Author Accepted Manuscript (corrected). This paper was presented at the 131st Convention of the Audio Engineering Society 2011, as paper number 8568. The full published version can be found at `http://www.aes.org/e-lib/browse.cfm?elib=16093`.

System Theory of Binaural Synthesis

Florian Völk

AG Technische Akustik, MMK, Technische Universität München, Arcisstraße 21, 80333 München, Germany

Correspondence should be addressed to Florian Völk (`florian.voelk@mytum.de`)

ABSTRACT

Binaural synthesis is widely used as efficient tool for the simulation of acoustical environments. Different headphones together with artificial as well as human heads are employed for the transfer function measurements involved, having considerable influence on the synthesis quality. Within this paper, a detailed system theoretical analysis of the signal paths and systems involved in a typical data based binaural synthesis scenario is given. The components to be equalized are identified, and equalization methods for every scenario are discussed. Further, restrictions and necessities when using artificial or human recording heads and for headphone selection are given. Most important results are the necessity of blocked auditory canal measurements and the selection of proper headphones for completely correct individual binaural synthesis.

1. INTRODUCTION

In a fundamental paper on binaural technology, Møller defined in 1992 the basic idea of binaural recording as follows: *"The input to the hearing consists of two signals: sound pressures at the eardrums. If these are recorded in the ears of a listener and reproduced exactly as they were, then the complete auditive experience is assumed to be reproduced, including timbre and spatial aspects."*

This so-called binaural recording of the sound pressure signals at the eardrums (the ear signals) is a time consuming process, especially if more than one listener, different head rotations, or different listening positions are to be considered. However, due to its simplicity, this procedure has been in use for many years (early work e. g. Bixler 1953, Wilkens 1972), for a review cf. Hammershøi and Møller (2002). A less time consuming method to acquire the necessary sound pressure signals is binaural synthesis (BS) (cf. Møller 1992). Within this procedure, the sound pressure signals are generated by convolution of the input signal of a single sound source with the impulse responses describing the paths between the sound source and the listener's eardrums in the corresponding real life situation. These impulse responses are dependent on the loudspeaker and body position and orientation as well as on the current geometrical arrangement of the listening situation considered. To account for possible topological variation, an adaptive signal processing is necessary that selects the impulse responses corresponding to the current listening situation at any instant of time based on the body and loudspeaker positions, especially if the listener moves, if the loudspeaker is repositioned, or if other parts of the situation change (cf. Mackensen et al. 1999). Complex acoustical situations with more than one source can be generated by linear superposition of the sound pressure signals evolving from each single source.

By now, a wide variety of BS systems is available and such systems are in use in virtual or augmented reality applications (cf. Begault 1999, Gröhn et al. 2007, Völk et al. 2007) or in psychoacoustic research (cf. Blauert et al. 2000, Djelani et al. 2000, Zahorik 2002). If certain prerequisites are met and necessary corrections are applied, presentation of the convolution products generated by a BS system using headphones can produce the same or at least approximately the same ear signals as would have been present in the corresponding real life situation. In the following, the necessary restrictions and corrections are identified and discussed step by step, based on a detailed system theoretical analysis of all involved components. With these requirements, different possible equalization procedures are presented with respect to the achievable overall system transfer functions. In addition, hardware selection criteria are introduced.

1.1. Considered Situation

The ultimate goal of a BS system is the perfect reconstruction of the auditory impression created by a real sound scene, for example a single loudspeaker box in a living room. This situation is defined as the reference scene for the following considerations. Then, the most important signals involved are the time varying sound pressure signals at the eardrums, the ear signals $p_{\mathrm{e_L}}(t)$ and $p_{\mathrm{e_R}}(t)$, the loudspeaker input voltage $u_{\mathrm{ls}}(t)$, as well as the digital sample sequence $s_{\mathrm{ls}}[n]$, encoding the sound signal to be played back by the loudspeaker. As basis for a linear and time invariant system theoretical analysis of the sound paths from a sound source's input to a listener's eardrums, all involved signals and systems have to be linear and time invariant (cf. Oppenheim et al. 1998). All systems involved in a typical BS situation can be modeled as linear and at least piecewise time invariant systems (within each period of time τ_i with $i \in [0, N]$). Therefore, it is possible to describe each of the systems completely by its impulse response $h(t)$ or by an array of impulse responses $h_{\tau_i}(t)$ with $i \in [0, N]$. The restriction to a loudspeaker box does not result in loss of generality since the voltage $u_{\mathrm{ls}}(t)$ can be regarded in any case as input to any ideal electroacoustic transducer (e. g. an acoustical monopole or dipole point source, cf. Zollner and Zwicker 1993). An additional advantage resulting from the linearity of all subsystems involved is the possibility to describe complex acoustical scenes including more than one listener and different sound sources as (linear) superposition of a basic scene consisting of one listener and one sound source.

1.2. Terminology

In the following, lower case letters denote time dependent signals or impulse responses. The voltage $u\,(t)$ is abbreviated for example as u without explicit notation of the time dependence. Upper case letters represent complex Fourier-spectra or transfer functions, for example $U\,(f) = \mathcal{F}\{u\,(t)\}$ denoted as U. Lower case subscripts differentiate between different signals and systems, additional upper case indexes denote the left (L) or right (R) side, if necessary. A major part of the following considerations is done for the left and right side in a similar manner, therefore variables representing the same components for both sides are summed up as vectors and therefore set in bold fonts, for example

$$\mathbf{U} = \begin{pmatrix} U_L \\ U_R \end{pmatrix}.$$

Throughout this article, the division or multiplication of two vectors is used as shorthand for element wise division or multiplication, i. e. the element representing the left side in the first vector is divided by or multiplied with the corresponding element of the second vector and vice versa. To underline that a signal or an impulse response is characteristic for one individual listener, the upper index *ind* is introduced. If the usage of an artificial head shall be indicated, the upper index *ah* is used. For the paper on hand, it is necessary at some points to denote explicitly if a signal is represented digital or analog. Digital sequences are denoted by the variable s in the time domain, the corresponding discrete Fourier transform (DFT) represented by S. If not denoted explicitly, it is not necessary here to distinguish if a signal or system is represented or works digital or analog. Within this paper, transfer functions are defined between different physical magnitudes. The most common relation in acoustics is that between two sound pressure spectra. To connect acoustic and electronic signals, transfer functions can for example be defined between a sound pressure and a voltage. The typical electrical transfer function is defined between two voltage spectra. To make a connection between physical signals and their digital representations, transfer functions are defined symbolically between them, too. It is a prerequisite for the division of two spectra or transfer functions that the dividend does not equal zero at any frequency. This is taken for granted for all following calculations as well as the invertibility of the mentioned impulse responses. Possible problems concerning these issues in practical implementations are not within the scope of this paper, although they could become a crucial factor in practical implementations (cf. e. g. Neely and Allen 1979).

2. REFERENCE SCENE

The reference scene for the following considerations is one person listening to a loudspeaker box. The position and orientation of the subject's head in an arbitrary coordinate system is given by

$$\mathbf{x}_\mathrm{h} = (x, y, z, r_x, r_y, r_z, \delta)^T.$$

Here x, y, and z denote the position, r_x, r_y, and r_z the orientation of the head (the rotation around the different axes). The additional parameter δ takes into account symbolically a possible rotation of the listener's head and/or torso with respect to the remaining body. In a straightforward manner, the position of the loudspeaker $\mathbf{x}_\mathrm{ls}$ is defined. Each ear signal occurring in this reference scene includes influences of the transfer characteristics of systems which could be grouped best by their affiliation to the generation or propagation parts of the sound transfer paths from the source to the eardrums.

2.1. Loudspeaker Playback System

The generation part of both paths includes the audio interface used for D/A-conversion of the source sequence s_ls to the output voltage u_da. The output of the audio interface drives the loudspeaker amplifier that provides the output voltage u_ls. These devices are described by their transfer functions

$$H_\mathrm{da} = \frac{U_\mathrm{da}}{S_\mathrm{ls}} \quad \text{and} \quad H_\mathrm{als} = \frac{U_\mathrm{ls}}{U_\mathrm{da}}. \tag{1}$$

The influence of the cables connecting the devices is neglected here for simplicity. Under certain circumstances it might be necessary to consider their transfer characteristics explicitly, which would be fully covered by linear system theory. The transfer function of the output system (the generation part of the transfer paths) can be summed up as follows:

$$H_\mathrm{o} = \frac{U_\mathrm{ls}}{S_\mathrm{ls}} = H_\mathrm{da} \cdot H_\mathrm{als} \tag{2}$$

2.2. Propagation Path

The propagation parts of the sound paths are here defined between the voltage u_ls at the loudspeaker input terminals and the sound pressure signals $\mathbf{p}_\mathrm{e}^\mathrm{ind}(\mathbf{x}_\mathrm{h}, \mathbf{x}_\mathrm{ls})$ at the eardrums. These paths can be described by following transfer functions:

$$\mathbf{H}^\mathrm{ind}_{u_\mathrm{ls},\mathbf{p}_\mathrm{e}}(\mathbf{x}_\mathrm{h}, \mathbf{x}_\mathrm{ls}) = \frac{\mathbf{P}_\mathrm{e}^\mathrm{ind}(\mathbf{x}_\mathrm{h}, \mathbf{x}_\mathrm{ls})}{U_\mathrm{ls}} \tag{3}$$

These transfer functions contain the transfer characteristics of the loudspeaker used as well as of the sound propagation paths between this loudspeaker and the eardrums, including all possible reflection and diffraction effects (e.g. room reflections or transformation characteristics of the outer ears and other body parts). For this reason, the latter transfer functions are highly dependent on the position and orientation of the listener's head and body and are therefore represented as array of transfer functions with the subject and loudspeaker position vectors as array parameters. If the reference scene is located in an anechoic chamber, the transfer functions $\mathbf{H}^\mathrm{ind}_{u_\mathrm{ls},\mathbf{p}_\mathrm{e}}(\mathbf{x}_\mathrm{h}, \mathbf{x}_\mathrm{ls})$ are often referred to as Head Related Transfer Functions (HRTFs) with corresponding Head Related Impulse Responses (HRIRs). However, the correct definition of head-related transfer functions (HRTFs) is the relation of the sound pressure spectra $\mathbf{P}_\mathrm{e}^\mathrm{ind}(\mathbf{x}_\mathrm{h}, \mathbf{x}_\mathrm{ls})$, recorded under anechoic conditions, to the spectrum of the sound pressure recorded at the midpoint of the recording head in the same situation, with the head being absent. If the reference scene is defined in a reflective environment, the impulse responses are commonly called Binaural Room Impulse Responses (BRIRs) or Transfer Functions (BRTFs).

For a more detailed discussion of the usual nomenclature see Blauert (1997), Møller (1992). For the paper on hand, none of these labels is used. Instead, every transfer function is defined in detail and referred to if necessary.

2.3. Ear Signals

The following equation can be given combining 2 and 3 to describe the connection between the sequence driving the loudspeaker and the ear signals in the reference situation (index *ref*):

$$\begin{aligned}\mathbf{H}_{\mathrm{ref}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}) =& \frac{\mathbf{P}_{\mathrm{e}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})}{S_{\mathrm{ls}}} \\ =& H_{\mathrm{o_{ref}}} \cdot \mathbf{H}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{e}}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})\end{aligned} \tag{4}$$

These transfer functions are valid as long as the overall situation does not change. This is important in BS if headphones are used for playback, which in general causes a change in radiation impedance compared to the reference situation (Vorländer 2000).

2.4. Input System

For measurement of the ear signals $\mathbf{p}_{\mathrm{e}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}})$, an input system described by its transfer function $\mathbf{H}_{\mathrm{i}}$ is necessary to transform the output voltages of the microphones to digital sample sequences. Its characteristics consist of the transfer functions of the microphone pre-amplifiers (that amplify the microphone output voltages $\mathbf{u}_{\mathrm{mic}}$ to the voltages $\mathbf{u}_{\mathrm{ad}}$) and the A/D-converters (producing output sequences $\mathbf{s}_{\mathrm{mic}}$):

$$\mathbf{H}_{\mathrm{am}} = \frac{\mathbf{U}_{\mathrm{ad}}}{\mathbf{U}_{\mathrm{mic}}}, \quad \mathbf{H}_{\mathrm{ad}} = \frac{\mathbf{S}_{\mathrm{mic}}}{\mathbf{U}_{\mathrm{ad}}} \tag{5}$$

Combining these equations, the following description of the input system can be given:

$$\mathbf{H}_{\mathrm{i}} = \frac{\mathbf{S}_{\mathrm{mic}}}{\mathbf{U}_{\mathrm{mic}}} = \mathbf{H}_{\mathrm{am}} \cdot \mathbf{H}_{\mathrm{ad}} \tag{6}$$

2.5. Probe Microphones

Ear signals are rather difficult to measure, since it would be necessary to capture the pressure distribution across the whole eardrum. An approximative measurement is possible with probe microphones (*pm*) right in front of the eardrums. The results are then valid only for the positions $\mathbf{x}_{\mathrm{pm}}$ of the probe microphones. The sound pressure spectra $\mathbf{P}_{\mathrm{pm}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{pm}})$ at the microphone positions are related to the probe microphone output spectra $\mathbf{U}_{\mathrm{pm}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{pm}})$ by the transfer functions

$$\mathbf{H}_{\mathrm{pm}} = \frac{\mathbf{U}_{\mathrm{pm}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{pm}})}{\mathbf{P}_{\mathrm{pm}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{pm}})}. \tag{7}$$

The propagation paths between the voltages u_{ls} and the pressures $\mathbf{p}_{\mathrm{pm}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{pm}})$ are given as follows:

$$\mathbf{H}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{pm}}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{pm}}) = \frac{\mathbf{P}_{\mathrm{pm}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{pm}})}{U_{\mathrm{ls}}} \tag{8}$$

In a similar manner, the propagation between the voltages u_{ls} and $\mathbf{u}_{\mathrm{pm}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{pm}})$ is represented by

$$\mathbf{H}_{u_{\mathrm{ls}},\mathbf{u}_{\mathrm{pm}}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{pm}}) = \frac{\mathbf{U}_{\mathrm{pm}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{pm}})}{U_{\mathrm{ls}}}. \tag{9}$$

The relation of the sequence driving the loudspeaker to the sound pressure signals at the microphones in the reference situation is defined as follows:

$$\begin{aligned}&\mathbf{H}_{\mathrm{ref_{pm}}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{pm_{ref}}}) = \\ &\quad = \frac{\mathbf{P}_{\mathrm{pm}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{pm_{ref}}})}{S_{\mathrm{ls}}}\end{aligned} \tag{10}$$

If the sound pressure at the microphones is assumed to represent the ear signals, the reference scene transfer functions $\mathbf{H}_{\mathrm{ref}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})$ can be approximated combining 6, 7, and 10:

$$\begin{aligned}\mathbf{H}_{\mathrm{ref}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}) &\approx \mathbf{H}_{\mathrm{ref_{pm}}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{pm_{ref}}}) \\ &= \frac{\mathbf{S}_{\mathrm{pm}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{pm_{ref}}})}{S_{\mathrm{ls}} \cdot \mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}_{\mathrm{i_{pm}}}} \\ &= H_{\mathrm{o_{ref}}} \cdot \mathbf{H}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{pm}}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{pm_{ref}}})\end{aligned} \tag{11}$$

2.6. Artificial Head

Approximative measurements of ear signals can also be done using an artificial head (AH) with microphones at the positions of the eardrums $\mathbf{x}_{\mathrm{ed}}$. The ear signal spectra are related to the artificial head microphone (*ahm*) output voltage spectra $\mathbf{U}_{\mathrm{ahm}}^{\mathrm{ah}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}})$ by the transfer functions

$$\mathbf{H}_{\mathrm{ahm}} = \frac{\mathbf{U}_{\mathrm{ahm}}^{\mathrm{ah}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}})}{\mathbf{P}_{\mathrm{ahm}}^{\mathrm{ah}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}})}. \tag{12}$$

The propagation parts between u_{ls} and $\mathbf{p}^{\mathrm{ah}}_{\mathrm{ahm}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}})$ are described by the transfer functions

$$\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}) = \frac{\mathbf{P}^{\mathrm{ah}}_{\mathrm{ahm}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}})}{U_{\mathrm{ls}}}. \tag{13}$$

Similarly, u_{ls} and $\mathbf{u}^{\mathrm{ah}}_{\mathrm{ahm}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}})$ are connected by

$$\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{u}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}) = \frac{\mathbf{U}^{\mathrm{ah}}_{\mathrm{ahm}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}})}{U_{\mathrm{ls}}}. \tag{14}$$

The relation of the sequence driving the loudspeaker to the sound pressures at the artificial head microphones in the reference situation is then defined by

$$\mathbf{H}^{\mathrm{ah}}_{\mathrm{ref}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}}) = \frac{\mathbf{P}^{\mathrm{ah}}_{\mathrm{ahm}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})}{S_{\mathrm{ls}}}. \tag{15}$$

The transfer functions $\mathbf{H}^{\mathrm{ind}}_{\mathrm{ref}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})$ describing the reference scene (given in 4) can be approximated as follows (with 6, 12, and 15):

$$\begin{aligned}\mathbf{H}^{\mathrm{ind}}_{\mathrm{ref}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}}) &\approx \mathbf{H}^{\mathrm{ah}}_{\mathrm{ref}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})\\ &= \frac{\mathbf{S}^{\mathrm{ah}}_{\mathrm{ahm}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})}{S_{\mathrm{ls}} \cdot \mathbf{H}_{\mathrm{i}_{\mathrm{ah}}} \cdot \mathbf{H}_{\mathrm{ahm}}}\\ &= H_{\mathrm{o}_{\mathrm{ref}}} \cdot \mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})\end{aligned} \tag{16}$$

3. RECORDING SITUATION

For recording, the propagation path impulse responses for BS with the defined reference, the loudspeaker of the reference situation has to be used. The recording can be done with probe microphones in the auditory canals or using an artificial head. It is also common practice to place miniature microphones (index m) at the entrances of a human subject's blocked auditory canals, since, according to Hammershøi and Møller (1996), at these points all directional information is present but as little individual information is included as possible. That should lead to as little error as possible if recordings from one subject are used for reproduction to others. An additional advantage of this procedure is the reduced complexity, since measurements at the entrances of the auditory canals are more easily done than measurements at the eardrums. Further, miniature microphones can be used, which provide a larger signal to noise ratio and more frequency independent transfer functions than probe microphones. In any case, the recordings contain transfer characteristics of three different systems: First the loudspeaker output system H_{o} (cf. section 2.1), second the input system $\mathbf{H}_{\mathrm{i}}$ (cf. section 2.4), and third the parts of the sound propagation paths $\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{u}_{\mathrm{mic}}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{mic}})$ captured by the selected recording setup. The transfer functions $\mathbf{H}^{\mathrm{ind}}_{\mathrm{rec}_{\mathrm{mic}}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{mic}})$ describing the connection between the digital output s_{ls} and the input sequences $\mathbf{s}^{\mathrm{ind}}_{\mathrm{mic}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{mic}})$ stemming from the microphones in the recording situation (index *rec*) are therefore given as follows:

$$\begin{aligned}&\mathbf{H}^{\mathrm{ind}}_{\mathrm{rec}_{\mathrm{mic}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{mic}_{\mathrm{rec}}}) =\\ &= \frac{\mathbf{S}^{\mathrm{ind}}_{\mathrm{mic}}(\mathbf{x}_{\mathrm{h}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{mic}_{\mathrm{rec}}})}{S_{\mathrm{ls}}}\\ &= H_{\mathrm{o}_{\mathrm{rec}}} \cdot \mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{u}_{\mathrm{mic}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{mic}_{\mathrm{rec}}}) \cdot \mathbf{H}_{\mathrm{i}_{\mathrm{rec}}}\end{aligned} \tag{17}$$

Usage of the reference situation output equipment ($H_{\mathrm{o}_{\mathrm{rec}}} = H_{\mathrm{o}_{\mathrm{ref}}}$) is taken for granted in the following, the usage of different output equipment could be taken into account by not assuming this equality.

3.1. Probe Microphone Recording

The transfer functions describing the recording situation with probe microphones are given using 7 and 17:

$$\begin{aligned}&\mathbf{H}^{\mathrm{ind}}_{\mathrm{rec}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{pm}_{\mathrm{rec}}}) = H_{\mathrm{o}_{\mathrm{rec}}} \cdot\\ &\cdot \mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{u}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{pm}_{\mathrm{rec}}}) \cdot \mathbf{H}_{\mathrm{i}_{\mathrm{pm}}}\\ &= H_{\mathrm{o}_{\mathrm{rec}}} \cdot \mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{rec}}}, \mathbf{x}_{\mathrm{pm}_{\mathrm{rec}}}) \cdot\\ &\quad \cdot \mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}_{\mathrm{i}_{\mathrm{pm}}}\end{aligned} \tag{18}$$

These transfer functions can be measured (with the usual difficulties in probe microphone measurements in the auditory canal) and are sometimes used for BS (cf. Wightman and Kistler 1989). They are under the assumptions described approximately identical to $\mathbf{H}^{\mathrm{ind}}_{\mathrm{ref}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})$ given in 4, multiplied by the input system and microphone transfer functions, if the loudspeaker, subject, and probe microphone positions are each coincident (cf. 11 and 17):

$$\begin{aligned}&\mathbf{H}^{\mathrm{ind}}_{\mathrm{rec}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{pm}_{\mathrm{rec}}}) \approx\\ &\qquad \approx \mathbf{H}^{\mathrm{ind}}_{\mathrm{ref}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}}, \mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}}) \cdot \mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}_{\mathrm{i}_{\mathrm{pm}}}\end{aligned} \tag{19}$$

3.2. Blocked Auditory Canal Recording

Measurements in the blocked auditory canals are indicated by the index b. The sound pressure spectra at the microphones $\mathbf{P}_{\mathrm{m}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{m}})$ are related to the microphone output voltages $\mathbf{U}_{\mathrm{m}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{m}})$ by the microphone transfer functions $\mathbf{H}_{\mathrm{m}}$:

$$\mathbf{H}_{\mathrm{m}} = \frac{\mathbf{U}_{\mathrm{m}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{m}})}{\mathbf{P}_{\mathrm{m}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{m}})} \tag{20}$$

The propagation paths between u_{ls} and the sound pressures $\mathbf{p}_{\mathrm{m}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{m}})$ in the blocked auditory canals are described by the transfer functions

$$\mathbf{H}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{m}}) = \frac{\mathbf{P}_{\mathrm{m}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{m}})}{U_{\mathrm{ls}}}. \tag{21}$$

In a similar manner, the propagation from u_{ls} to $\mathbf{u}_{\mathrm{m}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{m}})$ is represented by

$$\mathbf{H}_{u_{\mathrm{ls}},\mathbf{u}_{\mathrm{m}}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{m}}) = \frac{\mathbf{U}_{\mathrm{m}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h}}, \mathbf{x}_{\mathrm{ls}}, \mathbf{x}_{\mathrm{m}})}{U_{\mathrm{ls}}}. \tag{22}$$

The transfer functions for recording in the blocked auditory canals are then given by

$$\begin{aligned} &\mathbf{H}_{\mathrm{rec_m}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h_{rec}}}, \mathbf{x}_{\mathrm{ls_{rec}}}, \mathbf{x}_{\mathrm{m_{rec}}}) = \\ &= H_{\mathrm{o_{rec}}} \cdot \mathbf{H}_{u_{\mathrm{ls}},\mathbf{u}_{\mathrm{m}}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h_{rec}}}, \mathbf{x}_{\mathrm{ls_{rec}}}, \mathbf{x}_{\mathrm{m_{rec}}}) \cdot \mathbf{H}_{\mathrm{i_m}} \\ &= H_{\mathrm{o_{rec}}} \cdot \mathbf{H}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h_{rec}}}, \mathbf{x}_{\mathrm{ls_{rec}}}, \mathbf{x}_{\mathrm{m_{rec}}}) \cdot \\ &\quad \cdot \mathbf{H}_{\mathrm{m}} \cdot \mathbf{H}_{\mathrm{i_m}}. \end{aligned} \tag{23}$$

The latter transfer functions can be measured and are used in most implementations of BS (Hammershøi and Møller 2002). In general, they are different from $\mathbf{H}_{\mathrm{ref}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})$, even if the loudspeaker and subject positions are each coincident:

$$\begin{aligned} &\mathbf{H}_{\mathrm{rec_m}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{m_{rec}}}) = \mathbf{H}_{\mathrm{ref}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}) \cdot \\ &\quad \cdot \frac{\mathbf{H}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}^{\mathrm{ind,b}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{m_{rec}}})}{\mathbf{H}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{e}}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})} \cdot \mathbf{H}_{\mathrm{m}} \cdot \mathbf{H}_{\mathrm{i_m}} \end{aligned} \tag{24}$$

3.3. Artificial Head Recording

When using an artificial head, the following transfer functions describe the recording situation:

$$\begin{aligned} &\mathbf{H}_{\mathrm{rec_{ahm}}}^{\mathrm{ah}}(\mathbf{x}_{\mathrm{h_{rec}}}, \mathbf{x}_{\mathrm{ls_{rec}}}) = \\ &= H_{\mathrm{o_{rec}}} \cdot \mathbf{H}_{u_{\mathrm{ls}},\mathbf{u}_{\mathrm{ahm}}}^{\mathrm{ah}}(\mathbf{x}_{\mathrm{h_{rec}}}, \mathbf{x}_{\mathrm{ls_{rec}}}) \cdot \mathbf{H}_{\mathrm{i_{ah}}} \\ &= H_{\mathrm{o_{rec}}} \cdot \mathbf{H}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}^{\mathrm{ah}}(\mathbf{x}_{\mathrm{h_{rec}}}, \mathbf{x}_{\mathrm{ls_{rec}}}) \cdot \mathbf{H}_{\mathrm{ahm}} \cdot \mathbf{H}_{\mathrm{i_{ah}}} \end{aligned} \tag{25}$$

By comparing 25 to the reference scene transfer functions for artificial head usage (equation 16), the following connection can be derived:

$$\begin{aligned} &\mathbf{H}_{\mathrm{rec_{ahm}}}^{\mathrm{ah}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}) = \\ &\quad = \mathbf{H}_{\mathrm{ref_{ahm}}}^{\mathrm{ah}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}) \cdot \mathbf{H}_{\mathrm{ahm}} \cdot \mathbf{H}_{\mathrm{i_{ah}}} \end{aligned} \tag{26}$$

In general, these transfer functions are different from $\mathbf{H}_{\mathrm{ref}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})$, even if the loudspeaker and subject positions are each coincident:

$$\begin{aligned} &\mathbf{H}_{\mathrm{rec_{ahm}}}^{\mathrm{ah}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}) = \mathbf{H}_{\mathrm{ref}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}) \cdot \\ &\quad \cdot \frac{\mathbf{H}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}^{\mathrm{ah}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{e}}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})} \cdot \mathbf{H}_{\mathrm{ahm}} \cdot \mathbf{H}_{\mathrm{i_{ah}}} \end{aligned} \tag{27}$$

4. PLAYBACK SCENE

In the standard BS situation, the signals $\mathbf{s}_{\mathrm{hp}}$ are presented by headphones (at the positions $\mathbf{x}_{\mathrm{hp}}$ with the voltages $\mathbf{u}_{\mathrm{hp}}$ at their input terminals). Thereby, the transfer characteristics of the playback paths $\mathbf{H}_{\mathrm{play}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp_{play}}})$ are superimposed on the signals played back. It is also possible to use loudspeakers for signal presentation. In that case, it is necessary to assure that the signal intended to reach one ear does not reach the respective other ear. This procedure is fully covered by the introduced framework and could easily be included (for the necessary system theory cf. Völk et al. 2009), but is not taken into account here. When using headphones for playback, the path description consists of the transfer characteristics of the headphone (index *hp*) output equipment (see section 4.1 below), as well as of the sound propagation paths between the headphone input voltages and the sound pressures at the eardrums (under the headphones, superscript *h*):

$$\mathbf{H}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp}}) = \frac{\mathbf{P}_{\mathrm{e}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp}})}{\mathbf{U}_{\mathrm{hp}}} \tag{28}$$

The transfer functions accounting for the playback procedure (index *play*) are then given as follows:

$$\begin{aligned} \mathbf{H}_{\mathrm{play}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp_{play}}}) &= \frac{\mathbf{P}_{\mathrm{e}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp_{play}}})}{\mathbf{S}_{\mathrm{hp}}} \\ &= \mathbf{H}_{\mathrm{o_{hp}}} \cdot \mathbf{H}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp_{play}}}) \end{aligned} \tag{29}$$

4.1. Headphone Playback Equipment

The headphone output equipment consists of the audio interface used for conversion of $\mathbf{s}_{\mathrm{hp}}$ to $\mathbf{u}_{\mathrm{da}}$, as well as the headphone amplifiers:

$$\mathbf{H}_{\mathrm{da}} = \frac{\mathbf{U}_{\mathrm{da}}}{\mathbf{S}_{\mathrm{hp}}}, \quad \mathbf{H}_{\mathrm{ahp}} = \frac{\mathbf{U}_{\mathrm{hp}}}{\mathbf{U}_{\mathrm{da}}} \tag{30}$$

The transfer functions of the output equipment are then defined by the following equation:

$$\mathbf{H}_{\mathrm{o_{hp}}} = \frac{\mathbf{U}_{\mathrm{hp}}}{\mathbf{S}_{\mathrm{hp}}} = \mathbf{H}_{\mathrm{da}} \cdot \mathbf{H}_{\mathrm{ahp}} \tag{31}$$

4.2. Headphone Transfer Functions

Any attempt to measure the transfer functions accounting for the playback procedure must employ microphones in or attached to the auditory canals of a human or an artificial head. The transfer functions from the headphone driving sequences $\mathbf{s}_{\mathrm{hp}}$ to the corresponding microphone output sequences $\mathbf{s}_{\mathrm{mic}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp}}, \mathbf{x}_{\mathrm{mic}})$ are called Headphone Transfer Functions (HPTFs) here:

$$\mathbf{H}_{\mathrm{hptf}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp}}, \mathbf{x}_{\mathrm{mic}}) = \frac{\mathbf{S}_{\mathrm{mic}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp}}, \mathbf{x}_{\mathrm{mic}})}{\mathbf{S}_{\mathrm{hp}}} \tag{32}$$

The time domain representations are referred to as Headphone Impulse Responses (HPIRs).

4.3. Probe Microphone Measurement

An approximative description of the playback procedure is possible using probe microphones:

$$\begin{aligned} \mathbf{H}_{\mathrm{play}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp_{play}}}) &\approx \mathbf{H}_{\mathrm{play_{pm}}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{pm_{play}}}) \\ &= \frac{\mathbf{P}_{\mathrm{pm}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{pm_{play}}})}{\mathbf{S}_{\mathrm{hp}}} \\ &= \mathbf{H}_{\mathrm{o_{hp}}} \cdot \mathbf{H}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{pm}}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{pm_{play}}}) \end{aligned} \tag{33}$$

The HPTFs for probe microphone measurements are given using equation 32 by

$$\begin{aligned} &\mathbf{H}_{\mathrm{hptf_{pm}}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp}}, \mathbf{x}_{\mathrm{pm}}) = \\ &= \mathbf{H}_{\mathrm{o_{hp}}} \cdot \mathbf{H}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{pm}}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp}}, \mathbf{x}_{\mathrm{pm}}) \cdot \mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}_{\mathrm{i_{pm}}}. \end{aligned} \tag{34}$$

4.4. Blocked Auditory Canal Measurement

Another method for approximative measurements of the transfer functions describing the playback procedure (eq. 29) is using miniature microphones at the entrances to the blocked auditory canals. In that case, the sound pressures $\mathbf{p}_{\mathrm{m}}^{\mathrm{ind,b,h}}(\mathbf{x}_{\mathrm{hp}}, \mathbf{x}_{\mathrm{m}})$ detected by the microphones are always different from the signals of interest $\mathbf{p}_{\mathrm{e}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp}})$, which are not present when the auditory canals are blocked. For that reason, it is not possible to measure the transfer functions $\mathbf{H}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp}}, \mathbf{x}_{\mathrm{hp_b}}, \mathbf{x}_{\mathrm{m_b}})$ connecting the spectra of the two aforementioned sound pressure signals directly. An indirect definition is possible by

$$\begin{aligned} &\mathbf{H}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp}}, \mathbf{x}_{\mathrm{hp_b}}, \mathbf{x}_{\mathrm{m_b}}) = \\ &= \frac{\mathbf{P}_{\mathrm{e}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp}})}{\mathbf{P}_{\mathrm{m}}^{\mathrm{ind,b,h}}(\mathbf{x}_{\mathrm{hp}}, \mathbf{x}_{\mathrm{m}})} = \frac{\mathbf{H}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp}})}{\mathbf{H}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{m}}}^{\mathrm{ind,h,b}}(\mathbf{x}_{\mathrm{hp}}, \mathbf{x}_{\mathrm{m}})}. \end{aligned} \tag{35}$$

It is possible to compute the transfer functions describing the playback procedure using 29 and 35:

$$\begin{aligned} \mathbf{H}_{\mathrm{play}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp_{play}}}) &= \frac{\mathbf{S}_{\mathrm{m}}^{\mathrm{ind,b,h}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{m_{play}}})}{\mathbf{S}_{\mathrm{hp}} \cdot \mathbf{H}_{\mathrm{m}} \cdot \mathbf{H}_{\mathrm{i_m}}} \cdot \\ &\cdot \mathbf{H}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{hp_{play,b}}}, \mathbf{x}_{\mathrm{m_{play,b}}}) \end{aligned} \tag{36}$$

For blocked auditory canal miniature microphone measurement, the HPTFs are given using 29:

$$\begin{aligned} \mathbf{H}_{\mathrm{hptf_m}}^{\mathrm{ind,h,b}}(\mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{m_{hptf}}}) &= \mathbf{H}_{\mathrm{o_{hp}}} \cdot \mathbf{H}_{\mathrm{m}} \cdot \mathbf{H}_{\mathrm{i_m}} \cdot \\ &\cdot \frac{\mathbf{H}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp_{hptf}}})}{\mathbf{H}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{m_{hptf}}})} \end{aligned} \tag{37}$$

4.5. Artificial Head Measurement

Another approximation of the playback procedure can be done using an artificial head:

$$\begin{aligned} \mathbf{H}_{\mathrm{play}}^{\mathrm{ind,h}}(\mathbf{x}_{\mathrm{hp_{play}}}) &\approx \mathbf{H}_{\mathrm{play}}^{\mathrm{ah,h}}(\mathbf{x}_{\mathrm{hp_{play}}}) \\ &= \frac{\mathbf{S}_{\mathrm{ahm}}^{\mathrm{ah,h}}(\mathbf{x}_{\mathrm{hp_{play}}})}{\mathbf{S}_{\mathrm{hp}} \cdot \mathbf{H}_{\mathrm{ahm}} \cdot \mathbf{H}_{\mathrm{i_{ah}}}} \\ &= \mathbf{H}_{\mathrm{o_{hp}}} \cdot \mathbf{H}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}^{\mathrm{ah,h}}(\mathbf{x}_{\mathrm{hp_{play}}}) \end{aligned} \tag{38}$$

The HPTFs using an artificial head are given by

$$\begin{aligned} &\mathbf{H}_{\mathrm{hptf_{ahm}}}^{\mathrm{ah,h}}(\mathbf{x}_{\mathrm{hp_{hptf}}}) = \\ &= \mathbf{H}_{\mathrm{o_{hp}}} \cdot \mathbf{H}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}^{\mathrm{ah,h}}(\mathbf{x}_{\mathrm{hp_{hptf}}}) \cdot \mathbf{H}_{\mathrm{ahm}} \cdot \mathbf{H}_{\mathrm{i_{ah}}}. \end{aligned} \tag{39}$$

5. NON-EQUALIZED BINAURAL SYNTHESIS

Convolving an audio signal with the impulse responses recorded according to section 3 and presenting the convolution products to a listener by headphones as described in section 4 is here referred to as the (temporary) non-equalized BS scene (*tmp*). Using equations 17 and 29, the ear signals occurring in this non-equalized BS situation can be computed. It is assumed that $\mathbf{x}_{\mathrm{h_{rec}}} = \mathbf{x}_{\mathrm{h_{ref}}}$ and $\mathbf{x}_{\mathrm{ls_{rec}}} = \mathbf{x}_{\mathrm{ls_{ref}}}$ hold since a synthesis of the reference scene is intended. Then, the following ear signal spectra occur:

$$\begin{aligned} &\mathbf{P}^{\mathrm{ind,h}}_{\mathrm{e_{ne}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{mic_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}}) = S_{\mathrm{ls}} \cdot \\ &\cdot \mathbf{H}^{\mathrm{ind}}_{\mathrm{rec_{mic}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{mic_{rec}}}) \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathrm{play}}(\mathbf{x}_{\mathrm{hp_{play}}}) \end{aligned} \tag{40}$$

Comparing 4 and 40, it is obvious that the ear signals occurring in the non-equalized BS situation are not identical to those occurring in the reference scene. In the following, possible combinations of recording and playback situations are formulated. Therefore, for all considered situations, the transfer functions

$$\begin{aligned} &\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{ne}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{mic_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}}) = \\ &= \frac{\mathbf{P}^{\mathrm{ind,h}}_{\mathrm{e_{ne}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{mic_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}})}{S_{\mathrm{ls}}} \\ &= \mathbf{H}^{\mathrm{ind}}_{\mathrm{rec_{mic}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{mic_{rec}}}) \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathrm{play}}(\mathbf{x}_{\mathrm{hp_{play}}}), \end{aligned} \tag{41}$$

describing the connections between the source signal sequence and the ear signal spectra are given.

5.1. Non-Equalized Human Head Playback

Since the goal of binaural synthesis is the simulation of an auditory reference scene, the typical reproduction case is that with an individual human listener. When recordings from an individual head are used for human head playback, the same subject can be used for recording and playback (individual recording), but it is further possible to use a different subject for recording (non-individual recording).

5.1.1. Probe Microphone Recording

Combining 19 and 29 leads to the transfer functions giving the connection between the source signal and the ear signals when using probe microphones for recording and a human head for playback:

$$\begin{aligned} &\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{ne_{pm}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{pm_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}}) = \\ &= H_{\mathrm{O_{ref}}} \cdot \mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}}, \mathbf{p}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{pm_{rec}}}) \cdot \\ &\quad \cdot \mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}_{\mathrm{i_{pm}}} \cdot \mathbf{H}_{\mathrm{O_{hp}}} \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}}, \mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp_{play}}}) \end{aligned} \tag{42}$$

The transfer functions are given here with respect to the reference scene (approximated by probe microphone measurement) to make deviations from the intended goal clearly visible. Assuming the sound pressures at the probe microphones during the recording represent the ear signals, following approximation holds:

$$\begin{aligned} &\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{ne_{pm}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{pm_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}}) \approx \\ &\approx \mathbf{H}^{\mathrm{ind}}_{\mathrm{ref}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}) \cdot \mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}_{\mathrm{i_{pm}}} \cdot \mathbf{H}_{\mathrm{O_{hp}}} \cdot \\ &\quad \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}}, \mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp_{play}}}) \end{aligned} \tag{43}$$

5.1.2. Blocked Auditory Canal Recording

The transfer functions connecting source and ear signal spectra when presenting blocked auditory canal recordings to a human subject are given combining 24 and 29:

$$\begin{aligned} &\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{ne_{m}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{m_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}}) = \\ &= \mathbf{H}^{\mathrm{ind}}_{\mathrm{ref}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}) \cdot \frac{\mathbf{H}^{\mathrm{ind,b}}_{u_{\mathrm{ls}}, \mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{m_{rec}}})}{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}}, \mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})} \cdot \\ &\quad \cdot \mathbf{H}_{\mathrm{m}} \cdot \mathbf{H}_{\mathrm{i_{m}}} \cdot \mathbf{H}_{\mathrm{O_{hp}}} \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}}, \mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp_{play}}}) \end{aligned} \tag{44}$$

5.1.3. Artificial Head Recording

If an artificial head is used for recording, the playback on a human head can be described by a combination of 27 and 29:

$$\begin{aligned} &\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{ne_{ah}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{hp_{play}}}) = \\ &= \mathbf{H}^{\mathrm{ind}}_{\mathrm{ref}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}) \cdot \frac{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}}, \mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}}, \mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})} \cdot \\ &\quad \cdot \mathbf{H}_{\mathrm{ahm}} \cdot \mathbf{H}_{\mathrm{i_{ah}}} \cdot \mathbf{H}_{\mathrm{O_{hp}}} \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}}, \mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp_{play}}}) \end{aligned} \tag{45}$$

5.2. Non-Equalized Artificial Head Playback

In a addition to the playback on a human head, reproduction is possible on an artificial head. The latter is an essential step in the design and implementation processes of binaural synthesis systems and is therefore considered in the following paragraphs.

5.2.1. Probe Microphone Recording

For probe microphone recording, equations 18 and 29 can be combined to describe the playback situation on an artificial head by the transfer functions relating the source and ear signal spectra:

$$\begin{aligned} &\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{ne_{pm}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{pm_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}}) = \\ &= H_{\mathrm{O_{ref}}} \cdot \mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}}, \mathbf{p}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{pm_{rec}}}) \cdot \mathbf{H}_{\mathrm{pm}} \\ &\quad \cdot \mathbf{H}_{\mathrm{i_{pm}}} \cdot \mathbf{H}_{\mathrm{O_{hp}}} \cdot \mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}}, \mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}}) \end{aligned} \tag{46}$$

To display the transfer functions in relation to the artificial head reference, comparison to 16 gives

$$\begin{aligned}
&\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{ne_{pm}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{pm_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}}) = \\
&\quad = \mathbf{H}^{\mathrm{ah}}_{\mathrm{ref_{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}})\cdot \\
&\qquad \cdot \frac{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{pm_{rec}}})}{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}})}\cdot \\
&\qquad \cdot \mathbf{H}_{\mathrm{pm}}\cdot\mathbf{H}_{\mathrm{i_{pm}}}\cdot\mathbf{H}_{\mathrm{o_{hp}}}\cdot\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}}).
\end{aligned} \tag{47}$$

5.2.2. Blocked Auditory Canal Recording

Presenting the measurement results of the blocked auditory canal situation to an artificial head can be described by combining 23 and 29:

$$\begin{aligned}
&\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{ne_{m}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{m_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}}) = \\
&\quad = H_{\mathrm{o_{ref}}}\cdot\mathbf{H}^{\mathrm{ind,b}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{m_{rec}}})\cdot \\
&\qquad \cdot\mathbf{H}_{\mathrm{m}}\cdot\mathbf{H}_{\mathrm{i_{m}}}\cdot\mathbf{H}_{\mathrm{o_{hp}}}\cdot\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}})
\end{aligned} \tag{48}$$

A relation to the desired reference scene can be given by comparison to equation 16:

$$\begin{aligned}
&\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{ne_{m}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{m_{rec}}},\mathbf{x}_{\mathrm{hp_{play}}}) = \\
&= \mathbf{H}^{\mathrm{ah}}_{\mathrm{ref_{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}})\cdot\frac{\mathbf{H}^{\mathrm{ind,b}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{m_{rec}}})}{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}})}\cdot \\
&\quad\cdot\mathbf{H}_{\mathrm{m}}\cdot\mathbf{H}_{\mathrm{i_{m}}}\cdot\mathbf{H}_{\mathrm{o_{hp}}}\cdot\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}})
\end{aligned} \tag{49}$$

5.2.3. Artificial Head Recording

The most common situation in the development process of binaural synthesis is the presentation of artificial head recordings to an artificial head. Combining 26 and 29 gives a description of this situation:

$$\begin{aligned}
&\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{ne_{ah}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{hp_{play}}}) = \\
&= \mathbf{H}^{\mathrm{ah}}_{\mathrm{ref_{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}})\cdot\mathbf{H}_{\mathrm{ahm}}\cdot \\
&\quad\cdot\mathbf{H}_{\mathrm{i_{ah}}}\cdot\mathbf{H}_{\mathrm{o_{hp}}}\cdot\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}})
\end{aligned} \tag{50}$$

6. EQUALIZATION REQUIREMENTS

Usually, the correct reproduction of a reference situation's ear signals $\mathbf{p}^{\mathrm{ind}}_{\mathrm{e}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}})$ is assumed to be sufficient to recreate the auditory impression of the reference situation (cf. e. g. Møller 1992). This requirement seems completely sufficient since according to Fastl and Zwicker (2007) the human ear acts as pressure receiver in a wide frequency range. However, there is some uncertainty about this requirement, addressed in the following. Afterwards, the proper equalization for each of the introduced binaural synthesis situations is derived and discussed.

6.1. Necessary Premises

If headphones are used for reproduction of the computed signals in the binaural synthesis process, it is currently not completely clear if the claim of equal sound pressure signals in the auditory canals in reference and synthesis scene is a requirement sufficient to assure the same auditory impression in both situations. Some authors found a remarkable difference in the sound pressure levels measured in the auditory canals when the same perceived loudness is elicited for presentation over loudspeaker and headphones (cf. e. g. Fastl et al. 1985, Fastl 1986, Sivian and White 1933, Munson and Wiener 1952), whereas others could not confirm this difference (cf. Rudmose 1982). The latter author explains the difference between his and others' measurements by errors due to methodical shortcomings in the procedure.

Here, a different possible explanation for this somehow unexpected effect is presented: Comparing the situations with the same sound pressure levels caused by headphone and loudspeaker reproduction does not assure the same ear signals (time dependent sound pressure signals at the eardrums) in both situations. Since the recreation of the loudspeaker scene ear signals is the ultimate goal of binaural synthesis, it is to be expected that a repetition of the experiments conducted e. g. by Fastl et al. (1985) using binaural synthesis would be capable of reproducing the same loudness perception as the reference scene using headphones at identical sound pressure levels in the auditory canals. This explanation was formulated in a more descriptive manner by Theile (1980) as the so called *association principle*. This principle could be summed up in view of the problem considered here as follows: the auditory impression of the loudspeaker scene includes its perceived position (the so called auditory event position), which is different in the (traditional) headphone listening situation, since in that case the auditory events are located in the head (lateralization, cf. Jeffress and Taylor 1961, Plenge 1974). It could be concluded from the association principle that the same loudness perception (and the same overall auditory impression) can only be elicited if the auditory event position is identical in both cases. Based on these considerations, the recreation of the reference scene ear signals is assumed as the ultimate goal of binaural synthesis in the following.

6.2. Attempted Goal: Reproduction of Ear Signals

In section 5 it has been shown that the ear signals in the BS situation must equal the ear signals of the reference scene (cf. 4 and 40):

$$\begin{aligned}\mathbf{P}^{\mathrm{ind,h}}_{\mathrm{e_{bs}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{mic_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}}) &\overset{!}{=} \\ \overset{!}{=} \quad \mathbf{P}^{\mathrm{ind}}_{\mathrm{e}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})\end{aligned} \tag{51}$$

To reach this goal, in any case correction filters $\mathbf{h}^{\mathrm{ind}}_{\mathrm{eq}}$ have to be applied so that following equation holds:

$$\begin{aligned}&\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{ne}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{mic_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}})\cdot \\ &\cdot \mathbf{H}^{\mathrm{ind}}_{\mathrm{eq}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{mic_{rec}}}) = \mathbf{H}^{\mathrm{ind}}_{\mathrm{ref}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})\end{aligned} \tag{52}$$

The required equalization filters are therefore in general given as follows:

$$\begin{aligned}&\mathbf{H}^{\mathrm{ind}}_{\mathrm{eq}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{mic_{rec}}}) = \\ &= \frac{\mathbf{H}^{\mathrm{ind}}_{\mathrm{ref}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{ne}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{mic_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}})}\end{aligned} \tag{53}$$

In the following paragraphs, the required equalization filters for system theoretically correct BS are given for the different situations defined in the preceding section. In addition, their relation to the different kinds of HPTFs introduced in section 4 is derived, since a practical acquirement of equalization filters is commonly done by recording and inverting distinct HPIRs. It is taken for granted in the following that the same input equipment is used for the blocked auditory canal, the probe microphone, and the artificial head measurements, i. e. $\mathbf{H}_{\mathrm{i_m}} = \mathbf{H}_{\mathrm{i_{pm}}} = \mathbf{H}_{\mathrm{i_{ah}}}$. Further, the microphone transfer functions are neglected (or assumed to be approximately equivalent, i. e. $\mathbf{H}_{\mathrm{pm}} \approx \mathbf{H}_{\mathrm{m}} \approx \mathbf{H}_{\mathrm{ahm}}$).

6.3. Equalization for Human Head Playback

The eventual use case of binaural synthesis is the generation of the auditory events of a certain reference scene. Therefore, playback to human subjects is of major interest and is discussed in this section for the different recording situations. After adjusting the equalization requirements to the respective recording situation, the possible degree to which the requirements can be met in each case by using the different headphone impulse responses introduced is assessed.

6.3.1. Probe Microphone Recording

First, recording using probe microphones right in front of a listener's eardrums is considered. The equalization requirements formulated in 53 can be specialized for the playback of probe microphone recordings on a human head using 42:

$$\begin{aligned}&\mathbf{H}^{\mathrm{ind}}_{\mathrm{eq_{pm}}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{pm_{rec}}}) \approx \\ &\approx \frac{1}{\mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}_{\mathrm{i_{pm}}} \cdot \mathbf{H}_{\mathrm{o_{hp}}} \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp_{play}}})}\end{aligned} \tag{54}$$

This requirement can be approximated, as introduced before, using HPIRs acquired with probe microphones (4.3), blocked auditory canal measurements (4.4), or artificial heads (4.5).

Probe microphone headphone impulse responses With equation 34, 54 yields the relation of the equalization requirements for probe microphone measurement and human head playback to the HPTFs measured using probe microphones:

$$\begin{aligned}&\mathbf{H}^{\mathrm{ind}}_{\mathrm{eq_{pm}}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{pm_{rec}}}) \approx \\ &\approx \frac{1}{\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hptf_{pm}}}(\mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{pm_{hptf}}})}\end{aligned} \tag{55}$$

This could be read as follows: If probe microphones are used for all measurements, it is possible to equalize a BS system with the inverses of the HPIRs. Then, the synthesis equals the reference scene if following requirements are fulfilled:

- The sound pressure signals at the probe microphones right in front of the eardrums represent the ear signals to a degree sufficient.
- The probe microphone positions are the same for recording and HPIR measurement.
- The headphones are at the same position for playback as they were during the HPIR measurement.

This situation is the most direct way to asses binaural synthesis, but can never be proven completely for validity, since the physical measurement of an ear signal in its strict sense is not possible.

Blocked auditory canal headphone impulse responses With 37, 54 can be rewritten to describe the connection from the equalization requirements for probe microphone measurements and human head playback to the HPTFs measured in the blocked auditory canals:

$$\begin{aligned}\mathbf{H}^{\mathrm{ind}}_{\mathrm{eq_{pm}}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{pm_{rec}}}) \approx \\ \approx \frac{1}{\mathbf{H}^{\mathrm{ind,h,b}}_{\mathrm{hptf_m}}(\mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{m_{hptf}}})} \cdot \frac{\mathbf{H}_{\mathrm{m}} \cdot \mathbf{H}_{\mathrm{i_m}}}{\mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}_{\mathrm{i_{pm}}}} \cdot \\ \cdot \frac{1}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{m_{hptf}}})}\end{aligned} \tag{56}$$

Therefore, binaural synthesis using probe microphone recordings and human head playback could, with a remaining error, be equalized using HPIRs measured in the blocked auditory canals. The remaining error is given by the transfer functions from the sound pressures at the eardrums with free auditory canals to the sound pressure at miniature microphones in the blocked auditory canals. Since probe microphones were available and used for recording, they can also be used for the HPIR-measurements.

Artificial head headphone impulse responses With 39, 54 yields the link from the necessary equalization for probe microphone measurement and human head playback to the HPTFs measured using an artificial head:

$$\begin{aligned}\mathbf{H}^{\mathrm{ind}}_{\mathrm{eq_{pm}}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{pm_{rec}}}) \approx \frac{1}{\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{hptf_{ahm}}}(\mathbf{x}_{\mathrm{hp_{hptf}}})} \cdot \\ \cdot \frac{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{hptf}}})}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp_{play}}})} \cdot \frac{\mathbf{H}_{\mathrm{ahm}} \cdot \mathbf{H}_{\mathrm{i_{ah}}}}{\mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}_{\mathrm{i_{pm}}}}\end{aligned} \tag{57}$$

This means binaural synthesis using probe microphone measurements played back to a human listener could be equalized approximately by HPIRs measured using an artificial head. The error of this approximation is given by the deviation between the transfer functions from the headphones to the artificial and human head ears. Since a measurement of the latter is only possible approximately and even then difficult, the error can only be determined in listening experiments, e. g. in loudness adjustments (as mentioned in section 6.1).

6.3.2. Blocked Auditory Canal Recording

The second recording case to be considered for human playback is recording with miniature microphones in the blocked auditory canals. Using 44 and 53, the equalization requirement for playback of blocked auditory canal recordings for a human subject is given by

$$\begin{aligned}\mathbf{H}^{\mathrm{ind,b}}_{\mathrm{eq_m}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{m_{rec}}}) = \\ = \frac{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ind,b}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{m_{rec}}})} \cdot \\ \cdot \frac{1}{\mathbf{H}_{\mathrm{m}} \cdot \mathbf{H}_{\mathrm{i_m}} \cdot \mathbf{H}_{\mathrm{o_{hp}}} \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp_{play}}})}.\end{aligned} \tag{58}$$

In the following, this requirement is related to the considered headphone impulse responses.

Probe microphone headphone impulse responses With 34, 58 yields the connection of the equalization necessary for playback of blocked auditory canal recordings to a human subject to the HPTFs measured using probe microphones:

$$\begin{aligned}\mathbf{H}^{\mathrm{ind,b}}_{\mathrm{eq_m}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{m_{rec}}}, \mathbf{x}_{\mathrm{pm_{hptf}}}) \approx \\ \approx \frac{1}{\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hptf_{pm}}}(\mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{pm_{hptf}}})} \cdot \\ \cdot \frac{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ind,b}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{m_{rec}}})} \cdot \frac{\mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}_{\mathrm{i_{pm}}}}{\mathbf{H}_{\mathrm{m}} \cdot \mathbf{H}_{\mathrm{i_m}}}\end{aligned} \tag{59}$$

The equalization of blocked auditory canal recordings with headphone impulse responses measured using probe microphones leads to an error caused by the different observation points of the sound pressure in the auditory canals.

Blocked auditory canal headphone impulse responses The situation using miniature microphones in the blocked auditory canals for every involved measurement while playing back to a human head is assessed now. With 37, 58 yields the relation of the equalization requirements for the playback of blocked auditory canal recordings to a human subject to headphone transfer functions measured in the blocked auditory canals:

$$\mathbf{H}^{\mathrm{ind,b}}_{\mathrm{eq_m}}(\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{m_{rec}}}) = \frac{1}{\mathbf{H}^{\mathrm{ind,h,b}}_{\mathrm{hptf_m}}(\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{m_{hptf}}})}\cdot \\ \cdot\frac{1}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}(\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{m_{hptf}}})}\cdot \\ \cdot\frac{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ind,b}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{m_{rec}}})} \tag{60}$$

The transfer functions between the sound pressures at the microphones in the blocked auditory canals and at the eardrums with free auditory canals is (similar to equation 35) given as follows:

$$\mathbf{H}^{\mathrm{ind}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{ls}}(\mathbf{x}_{\mathrm{m_{rec}}}) = \frac{\mathbf{P}^{\mathrm{ind}}_{\mathrm{e}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{P}^{\mathrm{ind,b}}_{\mathrm{m}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{m_{rec}}})} = \frac{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ind,b}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{m_{rec}}})} \tag{61}$$

Then, 60 can be rewritten:

$$\mathbf{H}^{\mathrm{ind,b}}_{\mathrm{eq_m}}(\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{m_{rec}}}) = \frac{1}{\mathbf{H}^{\mathrm{ind,h,b}}_{\mathrm{hptf_m}}(\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{m_{hptf}}})}\cdot \\ \cdot\frac{\mathbf{H}^{\mathrm{ind}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{ls}}(\mathbf{x}_{\mathrm{m_{rec}}})}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}(\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{hp_{hptf}}},\mathbf{x}_{\mathrm{m_{hptf}}})} \tag{62}$$

This result is of special interest, because this situation is the most common implementation of individual BS. This synthesis equals the reference scene if inverted HPIRs measured in the blocked auditory canals are used as equalization filters and the following requirements are fulfilled:

- The microphone positions are identical for the recording and the headphone impulse response measurement situation.
- The headphones are at the same position for the playback as they were during the headphone impulse response measurement.
- $\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}(\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{m_{rec}}})$ defined in section 4.4 connecting the sound pressures at the miniature microphones in the blocked auditory canals and at the eardrums in the open auditory canals *under the headphones* has to be identical to $\mathbf{H}^{\mathrm{ind}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{ls}}(\mathbf{x}_{\mathrm{m_{rec}}})$ measured if no headphones are worn, as it is the case in the reference scene. In other words, for individual binaural synthesis using measurements with miniature microphones in the blocked auditory canals, headphones need to be used which do not alter the transformation from the sound pressure signals at the blocked auditory canals to the eardrums under the headphones compared to the situation without headphones.

Artificial head headphone impulse responses
Using 39, 58 can be rewritten to describe the connection between the equalization necessities for playback of blocked auditory canal recordings to a human subject and the headphone transfer functions acquired using an artificial head:

$$\mathbf{H}^{\mathrm{ind,b}}_{\mathrm{eq_m}}(\mathbf{x}_{\mathrm{hp_{play}}},\mathbf{x}_{\mathrm{m_{rec}}}) = \\ = \frac{1}{\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{hptf_{ahm}}}(\mathbf{x}_{\mathrm{hp_{hptf}}})}\cdot\frac{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{hptf}}})}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp_{play}}})}\cdot \\ \cdot\frac{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ind,b}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}},\mathbf{x}_{\mathrm{m_{rec}}})}\cdot\frac{\mathbf{H}_{\mathrm{ahm}}\cdot\mathbf{H}_{\mathrm{i_{ah}}}}{\mathbf{H}_{\mathrm{m}}\cdot\mathbf{H}_{\mathrm{i_m}}} \tag{63}$$

Here the difference between the human and the artificial head as well as the differences due to the different measurement points and circumstances are included.

6.3.3. Artificial Head Recording

The third and least complex recording method is an artificial head. Using 45 and 53, the equalization requirements for usage of artificial head recordings for playback to a human listener are given by

$$\mathbf{H}^{\mathrm{ind}}_{\mathrm{eq_{ah}}}(\mathbf{x}_{\mathrm{hp_{play}}}) = \frac{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}},\mathbf{x}_{\mathrm{ls_{ref}}})}\cdot \\ \cdot\frac{1}{\mathbf{H}_{\mathrm{ahm}}\cdot\mathbf{H}_{\mathrm{i_{ah}}}\cdot\mathbf{H}_{\mathrm{o_{hp}}}\cdot\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp_{play}}})}. \tag{64}$$

Probe microphone headphone impulse responses With 34, 64 yields the link between the equalization necessities for artificial head recordings played back to a human listener and probe microphone headphone transfer functions:

$$\mathbf{H}^{\mathrm{ind}}_{\mathrm{eq_{ah}}}(\mathbf{x}_{\mathrm{hp_{play}}}) \approx \frac{1}{\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hptf_{pm}}}(\mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{pm_{hptf}}})} \cdot \frac{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})} \cdot \frac{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{pm_{hptf}}})}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp_{play}}})} \cdot \frac{\mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}_{\mathrm{i_{pm}}}}{\mathbf{H}_{\mathrm{ahm}} \cdot \mathbf{H}_{\mathrm{i_{ah}}}} \tag{65}$$

The situation given in 65 suffers only from the deviation of the human from the artificial head.

Blocked auditory canal headphone impulse responses With 37 and 64, the connection between the equalization requirements for artificial head recordings played back to a human listener and HPTFs measured in the blocked auditory canals using miniature microphones can be computed:

$$\mathbf{H}^{\mathrm{ind}}_{\mathrm{eq_{ah}}}(\mathbf{x}_{\mathrm{hp_{play}}}) \approx \frac{1}{\mathbf{H}^{\mathrm{ind,h,b}}_{\mathrm{hptf_{m}}}(\mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{m_{hptf}}})} \cdot \frac{1}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{m_{hptf}}})} \cdot \frac{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})} \cdot \frac{\mathbf{H}_{\mathrm{m}} \cdot \mathbf{H}_{\mathrm{i_{m}}}}{\mathbf{H}_{\mathrm{ahm}} \cdot \mathbf{H}_{\mathrm{i_{ah}}}} \tag{66}$$

This equation describes the situation of artificial head recordings, individually equalized using the inverses of HPIRs measured using miniature microphones in the blocked auditory canals. It is evident that this solution is different from the reference scene even if the artificial head matches the human very closely, due to the different measurement points and circumstances.

Artificial head headphone impulse responses To get a relationship between the necessary equalization for artificial head recordings played back to a human listener and HPTFs measured using an artificial head, 39 and 64 can be combined as follows:

$$\mathbf{H}^{\mathrm{ind}}_{\mathrm{eq_{ah}}}(\mathbf{x}_{\mathrm{hp_{play}}}) = \frac{1}{\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{hptf_{ahm}}}(\mathbf{x}_{\mathrm{hp_{hptf}}})} \cdot \frac{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{hptf}}})}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp_{play}}})} \cdot \frac{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})} \tag{67}$$

This is the typical case when using artificial head recording in binaural synthesis. The degree to which the ear signals resemble the reference scene depends on how well the artificial head matches the human with regard to ear signals and headphone fit.

6.4. Equalization for Artificial Head Playback

As mentioned above, playback to an artificial head is an important tool in the design of binaural synthesis systems. For this reason, artificial head playback is considered here in detail in addition to human head playback. As done above for human head playback, after definition of equalization requirements in the respective recording situation, the degree to which the requirements can be met using the different headphone impulse responses introduced is assessed for each situation.

6.4.1. Probe Microphone Recording

The equalization requirement for playback of probe microphone measurements to an artificial head is given combining 47 and 53:

$$\mathbf{H}^{\mathrm{ah}}_{\mathrm{eq_{pm}}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{pm_{rec}}}) = \frac{1}{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}})} \cdot \frac{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{pm_{rec}}})} \cdot \frac{1}{\mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}_{\mathrm{i_{pm}}} \cdot \mathbf{H}_{\mathrm{o_{hp}}}} \tag{68}$$

As for human head playback, the equalization requirement is related to the different possibly occurring headphone impulse responses for artificial head playback in the following.

Probe microphone headphone impulse responses With 34, 68 yields the equalization necessities for playback of probe microphone recordings on an artificial head with regard to headphone transfer functions measured using probe microphones right in front of a subjects' eardrums:

$$\begin{aligned}\mathbf{H}^{\mathrm{ah}}_{\mathrm{eq}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}},\mathbf{x}_{\mathrm{pm}_{\mathrm{rec}}}) &= \\ &= \frac{1}{\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hptf}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}},\mathbf{x}_{\mathrm{pm}_{\mathrm{hptf}}})}\cdot \\ &\cdot \frac{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}},\mathbf{x}_{\mathrm{pm}_{\mathrm{hptf}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}})}\cdot \\ &\cdot \frac{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})}{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{pm}_{\mathrm{rec}}})}\end{aligned} \tag{69}$$

This result is of interest in that the deviation between artificial and human head would cancel if it were identical for recording and headphone impulse response measurement. This way, these aberrations can be quantified by measurement.

Blocked auditory canal headphone impulse responses To determine the equalization requirements for playback of probe microphone recordings on an artificial head dependent on HPTFs measured with miniature microphones in the blocked auditory canals, 37 and 68 can be combined:

$$\begin{aligned}\mathbf{H}^{\mathrm{ah}}_{\mathrm{eq}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}},\mathbf{x}_{\mathrm{pm}_{\mathrm{rec}}}) &= \frac{1}{\mathbf{H}^{\mathrm{ind,h,b}}_{\mathrm{hptf}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}},\mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}})}\cdot \\ &\cdot \frac{1}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}},\mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}},\mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}})}\cdot\frac{\mathbf{H}_{\mathrm{m}}\cdot\mathbf{H}_{\mathrm{i_m}}}{\mathbf{H}_{\mathrm{pm}}\cdot\mathbf{H}_{\mathrm{i_{pm}}}}\cdot \\ &\cdot \frac{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}})}\cdot\frac{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})}{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{pm}_{\mathrm{rec}}})}\end{aligned} \tag{70}$$

Artificial head headphone impulse responses With 39, 68 determines the equalization for playback of probe microphone recordings to an artificial head with regard to artificial head HPTFs:

$$\begin{aligned}\mathbf{H}^{\mathrm{ah}}_{\mathrm{eq}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}},\mathbf{x}_{\mathrm{pm}_{\mathrm{rec}}}) &\approx \frac{1}{\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{hptf}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}})}\cdot \\ &\cdot \frac{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})}{\mathbf{H}^{\mathrm{ind}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{pm}_{\mathrm{rec}}})}\cdot \\ &\cdot \frac{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}})}\cdot\frac{\mathbf{H}_{\mathrm{ahm}}\cdot\mathbf{H}_{\mathrm{i_{ah}}}}{\mathbf{H}_{\mathrm{pm}}\cdot\mathbf{H}_{\mathrm{i_{pm}}}}\end{aligned} \tag{71}$$

6.4.2. Blocked Auditory Canal Recording

Using equation 49 and equation 53, the equalization requirement for blocked auditory canal miniature microphone recordings played back to an artificial head is given:

$$\begin{aligned}\mathbf{H}^{\mathrm{ah}}_{\mathrm{eq}_{\mathrm{ah}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}},\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}}) &= \\ &= \frac{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})}{\mathbf{H}^{\mathrm{ind,b}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}})}\cdot \\ &\cdot \frac{1}{\mathbf{H}_{\mathrm{m}}\cdot\mathbf{H}_{\mathrm{i_m}}\cdot\mathbf{H}_{\mathrm{o_{hp}}}\cdot\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}})}\end{aligned} \tag{72}$$

Probe microphone headphone impulse responses With 34 and 72, the required equalization for blocked auditory canal recordings played back to an artificial head can be formulated in dependence of HPTFs measured with probe microphones:

$$\begin{aligned}\mathbf{H}^{\mathrm{ah}}_{\mathrm{eq}_{\mathrm{ah}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}},\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}}) &= \frac{\mathbf{H}_{\mathrm{pm}}\cdot\mathbf{H}_{\mathrm{i_{pm}}}}{\mathbf{H}_{\mathrm{m}}\cdot\mathbf{H}_{\mathrm{i_m}}}\cdot \\ &\cdot \frac{1}{\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hptf}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}},\mathbf{x}_{\mathrm{pm}_{\mathrm{hptf}}})}\cdot \\ &\cdot \frac{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}},\mathbf{x}_{\mathrm{pm}_{\mathrm{hptf}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}})}\cdot \\ &\cdot \frac{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})}{\mathbf{H}^{\mathrm{ind,b}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}})}\end{aligned} \tag{73}$$

Blocked auditory canal headphone impulse responses With 37, 72 yields the necessary equalization for blocked auditory canal recordings played back to an artificial head, using headphone transfer functions measured with miniature microphones in the blocked auditory canals:

$$\begin{aligned}\mathbf{H}^{\mathrm{ah}}_{\mathrm{eq}_{\mathrm{ah}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}},\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}}) &= \frac{1}{\mathbf{H}^{\mathrm{ind,h,b}}_{\mathrm{hptf}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}},\mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}})}\cdot \\ &\cdot \frac{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}})}\cdot \\ &\cdot \frac{1}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}(\mathbf{x}_{\mathrm{hp}_{\mathrm{play}}},\mathbf{x}_{\mathrm{hp}_{\mathrm{hptf}}},\mathbf{x}_{\mathrm{m}_{\mathrm{hptf}}})}\cdot \\ &\cdot \frac{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}})}{\mathbf{H}^{\mathrm{ind,b}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{ls}_{\mathrm{ref}}},\mathbf{x}_{\mathrm{m}_{\mathrm{rec}}})}\end{aligned} \tag{74}$$

Artificial head headphone impulse responses With 39, 72 yields equalization requirements for blocked auditory canal recordings played back to an artificial head with artificial head HPTFs:

$$\mathbf{H}^{\mathrm{ah}}_{\mathrm{eq_{ah}}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{m_{rec}}}) = \frac{1}{\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{hptf_{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}})} \cdot \cdot \frac{\mathbf{H}^{\mathrm{ah}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ind,b}}_{u_{\mathrm{ls}},\mathbf{p}_{\mathrm{m}}}(\mathbf{x}_{\mathrm{h_{ref}}}, \mathbf{x}_{\mathrm{ls_{ref}}}, \mathbf{x}_{\mathrm{m_{rec}}})} \cdot \cdot \frac{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{hptf}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}})} \cdot \frac{\mathbf{H}_{\mathrm{ahm}} \cdot \mathbf{H}_{\mathrm{i_{ah}}}}{\mathbf{H}_{\mathrm{m}} \cdot \mathbf{H}_{\mathrm{i_m}}} \tag{75}$$

6.4.3. Artificial Head Recording

Using 50 and 53, it is possible to formulate the equalization requirement for playback of artificial head recordings to an artificial head:

$$\mathbf{H}^{\mathrm{ah}}_{\mathrm{eq_{ah}}}(\mathbf{x}_{\mathrm{hp_{play}}}) = = \frac{1}{\mathbf{H}_{\mathrm{ahm}} \cdot \mathbf{H}_{\mathrm{i_{ah}}} \cdot \mathbf{H}_{\mathrm{o_{hp}}} \cdot \mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}})} \tag{76}$$

Probe microphone headphone impulse responses With 34, 76 can be formulated as follows to describe the required equalization for playback of artificial head recordings to an artificial head dependent on HPTFs measured using probe microphones:

$$\mathbf{H}^{\mathrm{ah}}_{\mathrm{eq_{ah}}}(\mathbf{x}_{\mathrm{hp_{play}}}) \approx \frac{1}{\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hptf_{pm}}}(\mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{pm_{hptf}}})} \cdot \cdot \frac{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{pm}}}(\mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{pm_{hptf}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}})} \cdot \frac{\mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}_{\mathrm{i_{pm}}}}{\mathbf{H}_{\mathrm{ahm}} \cdot \mathbf{H}_{\mathrm{i_{ah}}}} \tag{77}$$

Blocked auditory canal headphone impulse responses To relate the equalization requirements for playback of artificial head recordings to an artificial head with respect to HPTFs measured with miniature microphones in the blocked auditory canals, 37 and 76 are combined:

$$\mathbf{H}^{\mathrm{ah}}_{\mathrm{eq_{ah}}}(\mathbf{x}_{\mathrm{hp_{play}}}) = \frac{1}{\mathbf{H}^{\mathrm{ind,h,b}}_{\mathrm{hptf_m}}(\mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{m_{hptf}}})} \cdot \cdot \frac{1}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{m_{hptf}}})} \cdot \cdot \frac{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{e}}}(\mathbf{x}_{\mathrm{hp_{play}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}})} \cdot \frac{\mathbf{H}_{\mathrm{m}} \cdot \mathbf{H}_{\mathrm{i_m}}}{\mathbf{H}_{\mathrm{ahm}} \cdot \mathbf{H}_{\mathrm{i_{ah}}}} \tag{78}$$

Artificial head headphone impulse responses With 39, 76 yields the required equalization for playback of artificial head recordings to an artificial head dependent on the artificial head HPTFs:

$$\mathbf{H}^{\mathrm{ah}}_{\mathrm{eq_{ah}}}(\mathbf{x}_{\mathrm{hp_{play}}}) = \frac{1}{\mathbf{H}^{\mathrm{ah,h}}_{\mathrm{hptf_{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}})} \frac{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{hptf}}})}{\mathbf{H}^{\mathrm{ah,h}}_{\mathbf{u}_{\mathrm{hp}},\mathbf{p}_{\mathrm{ahm}}}(\mathbf{x}_{\mathrm{hp_{play}}})} \tag{79}$$

This situation represents the only case that can be equalized completely (in a straightforward manner).

7. RESULTING DESIGN CRITERIA

Based on the system theoretical framework introduced in the preceding chapters, recommendations for the design of binaural synthesis systems can be given. Since this work deals with the theoretical background, not with implementation details, it is based on the best case situation. That means all prerequisites necessary are mentioned but their feasibility is not proven. It is possible that certain implementation factors, for example the reproducibility of the headphone position, influence the analysis given below. These problems have to be accounted for in addition to the best case valuation given here. However, this best case scenario has to be a baseline for further considerations since this situation represents the theoretical basis and is therefore inevitably involved in every realization of binaural synthesis.

7.1. Human Head Recording

The preferable approach to binaural synthesis is using miniature microphones at the entrances to the blocked auditory canals of the subject for whom the synthesis is to be done. If the HPIRs are measured at the recording microphone positions, the headphone position is identical for playback and HPIR measurement, and appropriate headphones are employed, the reference scene ear signals can be recreated completely (cf. 62). Appropriate headphones in this context are headphones that do not alter the transfer functions of the sound pressure between the blocked auditory canals and the eardrums with open auditory canals compared to the reference situation (cf. 6.3.2). Mathematically this requires frequency independence of the transfer functions

$$\mathbf{H}^{\mathrm{ind}}_{\mathrm{hpsc}}(\mathbf{x}_{\mathrm{m_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{m_{hptf}}}) = = \frac{\mathbf{H}^{\mathrm{ind}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{ls}}(\mathbf{x}_{\mathrm{m_{rec}}})}{\mathbf{H}^{\mathrm{ind,h}}_{\mathbf{p}_{\mathrm{m_b}},\mathbf{p}_{\mathrm{e}},\mathrm{hp}}(\mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{m_{hptf}}})}. \tag{80}$$

Figure 1 shows the magnitude of $\mathbf{H}_{\mathrm{hpsc}}^{\mathrm{ind}}(\mathbf{x}_{\mathrm{m_{rec}}}, \mathbf{x}_{\mathrm{hp_{play}}}, \mathbf{x}_{\mathrm{hp_{hptf}}}, \mathbf{x}_{\mathrm{m_{hptf}}})$ for two different headphones commonly used in binaural synthesis (black - Sennheiser HD 800, gray - STAX SR lambda pro NEW), measured with one specimen of each model on a modified artificial head (Neumann KU 80). The specimen indicated

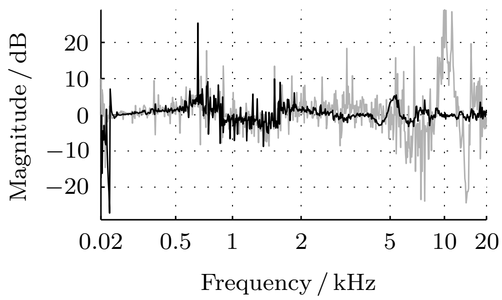

Fig. 1: Influence of headphones on the magnitude of the sound pressure transfer function between the blocked auditory canals and the eardrums with open auditory canals for two headphone models.

in gray will influence the binaurally synthesized ear signals remarkably whereas the model represented in black will have less undesired impact. For a detailed discussion see Völk (2010).

7.2. Artificial Head Recording

It might be desirable to use artificial head instead of individual recordings for binaural synthesis. The synthesized ear signals will then deviate from the intended due to the physical deviation between artificial and human heads. If probe microphone measured HPIRs are used to equalize the synthesis system (cf. 65) and signals measured by probe microphones are assumed to represent the ear signals completely, the deviation between artificial and human head recordings is the only cause of error in the resulting ear signals. Probe microphone measurements right in front of the eardrums are rather time consuming and require specific hardware. If this complexity is not justified or possible, one could use artificial head HPIRs (cf. eq. 67). In that case, the deviation from artificial to human head HPIRs cause additional error in the resulting ear signals.

8. CONCLUSIONS

A detailed system theoretical analysis of binaural synthesis systems is given. It is shown that system theoretically completely correct synthesis is possible using blocked auditory canal miniature microphone recordings. An inevitable requirement for authentic reproduction of the reference scene ear signals are appropriate headphones. Headphones suitable for individual binaural synthesis do not influence the sound pressure transfer functions between the blocked auditory canals and the eardrums with open canals compared to the reference situation without headphones. The selection of proper headphones is important since models typically used for binaural synthesis can show considerable influence on the aforementioned transfer paths. If individual measurements are not possible, artificial head measurements can be used for binaural synthesis. In that case, the deviation from the human to the artificial head defines the resulting error.

9. ACKNOWLEDGMENTS

The author is indebted to Prof. Dr.-Ing. H. Fastl for providing valuable insight on the subject and the possibility for this work, supported by grant FA 140/4 of Deutsche Forschungsgemeinschaft (DFG).

10. REFERENCES

Begault D. R.: Virtual Acoustic Displays for Teleconferencing: Intelligibility Advantage for “Telephone-Grade” Audio. J. Audio Eng. Soc. **47**, 824–828 (1999)

Bixler O.: A practical binaural recording system. Transactions of the IRE Professional Group on Audio **1**, 14–22 (1953)

Blauert J.: *Spatial Hearing – The Psychophysics of Human Sound Localization.* Revised Edition (The MIT Press, Cambridge, Massachusetts, London, 1997)

Blauert J., H. Lehnert, J. Sahrhage, H. Strauss: An Interactive Virtual-Environment Generator for Psychoacoustic Research. I: Architecture and Implementation. Acustica – Acta Acustica **86**, 94–102 (2000)

Djelani T., C. Pörschmann, J. Sahrhage, J. Blauert: An Interactive Virtual-Environment Generator for Psychoacoustic Research II: Collection of Head-Related Impulse Responses and Evaluation of Auditory Localization. Acustica – Acta Acustica **86**, 1046–1053 (2000)

Fastl H.: Gibt es *den* Frequenzgang von Kopfhörern? (Does *the* frequency response of headphones exist?). In *NTG-Fachberichte 91, Hörrundfunk 7*, 274–281 (VDE-Verlag, Berlin, 1986)

Fastl H., E. Zwicker: *Psychoacoustics – Facts and Models.* 3rd Ed. (Springer, Berlin, 2007)

Fastl H., W. Schmid, G. Theile, E. Zwicker: Schallpegel im Gehörgang für gleichlaute Schalle aus Kopfhör. od. Lautsprechern (Sound level in the ear canal for equally loud sounds from headp. vs loudsp.). In *Fortschritte der Akustik, DAGA '85*, 471–474 (DPG, Bad Honnef, 1985)

Gröhn M., T. Lokki, T. Takala: Localizing Sound Sources in a CAVE-Like Virtual Environment with Loudspeaker Array Reproduction. Presence **16**, 157–171 (2007)

Hammershøi D., H. Møller: Sound transmission to and within the human ear canal. J. Acoust. Soc. Am. **100**, 408–427 (1996)

Hammershøi D., H. Møller: Methods for Binaural Recording and Reproduction. Acta Acustica united with Acustica **88**, 303–311 (2002)

Jeffress L. A., R. W. Taylor: Lateralization vs Localization. J. Acoust. Soc. Am. **33**, 482–483 (1961)

Mackensen P., U. Felderhoff, G. Theile, U. Horbach, R. Pellegrini: Binaural Room Scanning – A new Tool for Acoustic and Psychoacoustic Research. In *137th Meeting of the Acoustical Society of America* (ASA, Melville, NY, 1999)

Møller H.: Fundamentals of Binaural Technology. Applied Acoustics **36**, 171–218 (1992)

Munson W. A., F. M. Wiener: In Search of the Missing 6 Db. J. Acoust. Soc. Am. **24**, 498–501 (1952)

Neely S. T., J. B. Allen: Invertibility of a room impulse response. J. Acoust. Soc. Am. **66**, 165–169 (1979)

Oppenheim A. V., A. S. Willsky, H. Nawab: *Signals and Systems.* 2nd Edition (Prentice-Hall, New Jersey, 1998)

Plenge G.: On the differences between localization and lateralization. J. Acoust. Soc. Am. **56**, 944–951 (1974)

Rudmose W.: The case of the missing 6 dB. J. Acoust. Soc. Am. **71**, 650–659 (1982)

Sivian L. J., S. D. White: On Minimum Audible Sound Fields. J. Acoust. Soc. Am. **4**, 288–321 (1933)

Theile G.: *Über die Lokalisation im überlagerten Schallfeld (On the localization in the superimposed soundfield)*, PhD thesis, Technische Universität Berlin (1980)

Völk F.: Messtechnische Verifizierung eines datenbasierten binauralen Synthesesystems (Metrological verification of data based binaural synthesis). In *Fortschritte der Akustik, DAGA 2010*, 1049–1050 (Dt. Gesell. für Akustik e. V., Berlin, 2010)

Völk F., T. Musialik, H. Fastl: Crosstalk Cancellation between Phantom Sources. In *126th AES Convention* (2009) (Convention Paper 7722)

Völk F., S. Kerber, H. Fastl, S. Reifinger: Design und Realisierung von virtueller Akustik für ein Augmented-Reality-Labor (Design and implementation of virtual acoustics for an augmented reality lab). In *Fortschritte der Akustik, DAGA 2007*, 673–674 (Dt. Gesell. für Akustik e. V., Berlin, 2007)

Vorländer M.: Acoustic load on the ear caused by headphones. J. Acoust. Soc. Am. **107**, 2082–2088 (2000)

Wightman F. L., D. J. Kistler: Headphone simulation of free-field listening. I: Stimulus synthesis. J. Acoust. Soc. Am. **85**, 858–867 (1989)

Wilkens H.: Kopfbezügliche Stereophonie – ein Hilfsmittel für Vergleich und Beurteilung verschiedener Raumeindrücke (Head-Related Stereophony – An Aid for the Comparison and Critical Examination of Different Room Effects). Acustica **26**, 213–221 (1972)

Zahorik P.: Assessing auditory distance perception using virtual acoustics. J. Acoust. Soc. Am. **111**, 1832–1846 (2002)

Zollner M., E. Zwicker: *Elektroakustik (Electroacoustics).* 3rd Edition (Springer, Berlin, 1993)