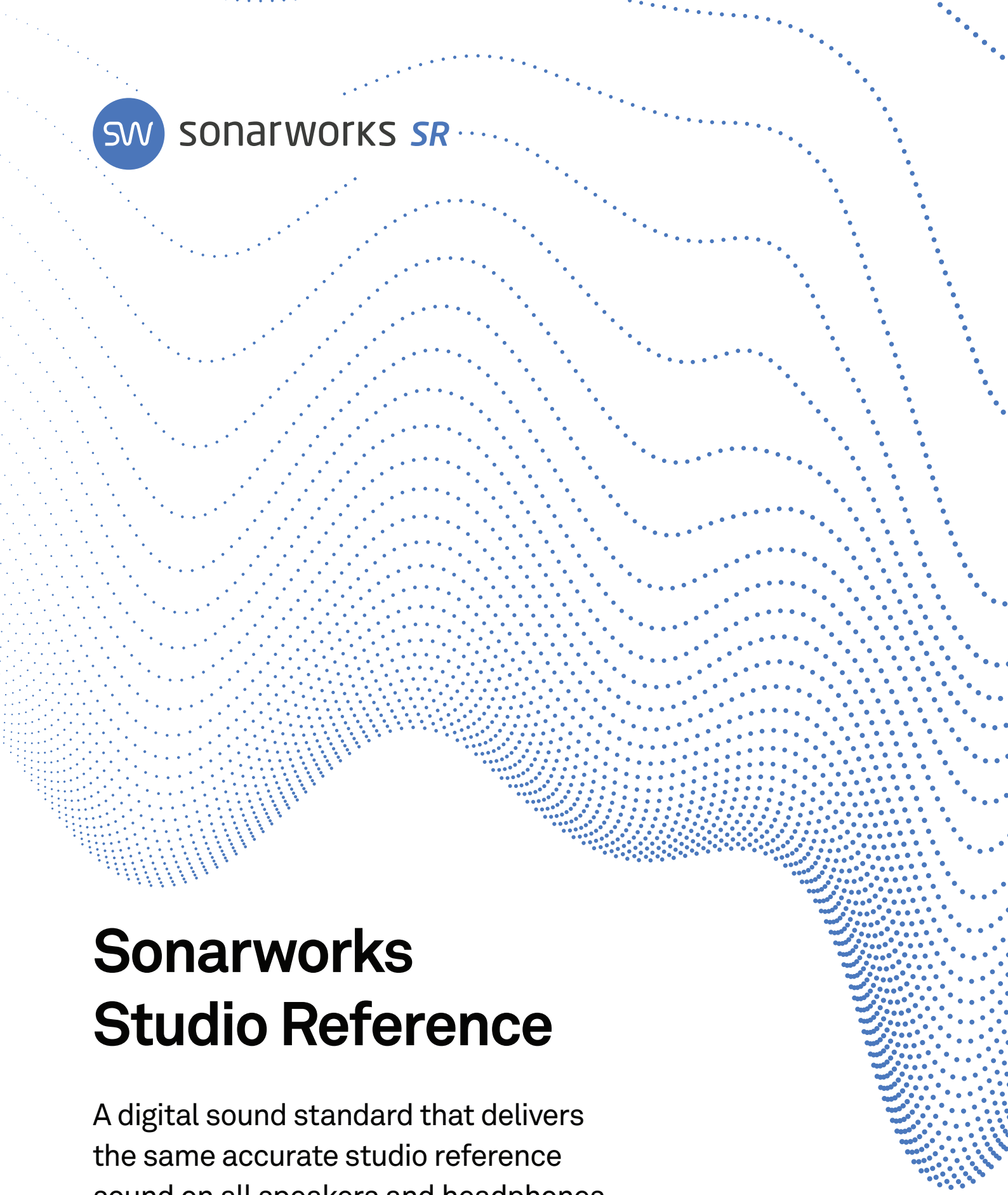

sonarworks SR

Sonarworks Studio Reference

A digital sound standard that delivers the same accurate studio reference sound on all speakers and headphones

Technical Overview / June 2018

Table of contents

1. Introduction

Speakers and headphones have been designed primarily for reproducing recorded audio. However, when a recorded piece of audio is played back through various sets of speakers and headphones, the listening experience varies due to the their individual designs. In the case of speakers, the environment they are placed in plays a major role in the produced sound quality. These factors contribute to speakers and headphones exhibiting varying **frequency response** (FR), which is a measure for describing how well a playback system reproduces the frequency spectrum of recorded audio.

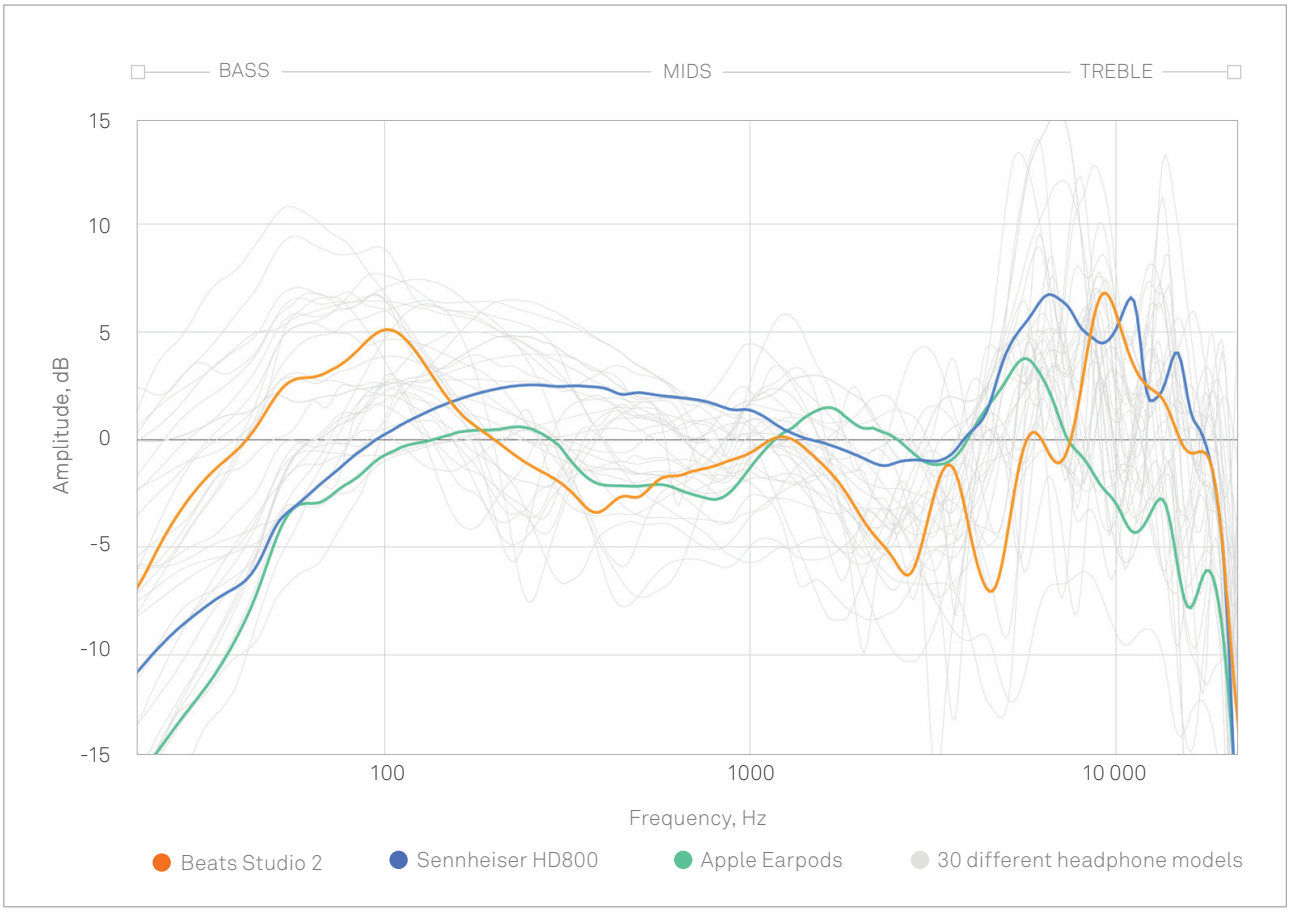

Figure 1: Frequency response variation of headphones
33 different headphone models

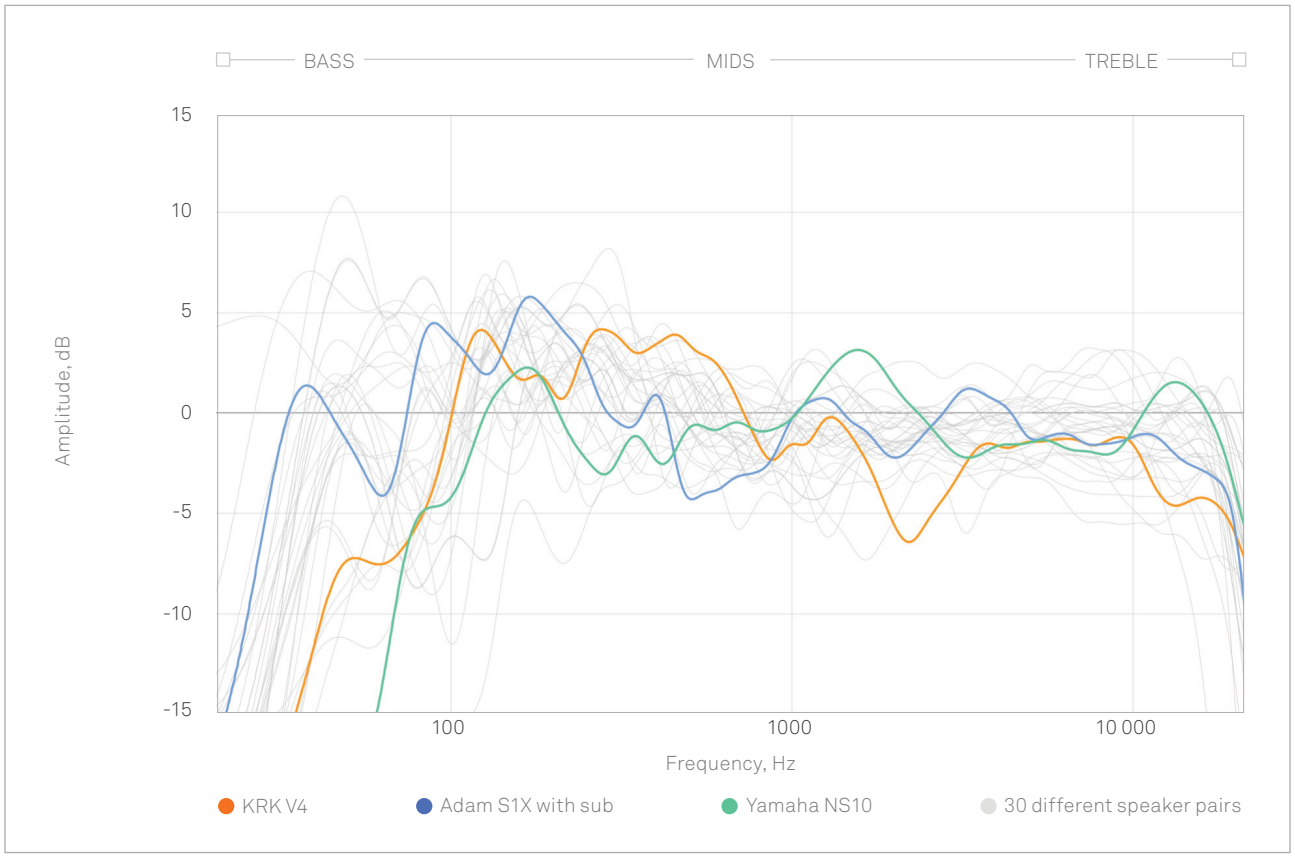

Figure 2: Frequency response variation of studio speakers
33 measured studio speaker pairs in various listening enviroments

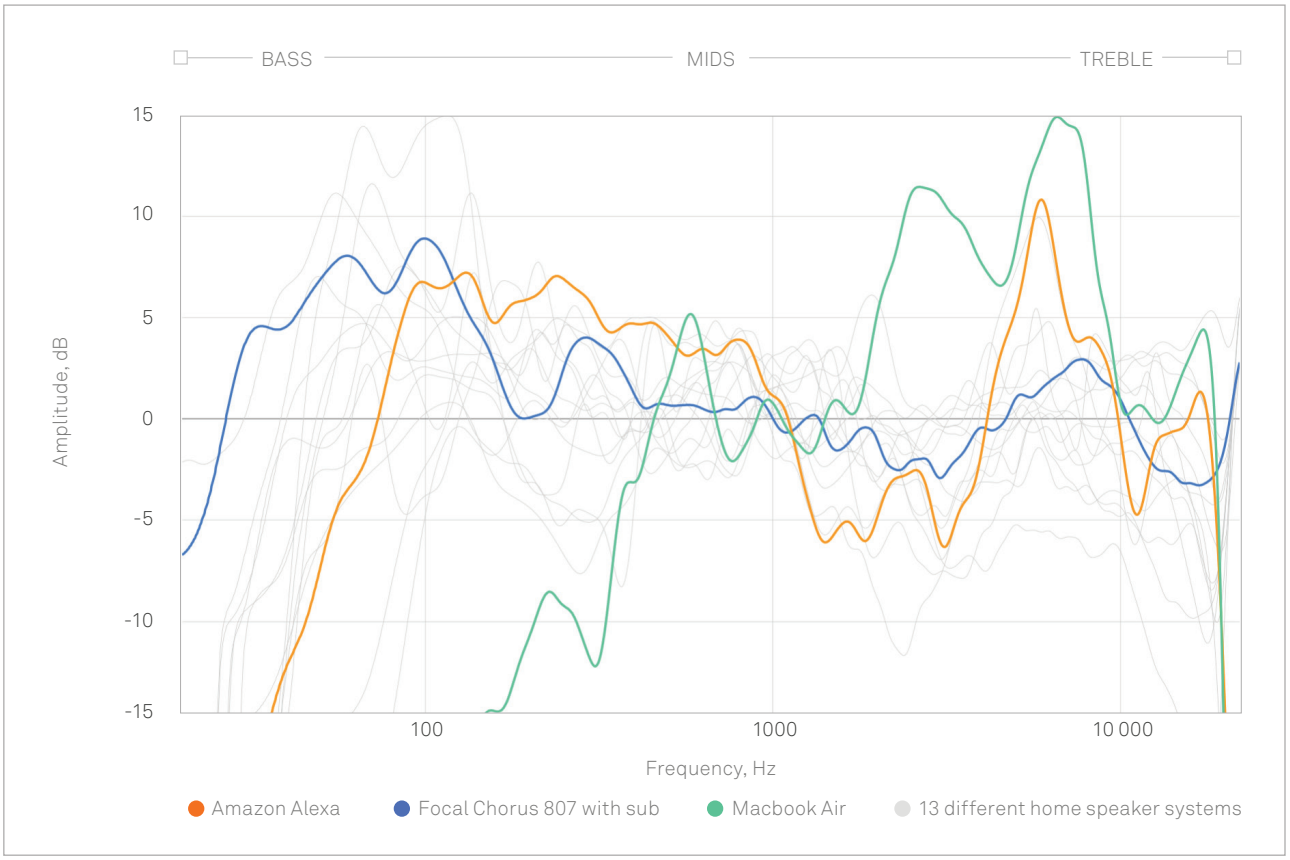

Figure 3: Frequency response variation of home speakers
16 home speaker pairs in various listening enviroments

As variations in playback systems' FR have always been present and unavoidable, it has been accepted as a norm by both audio content creators and audio consumers.

Artists and sound engineers, whose work is most affected by this state of affairs, have incorporated compromise into their workflow. By learning multiple sets of speakers and headphones and cross-referencing their work on them, they accept that a certain amount of compromise in favor of consumer sound systems is necessary to achieve good translation of their work across devices. Audio consumers, on the other hand, are perpetually confused by the question of whether their system sounds good or not.

We believe that this should not be the case. Digital technology has developed to the point where its power can push the quality of the sound of speakers and headphones beyond their analogue limitations, unlocking a venue of new possibilities and breaking the current paradigm of compromise in audio recording and reproduction. However, to have any chance of solving this problem at scale, the technology must be usable by anyone, its implementation for any playback system must be quick and seamless.

This paper proposes a vision for what the new paradigm of sound quality could and should become. Based on digital sound processing technologies developed by Sonarworks, a standard for speaker and headphone calibration is proposed with the aim of eradicating the ever-present issue of poor sound translation across audio playback devices.

2. The problem of translation

2.1. Background

Most of us have listened to the same song on different playback systems, be it different headphone models, laptop speakers, home entertainment system or car speakers. Even for people who do not listen to music on a daily basis the difference in how these devices sound is obvious. But which version of the song you experienced is the one you were meant to hear, the artist's true intent? The answer is quite simple - none. The acoustic properties of a music listener's playback system more often than not have little similarity with the setup the artist used when creating the song. Some of the music always gets lost in translation.

Furthermore, this is not a question on the audio consumer's part on how much to invest in headphones or speaker setup to be able to hear the artist's true vision - the retail price of headphones does not correlate to accuracy of sound reproduction[1] and even the best speaker performance can be significantly reduced depending on their position in a room, the listening position relative to the speakers and the acoustic properties associated with the listening environment (Fig 4).

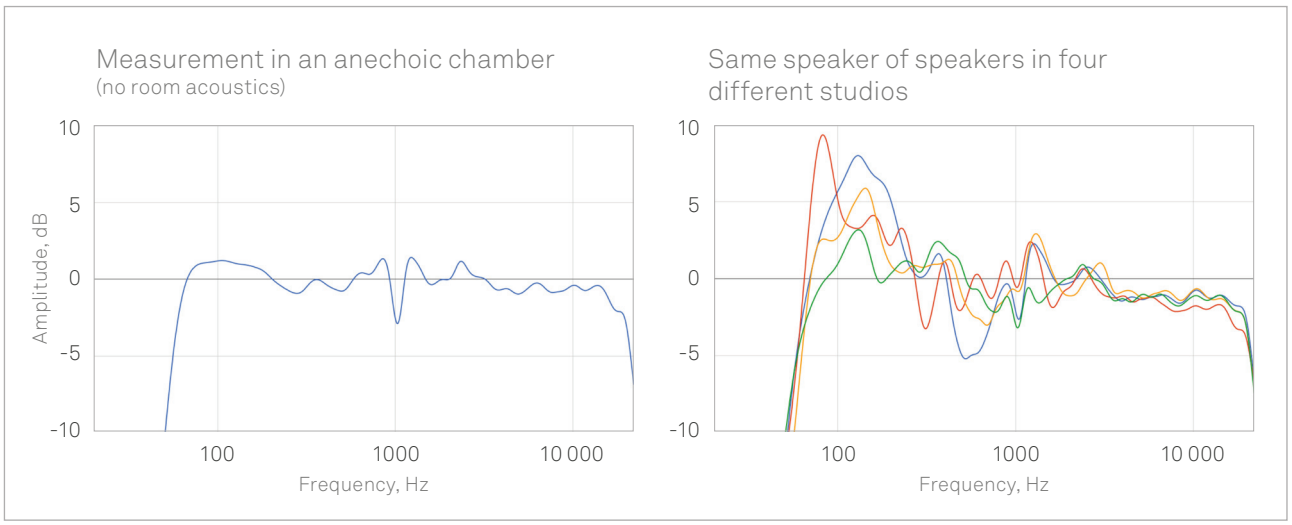

Figure 4: Speaker FR changes significantly when placed in a room

[1] *No correlation between headphone frequency response and retail price,* Jeroen Breebaart, The Journal of the Acoustical Society of America 141, EL526 (2017)

2.2. Sound translation and the music industry

The main purpose of great sound translation is the exact delivery of the artist's intent to the listener. In the current state of the music industry (in terms of devices made for audio playback), successful translation is highly unlikely, as artists create their work on different sounding studio speakers and production rooms and the end user experiences the audio on different sounding consumer speakers and headphones. For successful audio translation the acoustic properties of both the artist's and end-listener's playback systems and listening environments must be accounted for - a supposedly impossible task.

In general, translation problems lead to a lot of compromises being made along the audio production cycle. Since nobody in the cycle, neither the artists, nor engineers, nor listeners, have access to an objective reference point for how things should ideally sound, there is a lot of back and forth exchange of feedback that in large part is caused by each of the arguing parties having differently sounding playback equipment. For one of the parties the bass might be too punchy, for another it might not be there at all!

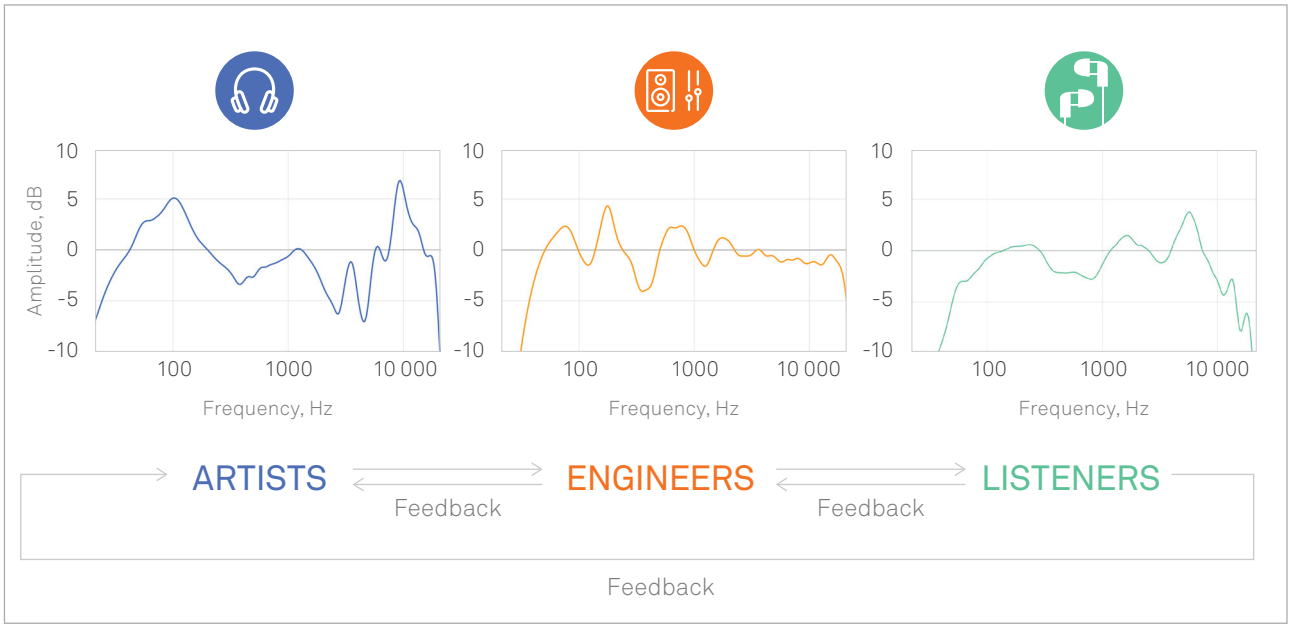

Figure 5: The cycle of compromise
Since nobody has access to a consistent sound standard, audio production is a continuous cycle of compromise

A similar idea has been described as 'The circle of confusion' by Floyd Toole in his seminal book *Sound Reproduction: Loudspeakers and Rooms.* At no point in the cycle of audio production is the hardware or music evaluated using a consistent standard for how things should sound. As a result, audio professionals can never be sure their work will translate on the listeners' systems and successfully convey their artistic vision. In general, professionals are intimately familiar with the sound translation problem, while the end audio consumers by and large remain unaware - it is considered as a natural state of things.

2.3. How do audio professionals cope?

Things being as they are, the general consensus among audio professionals is that working on a neutral-sounding system leads to best results. A neutral or 'flat' sounding system reproduces recorded sound as it is, without adding any unwanted FR coloration. Such systems yield better results not only because they allow hearing sound content accurately in terms of FR, but also because they reveal additional detail that can otherwise be masked by poor acoustics. Because the resulting sound material is not affected by the system it has been produced on, it, on average, sounds better on most other systems.

There are several approaches audio professionals use to strive for neutral sound. Applying acoustic room treatment is a common method for taming acoustics issues. Other more subjective methods include getting feedback from friends and family as well as the laborious process of learning the sound of their speakers by both listening to lots of reference material and listening to their own material on different systems - be it their car, club systems or their kitchen stereo.

Most audio engineers agree that acoustical treatment of the production room and the performance of the sound monitoring system are equally important in achieving a neutral-sounding system (*Mixing Secrets,* Mike Senior, 2011, p.17). Dampening the production room certainly helps, but to a limited extent. Not every studio owner has the significant financial resources required to eliminate all of the room's acoustic issues through room treatment. Even the top range studios with considerable investment in room treatment still have issues with acoustics, for instance, due to changes in studio environment caused by new gear being added to the studio.

Engineers then spend years learning the peculiarities of remaining acoustics issues to be able to mentally compensate for the sound coloration added to the mix. As mentioned above, a common approach is cross-referencing the mix on many different systems in many different environments to find the balance at which the audio track sounds 'equally good' on all systems. Note: good, not great. This method, while widely accepted and adapted by audio engineers, is flawed. Even the most super-human among sound engineers can only get an imperfect approximation of their system's performance and, more often than not, compromising between multiple playback systems leaves the final creation sounding sub-par to the artist's personal high standards.

2.4. The impossibility of translation in the current state of affairs

With engineers accepting the need for compromise and creating records to be optimally enjoyable on the majority of the playback systems, the released version can differ drastically from the complete message the artist wanted to convey. Having the most neutral-sounding production room in the world still does not solve the translation problem due to the unpredictable effects the end-listener's system can have on the sound content.

Artists and audio professionals shape the sound of their work based on imaginary impressions about what the average playback system sounds like. Many get stuck in endless iteration cycles tweaking the sound of their work just to find out it sounds horrible in their car which lead to more tweaks just to find out that their friends' headphones fail to reproduce the sonic impact they intended. There is no single standard to follow to tackle the problem of sound translation, both on professional audio industry and consumer side, and it is our goal at Sonarworks to remedy this problem once and for all. We also hope to raise awareness of the problem of sound translation in the audio industry and educate the end audio consumer on a solution to a problem most consider a natural state of things.

3. The Sonarworks solution

3.1. The Sonarworks SR standard

Sonarworks SR (Studio Reference) is a digital technology solution for delivering consistently accurate studio reference sound across all existing speakers and headphones. It includes in itself a solution for measuring the speakers and headphones as well as a sound processing engine for correcting the audio signal ensuring that the listeners hear accurate sound free from all unwanted coloration no matter what speakers or headphones they use. The Sonarworks SR standard sets the frequency response target to be completely neutral for speaker-based playback systems, i.e. a flat FR curve across all audible frequencies as perceived by the listener in the listening position. The headphone FR target is designed to emulate neutral-sounding speakers.

With the introduction of the Sonarworks SR standard we aim to solve the translation problem at its core. Our solution is not the first acoustics measurement and correction solution out there.

We believe, however, that it is the first one that is capable to successfully tackle the translation problem at scale. All other solutions for measuring and improving the frequency response of sound playback systems have some serious drawbacks that prevent their mass adoption. These solutions have either not been able to deliver consistent results across devices, are developed to cater only for speakers or only headphones, or are simply too complicated or expensive.

We believe that in order to achieve wide adoption, the technology must meet the following specification:

- **Consistency:**
  The sonic results delivered by the system should be consistent between
  • Different speakers and headphones.
  • Different rooms the speakers are placed in.
  • Different people operating the software.

- **Unified standard for speakers and headphones:**
  the same FR standard should be applied for speakers and headphones to ensure translation between these two modes of listening.

- **Ease of use:**
  Its use must be convenient and intuitive to the end-user. It should require no prior training or specialist knowledge.

- **Transparency:**
  The sound processing engine should not add any audible artefacts.

Sonarworks SR is a unique technology that meets the above mentioned criteria, introducing a standard in sound recording and reproduction across all sound playback platforms and listening environments. The standard enforces consistent sound delivery between sound reproduction devices by removing human error from the measurement process and remaining easy to implement, maintain and operate without any prior training.

The implementation of Sonarworks SR standard is designed to perform consistently between playback devices and work behind the scenes with ease in order to allow the user to focus on what matters - creating and experiencing recorded sound without any loss to its sonic character.

3.2. Speaker and room frequency response capture

In order to attain neutral frequency response for speaker systems, their FR first needs to be measured. There are three main problems for measuring speaker's FR in the listening environment:

- **Problem 1:**

  The frequency response of a speaker in a room is highly uneven.Two different measurements taken 10cm (~4 inches) apart produce considerably different FR results. (See Fig 6)

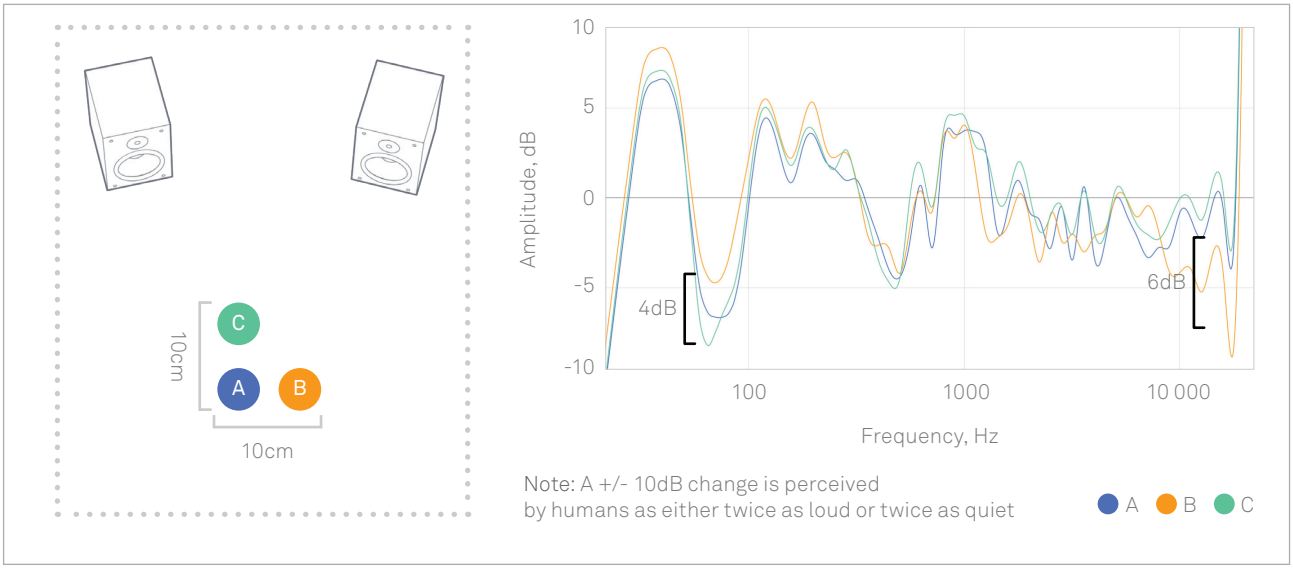

Figure 6: Speaker FR is uneven in a room
The same speakers measured in three different studio positions (10 cm apart)

- **Problem 2:**

  Human beings do not perceive sound the same way a measurement microphone does. Whereas for a measurement mic the room's FR changes significantly as it moves around (see above), it normally sounds much more even for a human listener due to brain interpretations of the sounds that we hear.

- **Problem 3:**

  Not all of the rooms acoustic phenomena can be solved through a digital calibration[2]. Things like standing waves can not be 'equalized out' of the room and must be dealt with by changing room acoustics.

The main challenge is developing a tool for consistently measuring FR of the speakers in listening environment in a manner that effectively deals with the problems mentioned above.

The unique solution that Sonarworks SR brings to solve this challenge is by taking speaker measurements across multiple points, while simultaneously recording coordinates of those measurements. As a result the software sees the individual measurements in context and builds itself a map of the room acoustics. Seeing the room's acoustic map enables our software to build a smart average of the individual measurements in a manner that is similar to how the human brain perceives the room acoustics as well as excludes attempts to correct for acoustic phenomena that cannot be corrected on a software level. On top of that the location of microphone with each measurement allows for precise user guidance during the process ensuring consistent results between independent measurements.

[2] *The Measurement and Calibration of Sound Reproducing Systems,* Floyd E. Toole, Journal of the Audio Engineering Society, Vol. 63, No. 7/8, p517, July/August 2015

The Sonarworks SR measurement solution currently captures a map of 37 measurement positions which yields a maximum error of +/- 0.9 dB between different measurements (Fig. 7).

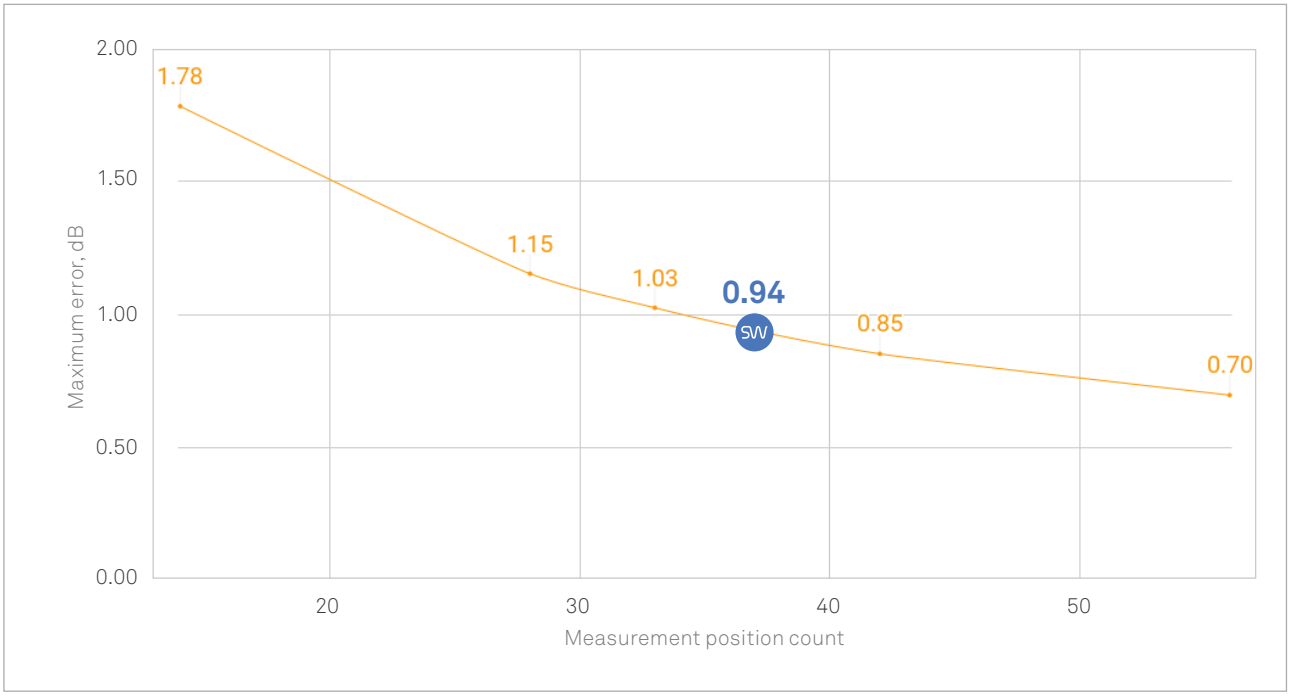

Figure 7: Sonarworks SR speaker measurement accuracy
A measurement of 37 points around the listening spot yields a maximum error of 0.94 dB

As a result, the Sonarworks SR solution delivers consistent repeated acoustic measurements of unique audio playback environments and equal sonic experience across all systems calibrated according to the standard.

3.3. Headphone frequency response capture

Headphones exhibit the same translation problems that speakers do, with the exception of room influence to some extent. Headphone frequency response capture is more complex than measuring speakers due to the close proximity of the headphone driver to the ear canal. Consequently, the position of the measurement microphone relative to the headphone driver and the seal the headphone under test has with the measurement rig heavily affects the measurement results.

Although there are a couple of devices out on the market designed for headphone measurement, they do not allow for a consistent measurement of headphones FR in a way that fulfills the goals described in this paper. The problem with existing headphone measurement tools (head and torso simulators, etc.) is that historically their main design goal has been to allow hardware manufacturers to comply with loudness and safety standards. While they can be used for frequency response measurement these tools can't achieve consistent sound across different headphone types and models.

Sonarworks has developed proprietary hardware and software analysis tools capable of consistently mapping headphone features. The target of headphone calibration is an internally developed speaker-to-headphone transfer curve, which has been rigorously tested with audio professionals, i.e. the people actually capable of evaluating the sound of playback systems.

The Sonarworks SR headphone measurement system:

- Yields accuracy of +/- 0.9 dB for each individually calibrated headphone pair.
- Yields +/- 3 dB accuracy for averaged profiles of a headphone model. The lower accuracy of averaged headphone measurements is caused by differences between different headphone pairs of the same model. See Figure 8 for maximum error of averaged headphone profiles by frequency band.
- Compensates for Left / Right channel loudness differences.
- Accounts for real-world headphone usage by compensating for various ways over-ear headphones can be positioned relative to the ear. Also, bass leakage effects due to improper seal around the ear are accounted for.
- Is compatible with all headphone types, both in-ear and over-ear.

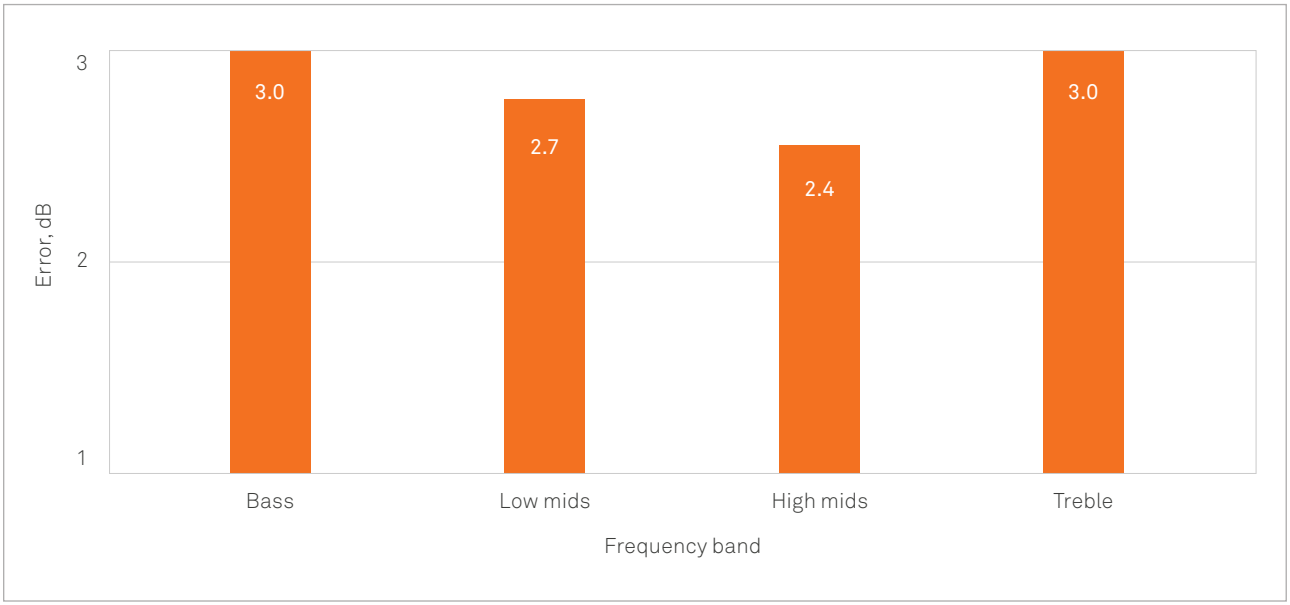

Figure 8: **Maximum error of averaged headphone profiles, dB**
Listening test results for averaged measurement profiles of 80 headphone models tested by trained listeners for how closely they match the Sonarworks SR sound standard

For consistent results, all headphones are individually tested by in-house engineers specifically trained for critical listening. All headphone profiles published by Sonarworks have gone through this rigorous testing protocol and are in compliance with SR standard.

3.4. Audio signal processing

Knowing your playback system is underperforming and what the particular problems are in objective detail is only a part of the battle. The objective is achieved only when the problems are eradicated, and in the field of audio playback consistency, this is done by the Sonarworks SR filter DSP engine. The main functionality of the filtering engine is conditioning the playback system to a target sound profile and doing so accurately. Any processing carried out by the filtering engine must be transparent to the listener and true to its purpose - processing an audio signal exactly to the specification without any addition of unwanted artefacts.

Processing audio signals in real time to correct for playback system's FR coloration effects is a resource-intensive process. Depending on the target application, e.g. mixing a track, real-time session monitoring or simply listening to music, a balance must be found between the required processing resolution, CPU load, effects of added phase distortion and the amount of maximum allowable latency to be added in the signal playback chain. For live monitoring, CPU load and latency minimization takes precedence whereas for mixing session latency is of lesser concern and processing resolution and phase distortion minimization takes over. The filter engine, as a consequence, must be adaptive to the purpose the audio signal is processed for.

Sonarworks solution implements cutting edge Finite Impulse Response (FIR) filter technology. The filter engine is of high precision, highly efficient and transparent to the listener when engaged. Supporting sampling frequencies up to 192kHz, Sonarworks filter technology adapts to the relevant use case, be it live monitoring or simple audio playback. Regardless of the operation mode, audio is processed at high resolution to capture the FR specifics of the playback system without any unwanted audio artefacts added to the signal due to signal processing. Currently implemented on Windows, MacOS, iOS and Android platforms, Sonarworks SR standard is designed to be easily integratable into virtually any audio signal processing and playback chain.

4. Conclusion

The advances in audio digital processing technologies have allowed the problem of sound translation to potentially become a thing of the past. Sonarworks intends to make this happen and introduces the Sonarworks SR standard to form a common ground in music producer and consumer sonic experience.

Already tested by tens of thousands of audio engineers and endorsed by Grammy-winning engineers, Sonarworks SR standard offers a scalable solution for audio translation in a single multi-functional package.

Here's to the future where everyone can hear music the way it was meant to be!

References

1. **“Factors that Influence Listeners’ Preferred Bass and Treble Balance in Head phones”,**

   Sean E. Olive, Todd Welti, Convention Paper 9382, AES, NY, USA, 2015

2. **“No correlation between headphone frequency response and retail price”,**

   Jeroen Breebaart, The Journal of the Acoustical Society of America 141, EL526, 2017

3. **“Sound Reproduction: Loudspeakers and Rooms”,**

   Floyd E. Toole, Focal Press, MA, USA, 2013

4. **“Acoustics and Psychoacoustics”,**

   4th Edition, David M. Howard, Jamie A.S. Angus, Focal Press, MA, USA, 2009