphone is recognized by the headphone enthusiast community as being especially neutral

or faithful sounding to recordings, the more they resemble Fig. 4. The criticisms people make, such as the HD 600 having a slight thickness in the upper bass and lower mids, show up as departures from Fig. 4 when we overlay their measurements with Fig. 4. Headphones that people report as less faithful to the source show greater departures from Fig. 4. We can't keep calling this Fig. 4. Prosaically, it's one of many over-ear/on-ear headphone neutrality candidates. Since no one knows how many previous candidates exist and it's hopefully at least close to the final version, I'm going to dub it OEHP Neutral #9, or OHN-9 for short.

So OHN-9 shows us at least the approximate equivalent of daylight white light for headphones. OHN-9 gets us in the right ball park of a practical tuning, or frequency response, for a headphone — one that should come very close to accurately reproducing the loudness aspect of recorded sound. OHN-9 is based on the speaker measurements of one research team (Olive and Welty at Harman International). It would be nice to have it validated or refined by further research, but there is no reason to expect more than a few percent change to result. Ditto with the Brüel & Kjær room gain research that resulted in Fig. 3.

However, there are still some further factors to consider. (Are we having fun yet?)

Another factor that impinges on OHN-9 is the great variety of shapes and sizes people's ears come in. And this is true for the ear canal as well as the more obvious outer ear. Again, for loudspeaker listening this isn't an issue. Each person's brain knows how to compensate for his/her own ear anatomy.

Fig. 7: typical differences in human ear shape (source: Wikipedia Commons)

But for headphone listening it very definitely is an issue. OHN-9 may be fairly close to accurate for a person with a hypothetical average human outer ear and ear canal, listening at some reasonable loudness. But we see from Fig. 7 how widely divergent ear shapes are.

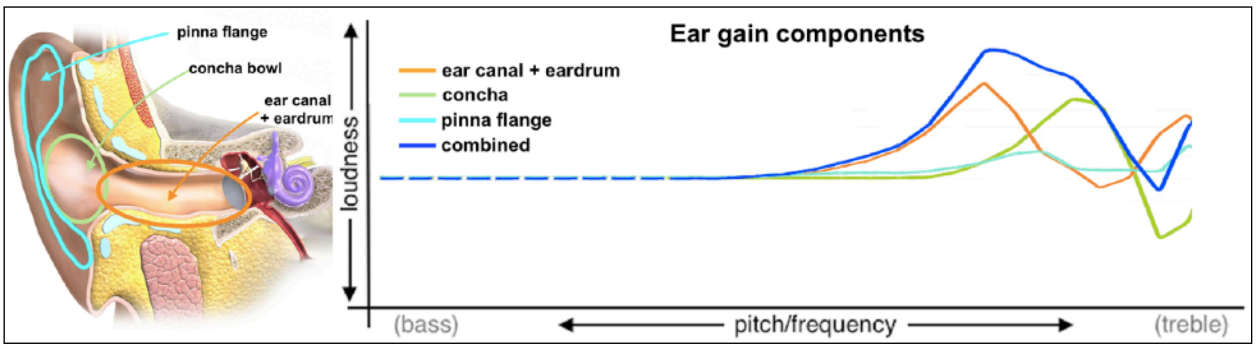


Fig. 8: contribution to ear gain from ear regions (source: Wikipedia & InnerFidelity)

Fig. 8 shows how much three main regions of the ear contribute to the total ear gain. If your ear canal contributes more than an average amount but the concha and pinna of your outer ear contribute less, then the shape of the blue line total ear gain in Fig. 8 — as well as Figures 2 and 4 — will differ accordingly. Your particular total ear gain is the main component of what is called your *head-related transfer function* or HRTF. (Your head, neck and upper body also have an acoustic effect, but taken together they have less effect on the over-all shape of the curve.)

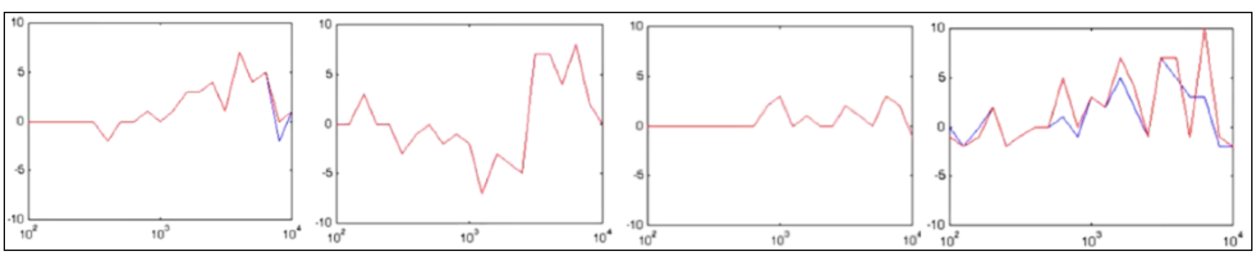


Fig. 9: insert microphone measurements of four different individuals wearing same HD 600 headphone (source: David Griesinger, includes audio)

These individual differences also affect the high frequency cues our brains use to localize the source of a given sound. So the same headphone will produce a different soundstage and imaging for different people.

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. All this very much applies to live music and loudspeaker listening.

But for headphones with their direct coupling all bets are off. Given a chance, your brain *may* 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. Not a lot seems to be known about this, but the idea is that your brain may create a model of how your headphone departs from accurate at each frequency that falls in the frequency bands it has evolved to manipulate. It would then use that model to correct for those inaccuracies. This is part of a field of study called psychoacoustics. The Harman research we'll be looking at below indicates only a small variation in personal preference across the entire ear gain region that Figs. 8 and 9 tell us should be the most variable.

Anecdotally, however, there seem to be serious limits, particularly at frequencies outside the vocal range, on how much inaccuracy your brain can correct for. And switching from one model headphone to a different model can at least temporarily defeat this ability pretty thoroughly. But given a chance, many find their brains will learn to compensate for the inaccuracies of a reasonably accurate headphone. Audio enthusiasts call this familiar phenomenon *brain burn-in*. Headphone enthusiasts are also quite clear that no amount of brain burn-in is going to make a piercing treble spike or a shouty mid-range play nice.

Given all the above, something like OHN-9 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 OHN-9 in some areas. Too much departure from your personal version of OHN-9 may well result in a persistent frequency response annoyance, such as boominess, muddiness, sibilance or shrillness.

The search for the fun factor

So OHN-9 might be a good target for headphone designers to shoot for if the goal were to please the perhaps five percent of potential customers who want accuracy and who listen at moderate loudnesses. Presumably, the other 95 percent are looking for something more like this:

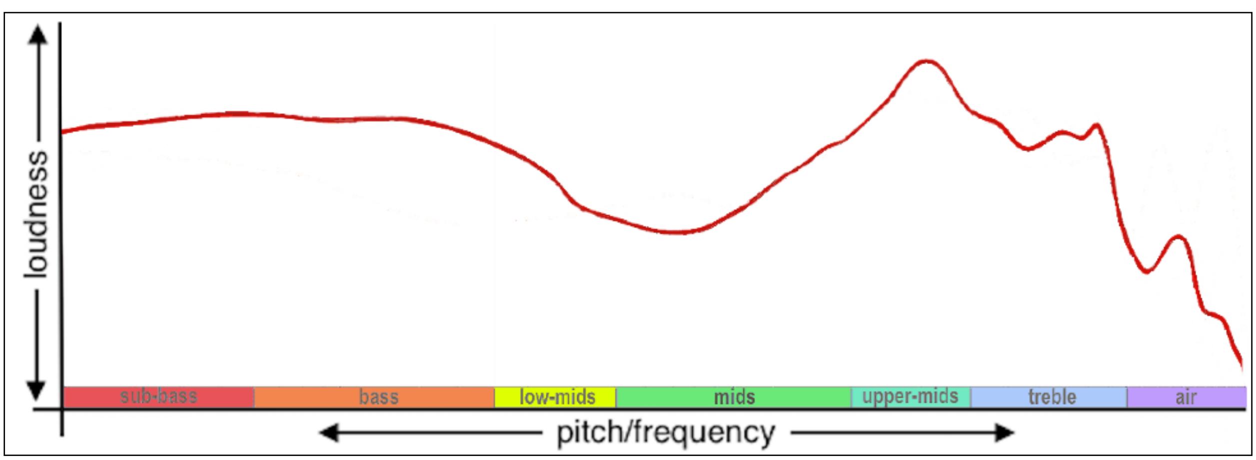


Fig. 10: example consumer preference curve. (based on rtings.com graph)

And many headphones actually measure more or less like Fig. 10. This is what consumer demand dictates. Immediate grab. Instant party time. 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:

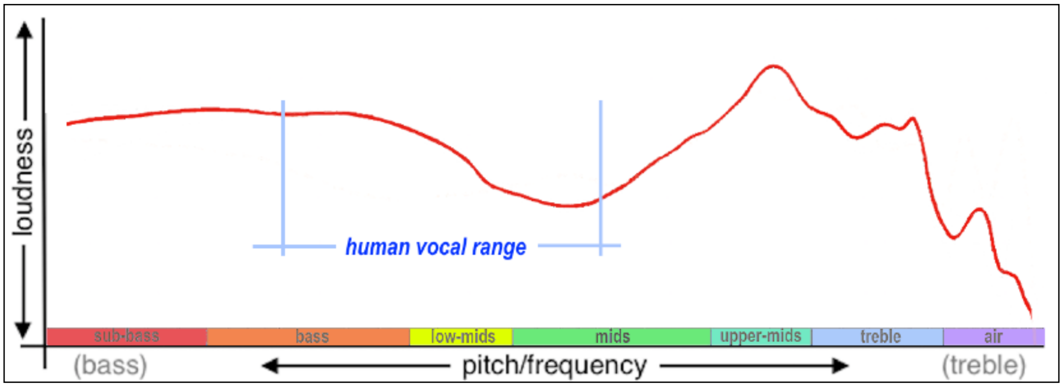


Fig. 10b: vocal range

Compare the portion of the red line in Fig. 10b between the two blue lines with the same portion of OHN-9, which actually rises slightly rather than plummeting. It's generally still possible to recognize a singer's voice even when distorted as extremely as shown in Fig. 10b. It's also possible to recognize yourself stretched and pinched in a fun house mirror.

Whether that's what you want to look like is another matter. (Of course, I exaggerate. Aficionados simply call it recessed mids and no big deal.)

Sean Olive, Todd Welty 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 like this:

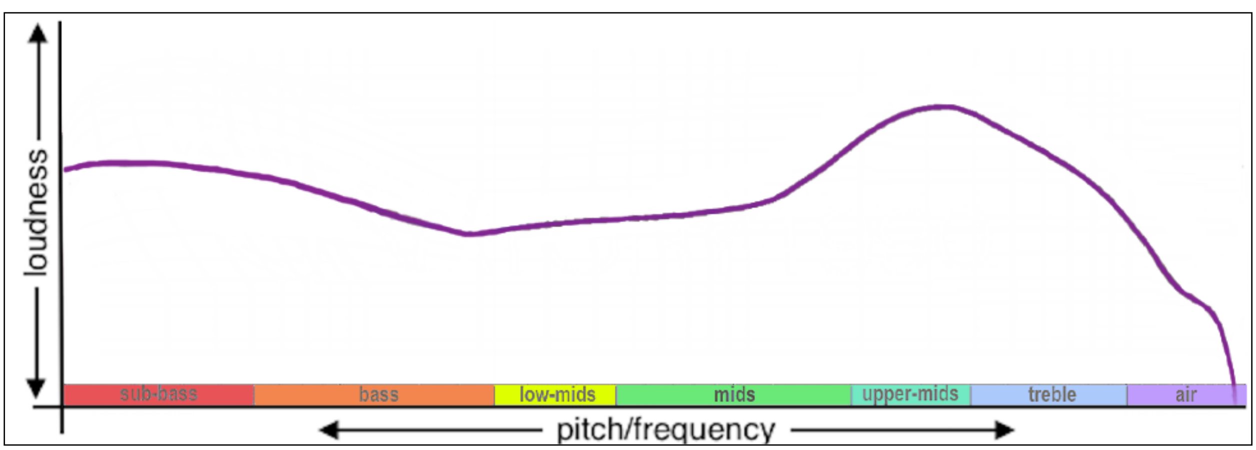


Fig. 11: Harman target. Source: InnerFidelity.com.

As a nod to listener preference there is definitely a bass elevation. But it gently tapers off where the bass range does — before the mid/vocal range. Also, the ear amplification peak is slightly lower than in OHN-9 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 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:

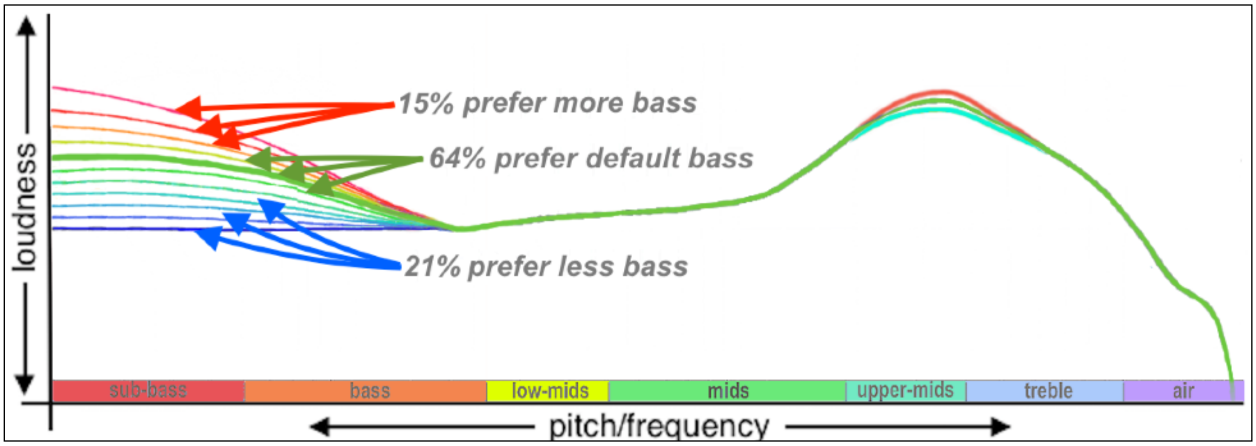


Fig. 12: Harman target plus some individual variations. Source Oratory1990, Sean Olive.

(And, yes, at least one individual actually preferred a slight bass reduction.) The purple line curve gets all the attention, but the research is more nuanced than just that, as we see in Fig. 12. Fig. 8 tells us there is no anatomical basis for this range of bass preference. Anatomically, it should be the right half of the graph, not the left, that shows all that variance. But if you're trying to manufacture headphones that sell, Fig. 12 is what you need to pay attention to.

Circling back to the initial analogy, if the neutrality curve of OHN-9 is a daylight white light bulb, the median Harman target curve in Fig. 11 is a warm white light bulb. In other words, a compromise between strict practicality and the dictates of the human auditory pleasure response.

However, it's not that the default six decibel rise of the standard Harman target is without consequences. The presence of that much bass effects the total tonal *balance* of a headphone's (or loudspeaker's) sound signature. This is easily heard when listening to a dense orchestral passage in which bass, mid-range and high-pitched instruments are all sounding at the same time. No live professional orchestral performance is going to have the sound that comes from the Harman target's balance. It also skews the bass/mids/treble balance a solo pianist is working so carefully to achieve. But for the smaller ensembles typical of popular music the effect is more easily ignored. Nevertheless, Sean Olive himself laments that every headphone would ideally come with a bass adjustment knob.

Some take-aways so far

In the world of 2-channel loudspeakers, an accurate loudspeaker is easily described by its flat-line frequency response once room effects are removed. But those very room effects are the great challenge for 2-channel.

In contrast, in the world of over-ear and on-ear headphones, external room effect variations are virtually eliminated. The challenge is in re-creating a natural eardrum response. The combination of this re-created loudspeaker plus room gain is at least a good approximation of the frequency response of a hypothetical perfect loudspeaker system in an hypothetical ideal room. That said, there can be no one truly correct/accurate headphone tuning, if only due to personal ear shape variance. Even so, brain adaptation can and often does at least partially compensate for some departure from the ideal.

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 in its default form it is a compromise that will satisfy about two-thirds of potential consumers.

Alternate approaches

Early on, first the free field, then the diffuse field ear measurement approaches were co-opted as tuning regimes for headphones. But this ignored the tempering effect of room gain. We've explored the linear speaker plus room gain approach, but are there other possibilities?

Alternative 1: Harman linear

Ironically, simply leaving off the bass rise from the Harman target creates what many will find to be a satisfyingly neutral target curve, with an ear gain peak just slightly lower than OHN-9's. And, as we saw back in Fig. 7b, eliminating the bass elevation is a perfectly valid aspect of the Harman target research. 21% find satisfaction with a lower level of bass loudness, including some fanatical fraction who opt for no elevation at all.

Planar magnetic driver headphones are very popular among headphone enthusiasts. Few planar magnetics have a bass elevation, yet many of their owners feel no need to change that. A Harman linear target is close enough to OHN-9 to be in contention for the ISO standards crown.

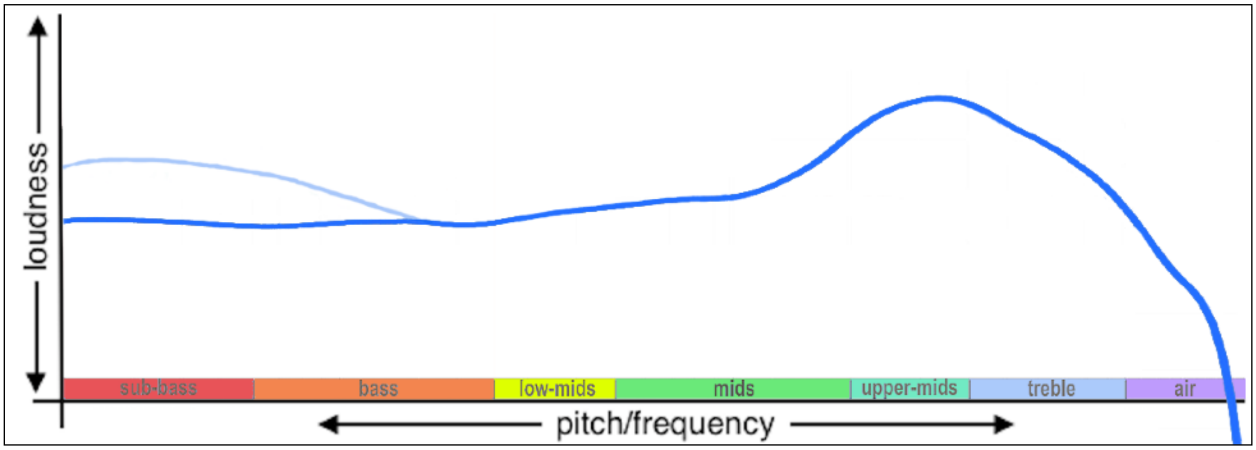


Fig. 16: alternative 1 — 2018 Harman 2018 target with linear bass

(**Note**: Closely related to the Harman linear target is the Oratory1990 Optimum Hifi target. Unfortunately, I can't find enough information about the thinking behind it at this time to put it in further context.)

Alternative 2: Average 24

Above, we used the Sennheiser HD 600 and the Focal Clear as a quick check on the validity of OHN-9. However, over time Sennheiser moved on from the HD 600, first to the HD 650, then to the HD 660S:

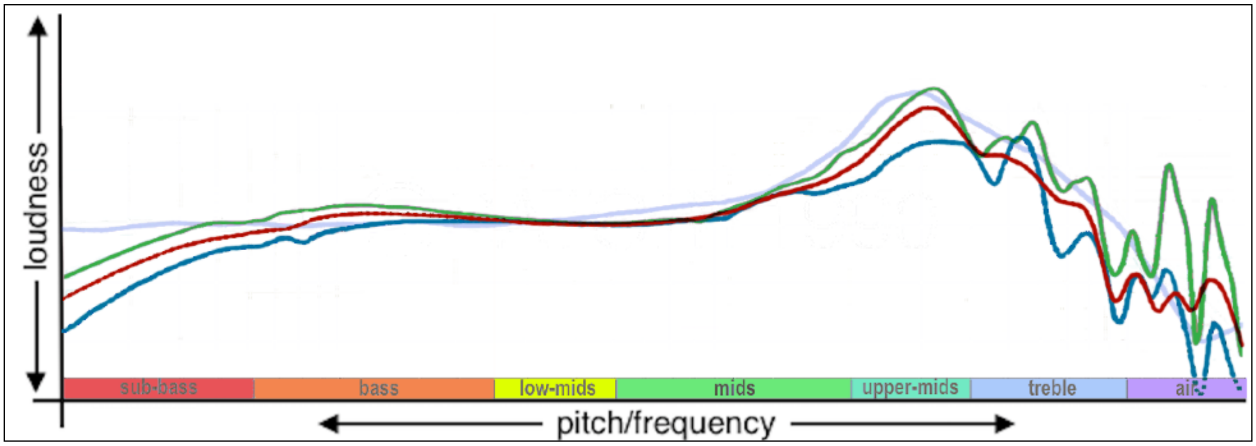


Fig. 13: Sennheiser HD 600 (green), HD 650 (red), HD 660S (blue) and OHN-9 (pale blue) (source: HeadPhoneDatabase.com)

We see the progression toward an ever-lower ear gain peak. Similarly with Focal's two neutral-leaning models:

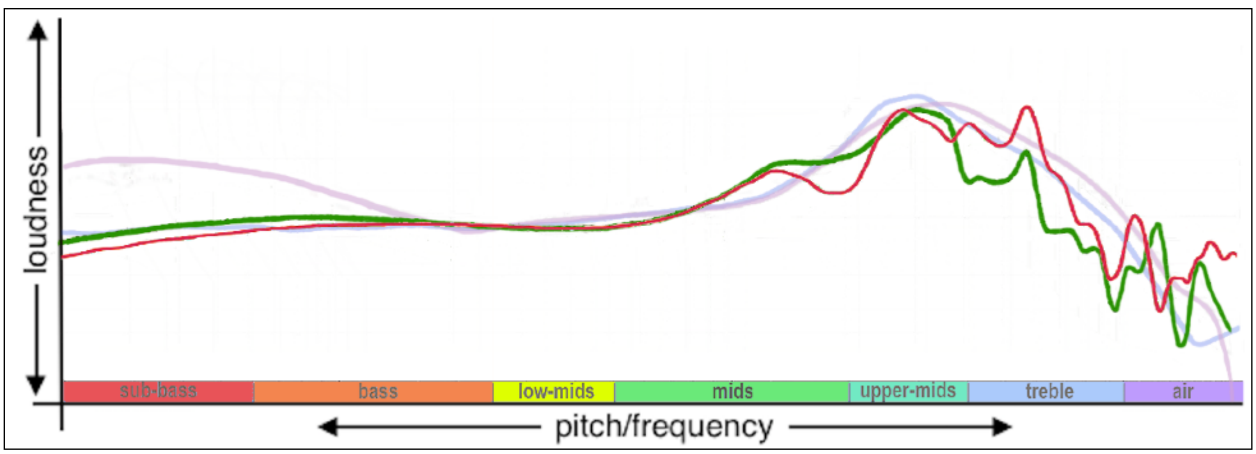


Fig. 14: Focal Utopia (orange), Clear (green), OHN-9 (pale blue) and Harman target (pale purple) (source: HeadPhoneDatabase.com)

In Fig. 14, I've introduced the Harman target to show how much more closely they match it's ear gain rise than that of OHN-9. But either way we see an undershoot.

Let's pursue this a little further. For Fig. 15, below, I took the average of the frequency responses of 24 of the most "neutral" headphones I could find, but this time including those with a Harman-like elevated bass. Namely:

- AKG K371*
- Audeze LCD-1
- Audeze Mobius*
- Audio-Technica M40x
- Audio-Technica M70x
- Audio-Technica R70x
- Aurora Audio Borealis
- beyerdynamic DT-880/250
- beyerdynamic TYGR 300R*
- Dan Clark Aeon 2 Noire*
- Focal+Drop Elex
- Focal Celestee*
- Focal Clear
- Focal Utopia
- Hifiman Sundara (2020 rev)
- NAD Viso HP50*
- Oppo PM-3*
- Philips Fidelio X2*
- Sennheiser HD 560S
- Sennheiser HD 600
- Sennheiser HD 650/6XX
- Sennheiser HE1
- Warwick Sonoma Model 1*
- ZMF Auteur

The eight starred headphones in the list have at least a partial Harman-like bass boost. (I've subsequently found a few more, such as the Fostex TX-00 Ebony, but adding them wouldn't materially change the results.

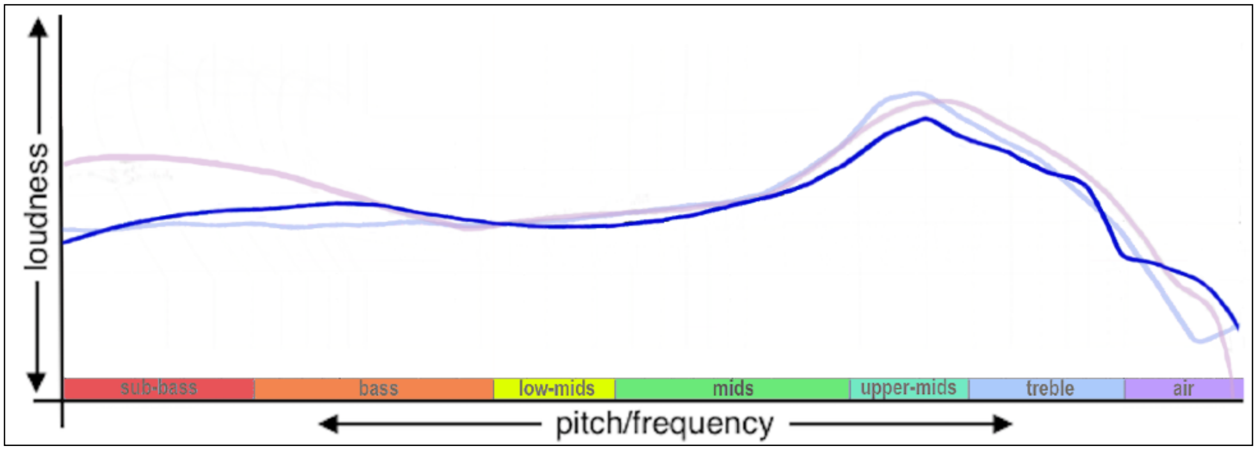


Fig. 15: alternative 2 — average frequency response of 24 neutral headphones (blue), OHN-9 (pale blue) and Harman (pale purple)

Forgive me for slipping into full geek mode for a minute: With the average max bass of 1½ dB and an average ear gain peak at 9 dB, we can see that manufacturers are putting their money on an undershoot of the Harman target curve's ear gain rise. While there is a rise between the 250 hertz start of the midrange and 1 kilohertz, it starts late at 500 hertz. Interestingly, in terms of left/right position, the ear gain peak splits the difference between the 2.7 kilohertz maximum point of the Etymotic and the Fig. 3 curves versus the 3.5 kilohertz maximum of the Harman target. The peak is followed by a slight scoop between 3 and 5 kilohertz that we see so consistently employed by manufacturers. Then there is a sharp drop-off from 8 to 10 kilohertz which supports the theory that a drop in this area is an important localization cue.

Yet, the Harman research shows a wide cross-section of people in various demographics preferred the somewhat higher ear gain peak of the Harman target. I suspect the undershoot of the Harman target ear rise we see in Fig. 15 is at least in part due to the difficulty in matching the Harman target's ear gain area precisely enough. By undershooting it I suspect designers are preferring to steer clear of the inevitable complaints that this treble spike or that treble spike is piercing. This is due to something called ear canal resonances, which differ from person to person both in their strength and in their location along the frequency response axis.

So, regardless of what OHN-9 or the Harman target indicate, Fig. 15 may be the practical neutrality target, given today's manufacturing technology for purely passive (non-DSP) headphones.

Alternative 3: community chaos
The 2-channel and home theatre community is filled to the brim with individuals creating their own signature room gain curve. The headphone enthusiast community is divided between between diffuse field traditionalists (they actually mean diffuse field + room gain), Harman

target proponents (either the original 2013 version or the final 2018 version) and the pox-on-all-your-houses, I-like-what-I-like majority. But there are also a number of influencers promoting their own neutrality candidates, whether based on theory or on sheer subjective preference. I'm sceptical that any of these I've seen so far will pass an eventual ISO standards litmus test. But I've been wrong before…

Wrap-up

So what is a neutral/accurate frequency response for a headphone? Depending on your purpose it could be different things:

- For a theoretical neutral reference based on the ISO 711 standard ear, the OHN-9 Olive & Welti loudspeaker + B&K room gain target is arguably as close as we have at this point.
- For personal use where maximum accuracy is essential, you would need to have your own head-related transfer function measured using insert microphones.
- For personal use in which you simply want to hear things as closely matched to the intention of the recording but not knowing your HRTF, start with OHN-9 or the linear bass Harman target or similar, then use EQ to adjust bass level and ear gain peak level to remove any sense of excess or deficiency.

For listening pleasure with a focus on hearing the recording as intended but without a lot of technical fuss, you'll simply have to gamble on which of the 24 headphones listed above you want to take a chance on. None match any target curve precisely, although the AKG K371 comes very close to matching the Harman target. You can use the graphing tool on headphonedatabase.com to compare various headphone models to both the standard Harman target and the linear bass Harman target (Oratory Optimum HiFi being very nearly identical to that). Comparing to any of the other targets discussed above is problematic. Perhaps print a graph, then draw or trace the target curve on top (Fig. 11). Some examples:

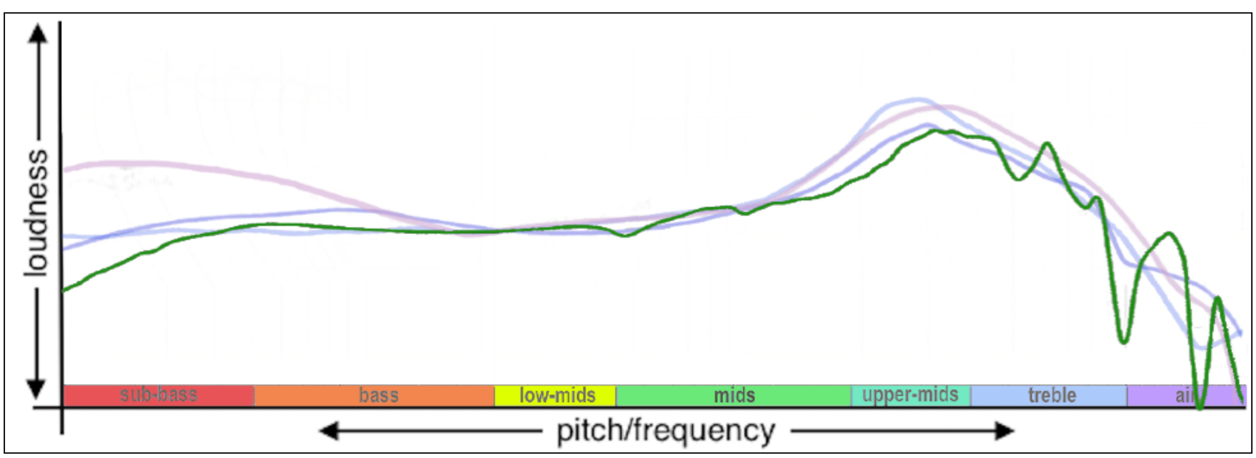


Fig. 17: Hifiman Sundara (2020 revision, green) with some target curves for comparison

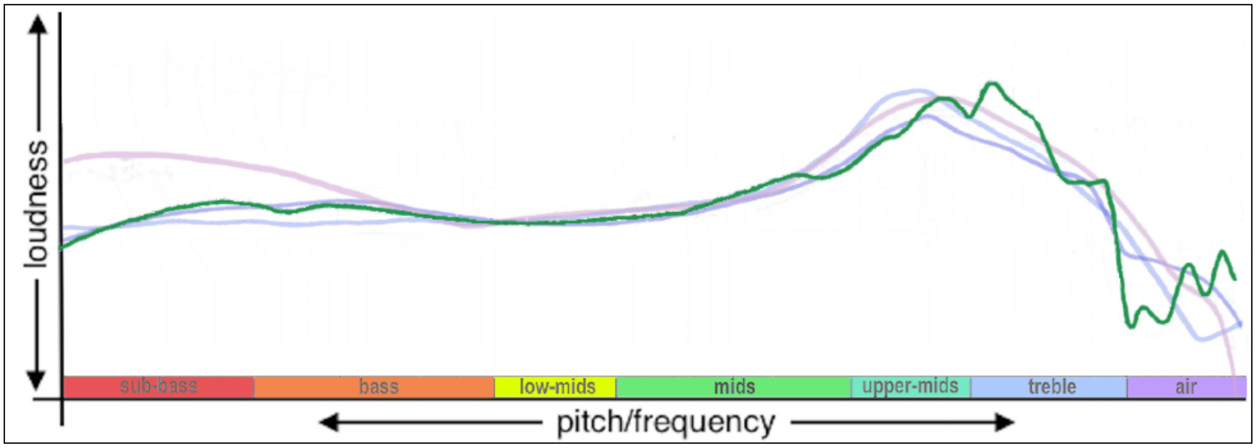


Fig. 18: Sennheiser HD560S (green)

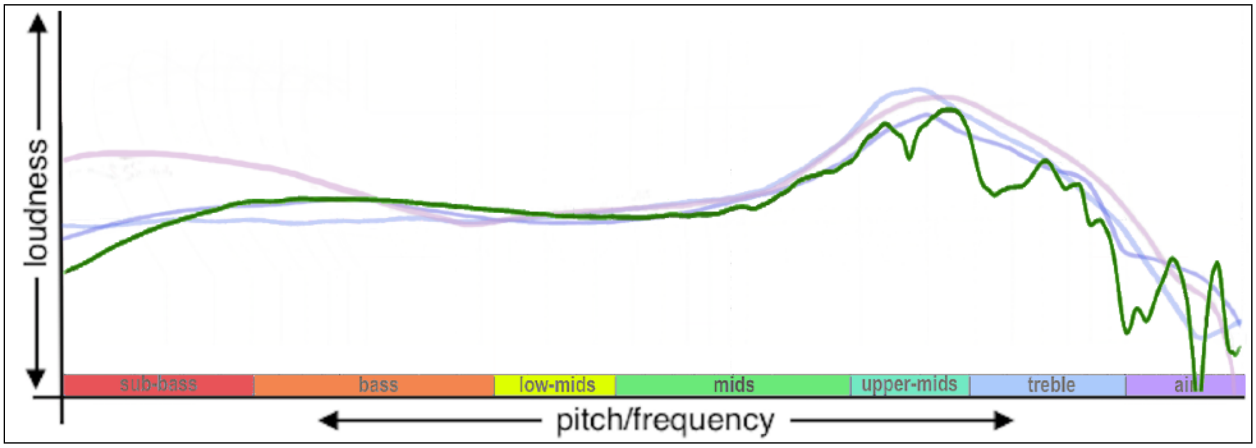


Fig. 19: Audio-Technica ATH-R70x (green)

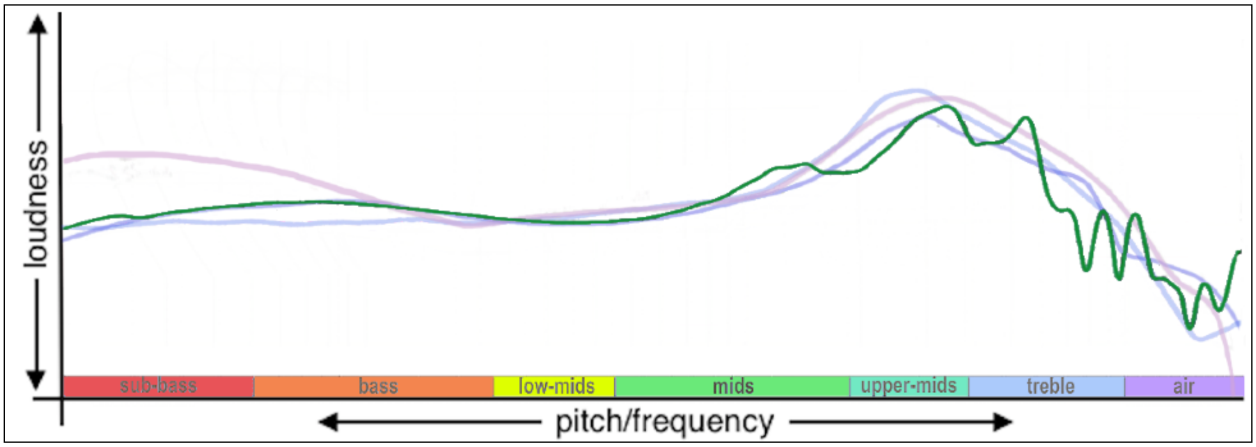


Fig. 20: Focal+Drop Elex (green)

Resources

The ⅓-octave decibel numbers I used for plotting these graphs may be found on my website. Several, including the Hammershøi & Møller and the Brüel & Kjær numbers, are estimates based on visual inspection of the published graphs and have no greater authority than that.

Some source material and further reading/viewing:

1. D. Hammershøi and H. Møller, "Determination of Noise Emission from Sound Sources Close to the Ears”, *Acta Acustica*, Vol.94 No.1 (January 2008)
2. 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
3. Tyll Herstens, “Headphone Measurements Explained - Frequency Response Part One”, https://www.stereophile.com/content/innerfidelity-headphone-measurements-explained
4. Warren TenBrook, “Warren TenBrook's Summary of Head Measurements at Harman”, https://web.archive.org/web/20190711222725/https://www.innerfidelity.com/content/warren-tenbrooks-summary-head-measurements-harman
5. Warren TenBrook, “An Acoustic Basis for the Harman Listener Target Curve”, https://web.archive.org/web/20160924084851/http://www.innerfidelity.com/content/acoustic-basis-harman-listener-target-curve
6. David Griesinger, “Binaural hearing, Ear canals, and Headphones” (2016), https://www.youtube.com/watch?v=a-JGAobDwGs
7. Oratory1990, “Differences between the Harman Curve and Diffuse Field”, https://www.reddit.com/r/headphones/comments/78x77b/initial_impressions_of_2016_audeze_lcd2f_with/doyj84e/
8. Oratory1990, list of EQ presets [and measurements], https://www.reddit.com/r/oratory1990/wiki/index/list_of_presets and https://headphonedatabase.com/oratory
9. Sean Olive, Do Listeners Agree on What Makes a Headphone Sound Good?, https://www.youtube.com/watch?v=f1EVZVDaeLw
10. Olive and Welti, AES 139 Presentation 2015 Olive and Welti Preferred Bass and Treble Levels, https://www.youtube.com/watch?time_continue=2&v=ySQV5OR71e4&feature=emb_logo
11. Wikipedia, “Head-related transfer function”, https://en.wikipedia.org/wiki/Head-related_transfer_function