The Basics of High Fidelity
Part 3: The Ideal Loudspeaker, Diffuse Field Equalization

Part 1 dealt with the problem of transparency, Part 2 with the problem of applying this idea on the transduction of audio into sound. In this part we will elaborate further on the idea of transparency in order to be able to design the ideal loudspeaker system.

In Part 2 I have explained the two basic definitions of transparency when listening to a loudspeaker (or headphone), one related to the ideal “Here and Now” (Augmented Reality) and one related to the ideal “There and Then” (Virtual Reality). The “Here and Now” transparency creates the illusion that the sound source is present in the room where you are sitting while “There and Then” creates the illusion that you are present at the place where the recording was made. These two different transparency ideals can never be fulfilled at the same time and require complete different designs of the recording and reproduction chain.

In general we will choose for the illusion “There and Then” because of the wonderful acoustic designs of concert halls. When pursuing this ideal with a loudspeaker we will have to deal with the acoustic environment of our reproduction room, so one would be tempted to use headphone reproduction. However there are severe limitations in using headphones:

- Acoustic recordings are seldom made with an artificial head, and even if they are made with an artificial head, the adaptation of such a recording towards the individual HRTF’s of a listener is extremely difficult (Head Related Transfer Functions, one for the left and one for the right ear).
- Low frequencies are for a major part perceived with our body so we miss the low frequency impact.
- Head movements are difficult to take into account.
- The ideal wearing comfort does not exist, we feel excluded from our natural environment.

So we are stuck with our loudspeakers, but new smart signal processing techniques are currently promising to solve the reproduction transparency problem using wave front synthesis. A very expensive solution that seldom can deal correctly with the reproduction room reflections. And what about the more simple surround (5.1, 7.1 or X.Y)? Well, they require special recording techniques and in general the whole surround approach is focused on movies where back localization is important, contrary to music reproduction, where back localization is irrelevant. In rare occasions one may need back localization, but in most cases back localization is a reproduction error that can lead to annoying artifacts. The only reason for choosing surround sound is the improved reproduction of the diffuse field of the recorded sound. However when using a surround set the reproduction room will still add coloration to the sound. A much simpler and cheaper approach that can be used with any normal stereo set, without the need to make surround recordings, is diffuse field equalization. In this approach the room reflections are exploited and used to solve a fundamental loudspeaker reproduction problem, allowing to create a high quality “There and Then” illusion if the recording that is used contains the correct acoustic information (a high quality is example is this BIS recording).

How does it work? Let's start by explaining a basic loudspeaker problem, the bundling of sound. As you probably know the bundling of a sound depends on the ratio of the wavelength that is reproduced and the physical dimensions of the object that radiates the sound. First formulated by Huygens in 1678, the bundling can be visualized by drawing secondary circular waves at each point of a wave front. When the object that is radiating the sound is small compared to the wavelength of the sound, the resulting radiation is circular, e.g. when reproducing low frequencies with a normal loudspeaker. When the object that is radiating the sound is large compared to the wavelength the radiation pattern is progressively more bundled. A simple solution to the bundling problem is to use a small loudspeaker (tweeter) for the high frequency range. However in the cross-over frequency range, where the tweeter takes over from the woofer, we get a dip in the diffuse field response. Just below the cross-over frequency we have strong bundling and just above the cross-over frequency we have almost no bundling.

We could equalize the frequency response at the listening position, however, this will introduce an even bigger problem than we are trying to solve, the direct, on axis frequency response will now show a peak at the cross-over frequency. Because our perception is dominated by the first wave front (the so called [Haas effect](#)) the resulting sound reproduction of the system will be unnatural (see Figure 1).

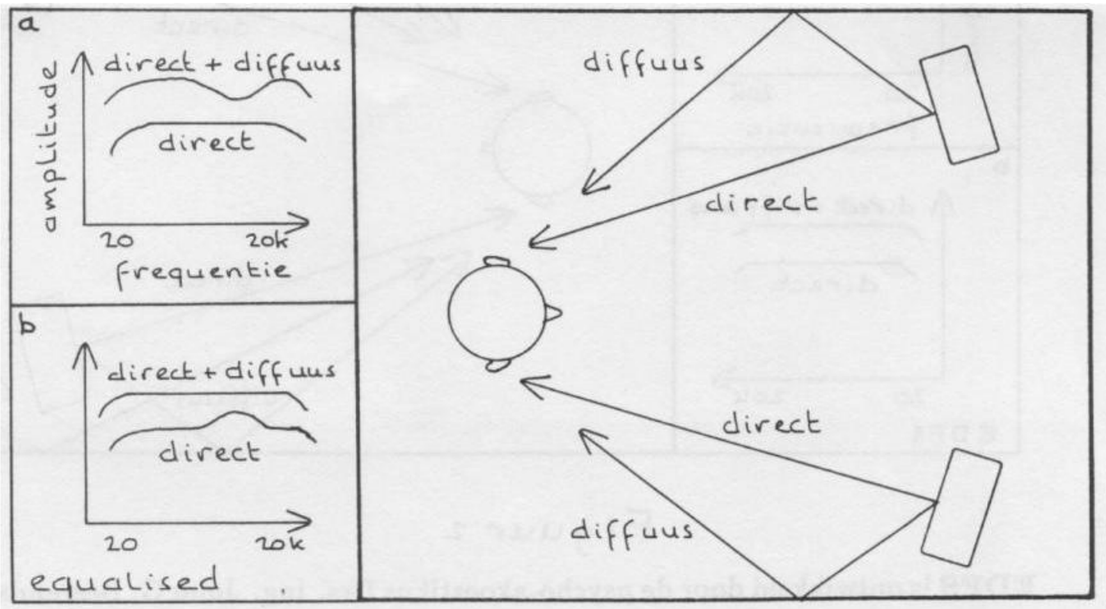

Figure 1. Frequency characteristic of a classical two-way loudspeaker reproduction system. If we do not use any equalization we get a dip in the frequency response at the listening position (direct+diffuse field, Figure 1a). If the response is equalized to be flat the resulting direct, on axis, response will show a peak at the cross-over frequency (Figure 1b).

The correct way to equalize the sound field is to use a second set of loudspeakers placed behind the main loudspeakers and apply the equalization only to this second pair of loudspeakers (see Figure 2). By manipulating the power division between the back and front radiating loudspeakers one can balance the diffuse field according to ones preferences, while at the same time keeping the direct+diffuse field response flat, creating an absolute fantastic sound reproduction system not available from any commercial vendor. The most striking error in expensive surround systems is that they provide measurement microphones and measurement signals to equalize the reproduction chain but without separating the direct and diffuse field transfer function. All consumer surround systems that I know of are only useful in providing side and back localization as required with movie sound. For music reproduction however, side and back localization are almost never required and to the despair of many HiFi freaks standard consumer surround systems often provide disturbing artifacts when used for music reproduction. Only in rare occasions, e.g. when reproducing some obscure Stockhausen piece, side and back localization are an advantage. The general preference with surround sound for music reproduction is to use the side and back loudspeakers only for diffuse field reproduction in combination with some kind of diffuse field equalization. The first ideas of diffuse field equalization were already formulated in the eighties and implemented in a loudspeaker design [1], but I have not seen any modern surround system on the market using this approach.

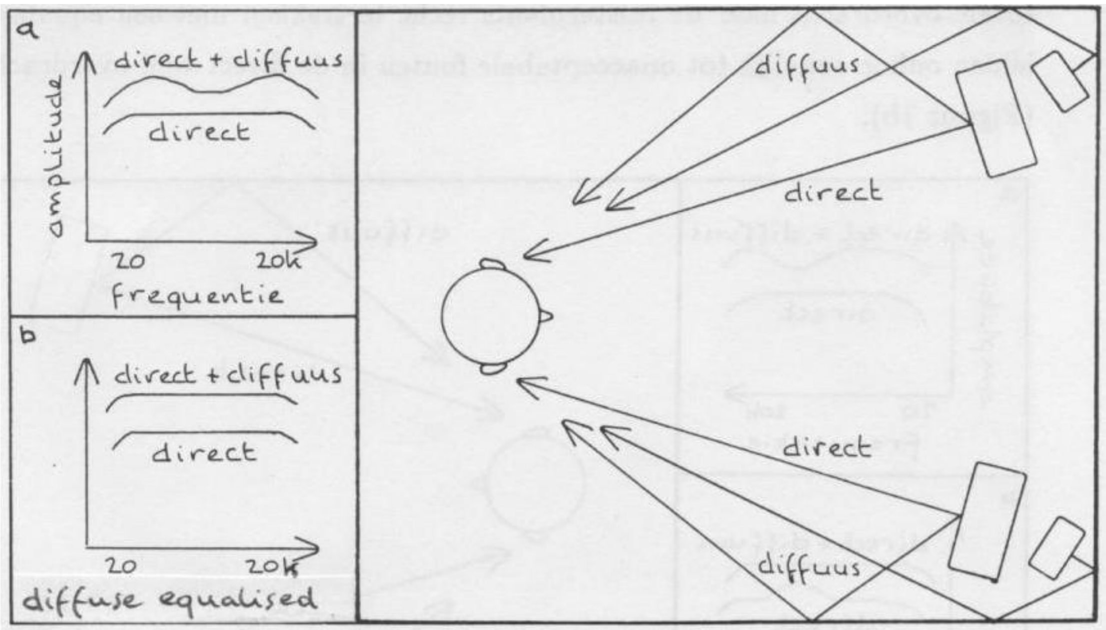

Figure 2. Frequency characteristic of a classical two-way loudspeaker reproduction system. If we do not use any equalization we get a dip in the frequency response at the listening position (direct+diffuse field, Figure 2a). If the response is equalized using only the second set of back radiating loudspeakers the resulting response is flat for both the direct field as well as the direct+diffuse field (Figure 2b).

A problem in assessing the reproduction quality of a loudspeaker system is that the transparency approach as explained in Part 1 is difficult to implement in the acoustic domain due to the fact that subjects cannot be provided with an acoustic reference signal. Loudspeaker reproduction quality assessment by subjects is always based on an unknown, internal, ideal reference that is formed by their listening experience. If we want to develop an objective perceptual assessment algorithm equivalent to the one described in Part 1 we will need to construct an ideal reference signal. This can be carried out by making binaural recordings with a head and torso simulator of a set of music signals, using the best quality loudspeakers, in the ideal listening position in the best quality listening environment, see [2]. For each music signal the ideal reference signal is defined as the recording with the highest subjective quality and these signals are then compared to the recordings of the acoustic output of the loudspeaker that is under test. This method thus allows prediction of the subjectively perceived sound quality of loudspeakers, taking into account the influence of the listening room, the listening position and the type of music signals that are considered to be relevant.

[1] K. L. Kantor and A. P. Koster, “A psychoacoustically optimized loudspeaker,” J. Audio Eng. Soc., vol. 34, pp. 990-996, (1986 Dec.).

[2] J. G. Beerends, K. van Nieuwenhuizen, and E. vd Broek, “Quantifying Sound Quality in Loudspeaker Reproduction”, J. Audio Eng. Soc., vol. 64, pp. 784-799 (2016 Oct.).

Go to Part 4: The Ideal Loudspeaker, Reflections and Resonances.

John G. Beerends
Published in Hifi Video Test 12/2007 (in Dutch), translated and updated over the period 2012-2020.