Proceedings of 20th International Congress on Acoustics, ICA 2010

23–27 August 2010, Sydney, Australia

Advantages of binaural room synthesis for research and fitting of hearing aids, cochlear implants, electro-acoustical stimulation, and combined systems

Florian Völk and Hugo Fastl

AG Technische Akustik, Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München, Germany

florian.voelk@mytum.de

PACS: 43.66.Ts, 43.55.Ka

ABSTRACT

Binaural room synthesis is a sound reproduction technology that is, for a normal hearing listener, based on convolution of the sound signals to be reproduced with the impulse responses of the sound pressure propagation paths from the sources to be simulated to the listener's eardrums. The convolution products are then presented by adequately equalized, appropriate headphones. The simulation of arbitrary acoustical reproduction environments is possible as long as they can be regarded as linear systems, which is the case for almost all practical scenarios. If the signal processing is implemented with real-time capability and adaptively, the impulse responses can be adjusted based on the listener's head position and orientation, gathered by a tracking system. This allows the listener to move freely while the simulated acoustical scenario, encoded by the sound pressure signals at the eardrums, remains correct. For research and fitting purposes in the field of hearing aids or cochlear implants, binaural synthesis could save effort and time by providing arbitrary acoustical environments in the laboratory. Since conventional binaural synthesis is based on the reproduction of an original scene's sound pressure signals in front of the listener's eardrums, it is not directly applicable to listeners using hearing aids or cochlear implants, especially when the sound pressure at the eardrums is not present or not involved in the hearing process. If cochlear implants are used in connection with hearing aids or the remaining normal hearing system, as for example with electro-acoustical stimulation, the situation becomes even more complicated. Within this contribution, the conventional theory of binaural room synthesis is adapted to the typical applications with hearing impaired listeners. Further, possible application scenarios and advantages compared to traditional methods for research and fitting in the context of hearing aids and cochlear implants are discussed.

INTRODUCTION

The underlying principle and basic procedure of binaural technology is comprehensively defined by Møller (1992) 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 explanation shows very demonstratively the expansive capabilities of the described procedure. The direct binaural recording of the sound pressure signals at the eardrums (the ear signals) is rather time consuming, especially if more than one listener, different head rotations, or listening positions are to be considered. Nevertheless, due to its simplicity, this procedure has been in use for many years (e. g. Bixler 1953, Wilkens 1972, Hammershøi and Møller 2002). A less time consuming method to acquire the necessary ear signals is binaural room synthesis (BRS, cf. Møller 1992). Here, the ear signals are generated by convolution of the source input signals with the impulse responses describing the acoustical propagation paths between the sound source and the listener's eardrums in the corresponding real life situation (cf. Völk et al. 2009b). These impulse responses are dependent on the loudspeaker and body position and orientation as well as on the geometrical setup of the listening situation under consideration. For that reason, an additional adaptive signal processing is required that selects the impulse responses corresponding to the current listening situation at any instant of time based on the body and loudspeaker positions (cf. Mackensen et al. 1999). If more than one source is to be simulated, the ear signals can be generated by linear superposition of the ear signals produced by each source.

A wide variety of binaural room synthesis systems for normal hearing listeners based on the described procedure is available. 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 BRS 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 (see Völk 2010). This way, the auditory impressions of the real life situation are reproduced quite realistically (e. g. Völk et al. 2008a, Völk 2009), even if complex acoustical scenarios are to be simulated (cf. Völk et al. 2008b, 2010).

The major aim of this paper is to identify the restrictions and corrections necessary for applying the BRS procedure to hearing impaired listeners using typical hearing aids, based on a detailed system theoretical analysis of all involved components. In addition, hardware selection criteria are discussed for every considered hearing aid fitting.

SYSTEM THEORETICAL BASICS

In the following, lower case letters denote impulse responses or time dependent signals, for example the voltage $u(t)$ is abbreviated as u without explicit notation of the time dependence. Upper case letters represent complex Fourier-spectra or transfer functions, for example $U(f) = \mathscr{F}\{u(t)\}$ is denoted as U.

Lower case subscripts differentiate between specific signals and systems, additional upper case indexes denote the left (L) or right (R) side, if necessary. A major part of the calculations 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 set in bold fonts, e. g.:

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

The division or multiplication of two vectors or matrices is used as shorthand for element wise division or multiplication throughout this paper, 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. For the paper on hand, it is necessary at some points to denote explicitly if a signal is represented digitally or analogous. Digital sequences are represented by the variable s in the time domain, the corresponding Discrete Fourier Transform (DFT) by S. If not denoted explicitly, it is not necessary here to distinguish if a signal or system is represented or works digitally or analogous.

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 could be for example 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 can be defined between them symbolically, too. In any case, an appropriate constant has to be introduced to convert between the different units. This is taken for granted here. 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 as well as the invertibility of the mentioned impulse responses. Problems concerning these issues in practical implementations are not within the scope of this paper although they could become a crucial factor.

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 and Wilskey 1983). If the situation is restricted to constant temperature and sound sources at distances more than 0.5 m and less than 15 m from the listener, this holds true at least within short temporal samples to a sufficient degree for all considered subsystems. For larger distances, the damping of air, for smaller distances shadowing of the head cause the system to be nonlinear (according to Blauert 1997). For that reason, if the consideration is restricted to sources at distances between 0.5 and 15 m, all involved systems can be regarded as linear and at least piecewise time invariant (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 or by an array of impulse responses $h(\tau_i)$ with $i \in [0,N]$.

BINAURAL ROOM SYNTHESIS WITH NORMAL HEARING LISTENERS

The aim of BRS is the 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 (index *ref*) here. The position and orientation of the subject's head in an arbitrary coordinate system is given by the vector $\vec{x}_\mathrm{h}$:

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

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 parameter δ takes into account 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 is defined by:

$$\vec{x}_\mathrm{ls} = (x, y, z, r_x, r_y, r_z)^T$$

Further positions that are indicated in the same coordinate system are the positions of pairs of microphones

$$\vec{\mathbf{x}}_\mathrm{mic} = (x_L, x_R, y_L, y_R, z_L, z_R, r_{x_L}, r_{x_R}, r_{y_L}, r_{y_R}, r_{z_L}, r_{z_R})^T,$$

and the hearing aids' positions on a listener's head $\vec{\mathbf{x}}_\mathrm{ha}$, as well as possibly present output loudspeakers' positions in the auditory canals, the latter represented by the additional position vectors $\vec{\mathbf{x}}_\mathrm{hals}$:

$$\vec{\mathbf{x}}_\mathrm{ha} = (x_L, x_R, y_L, y_R, z_L, z_R, r_{x_L}, r_{x_R}, r_{y_L}, r_{y_R}, r_{z_L}, r_{z_R})^T$$
$$\vec{\mathbf{x}}_\mathrm{hals} = (x_L, x_R, y_L, y_R, z_L, z_R, r_{x_L}, r_{x_R}, r_{y_L}, r_{y_R}, r_{z_L}, r_{z_R})^T$$

The restriction to a loudspeaker box does not result in loss of generality, since the voltage u_ls can be regarded in any case as input to any ideal electro-acoustic transducer (e. g. an acoustical monopole or dipole point source, cf. Zollner and Zwicker 1993). An additional advantage resulting from the linearity of all involved subsystems 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.

Important signals involved are the loudspeaker input voltage u_ls as well as the digital sample sequence s_ls, encoding the sound signal to be played back by the loudspeaker. Further, the time varying sound pressure signals at the eardrums, called ear signals $\mathbf{p}_\mathrm{e}^\mathrm{ind}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls})$.

Each ear signal occurring in the reference scene includes influences of the transfer characteristics of systems which could be grouped best by their affiliation to the sound generation or propagation parts of the sound transfer paths from the source to the eardrums. These contributing partial paths will be considered in the following.

Loudspeaker playback system

The generation part of all paths consists of the audio interface used for D/A-conversion from the source sequence s_ls to the output voltage u_da, represented by its transfer function:

$$H_\mathrm{da} = \frac{U_\mathrm{da}}{S_\mathrm{ls}} \tag{1}$$

The output of the audio interface drives the loudspeaker amplifier, which in turn provides the output voltage u_ls. This device is described by the transfer function

$$H_\mathrm{als} = \frac{U_\mathrm{ls}}{U_\mathrm{da}}. \tag{2}$$

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 overall transfer function of the output system (the generation part of the transfer paths) is then given as follows:

$$H_\mathrm{o} = \frac{U_\mathrm{ls}}{S_\mathrm{ls}} = H_\mathrm{da} \cdot H_\mathrm{als} \tag{3}$$

Microphone transfer functions

For the study on hand, probe microphone (*pm*) recordings in front of the listener's eardrums as well as recordings of hearing aid microphones (*ham*) are of interest. If necessary, the specific system is indicated by the respective index, if no explicit distinction is necessary, the index *mic* (microphone) is used for convenience.

Microphone transfer functions describe the conversion of the sound pressure signals $\mathbf{p}^\mathrm{ind}_\mathrm{mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})$ to the voltage signals $\mathbf{u}^\mathrm{ind}_\mathrm{mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})$, and are here defined as follows:

$$\mathbf{H}_\mathrm{m} = \frac{\mathbf{U}^\mathrm{ind}_\mathrm{mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})}{\mathbf{P}^\mathrm{ind}_\mathrm{mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})} \tag{4}$$

Input systems

An input system described by its transfer functions $\mathbf{H}_\mathrm{i}$ is used to transform the microphone output voltages $\mathbf{u}^\mathrm{ind}_\mathrm{mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})$ to digital sample sequences $\mathbf{s}^\mathrm{ind}_\mathrm{mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})$. The input system's characteristics are a combination of the microphone preamplifier transfer functions

$$\mathbf{H}_\mathrm{am} = \frac{\mathbf{U}^\mathrm{ind}_\mathrm{ad}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})}{\mathbf{U}^\mathrm{ind}_\mathrm{mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})} \tag{5}$$

that amplify the input voltages $\mathbf{u}^\mathrm{ind}_\mathrm{mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})$ to the output voltages $\mathbf{u}^\mathrm{ind}_\mathrm{ad}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})$ and the A/D-converters, producing the digital output sequences $\mathbf{s}^\mathrm{ind}_\mathrm{mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})$:

$$\mathbf{H}_\mathrm{ad} = \frac{\mathbf{S}^\mathrm{ind}_\mathrm{mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})}{\mathbf{U}^\mathrm{ind}_\mathrm{ad}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})} \tag{6}$$

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

$$\mathbf{H}_\mathrm{i} = \frac{\mathbf{S}^\mathrm{ind}_\mathrm{mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})}{\mathbf{U}^\mathrm{ind}_\mathrm{mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})} = \mathbf{H}_\mathrm{am} \cdot \mathbf{H}_\mathrm{ad} \tag{7}$$

Propagation paths for normal hearing listeners

For a normal hearing listener, the propagation parts of the sound paths are usually defined between the voltage u_ls at the loudspeaker input terminals and the ear signals $\mathbf{p}^\mathrm{ind}_\mathrm{e}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls})$. These paths are described by the following transfer functions:

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

These transfer functions contain the transfer characteristics of the loudspeaker 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, they are highly dependent on the position and orientation of the listener's head and body and therefore represented as array of transfer functions with the subject and loudspeaker position vectors as array parameters.

Reference scene for normal hearing listeners

The following equation is given to describe the ear signal spectra in the considered reference situation (index *ref*):

$$\mathbf{P}^\mathrm{ind}_\mathrm{e}(\vec{x}_{\mathrm{h}_\mathrm{ref}}, \vec{x}_{\mathrm{ls}_\mathrm{ref}}) = S_\mathrm{ls} \cdot H_{\mathrm{o}_\mathrm{ref}} \cdot \mathbf{H}^\mathrm{ind}_\mathrm{ls,e}(\vec{x}_{\mathrm{h}_\mathrm{ref}}, \vec{x}_{\mathrm{ls}_\mathrm{ref}}) \tag{9}$$

The transfer functions describing the connection between the sequence driving the loudspeaker and the ear signals in the reference situation are defined as follows:

$$\begin{aligned} \mathbf{H}^\mathrm{ind}_\mathrm{ref}(\vec{x}_{\mathrm{h}_\mathrm{ref}}, \vec{x}_{\mathrm{ls}_\mathrm{ref}}) &= \frac{\mathbf{P}^\mathrm{ind}_\mathrm{e}(\vec{x}_{\mathrm{h}_\mathrm{ref}}, \vec{x}_{\mathrm{ls}_\mathrm{ref}})}{S_\mathrm{ls}} \\ &= H_{\mathrm{o}_\mathrm{ref}} \cdot \mathbf{H}^\mathrm{ind}_\mathrm{ls,e}(\vec{x}_{\mathrm{h}_\mathrm{ref}}, \vec{x}_{\mathrm{ls}_\mathrm{ref}}) \end{aligned} \tag{10}$$

These transfer functions are valid as long as the overall situation does not change. This is important in binaural synthesis, since usually headphones are used for playback, which can cause a change in radiation impedance compared to the reference situation (Völk 2010).

Recording scene for normal hearing listeners

For recording (*rec*) of the propagation paths' impulse responses necessary for binaural synthesis with the defined reference, the loudspeaker of the reference situation has to be used. For normal hearing listeners, the recording can be done with probe microphones in the auditory canals. This technique may be used in a hearing impaired listener's auditory canal too, when usage of the hearing aid's microphones is not sufficient or feasible.

In every case, the recordings contain transfer characteristics of three different systems: First the loudspeaker output system H_o (cf. section *Loudspeaker playback system*), second the input system $\mathbf{H}_\mathrm{i}$ (cf. section *Input system*), and third the parts of the sound propagation paths captured by the selected recording setup. These parts are here defined between the loudspeaker's input voltage u_ls and the microphones' output voltages $\mathbf{u}^\mathrm{ind}_\mathrm{mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})$ and described by their transfer functions:

$$\mathbf{H}^\mathrm{ind}_\mathrm{ls,mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic}) = \frac{\mathbf{U}^\mathrm{ind}_\mathrm{mic}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})}{U_\mathrm{ls}} \tag{11}$$

The transfer functions $\mathbf{H}^\mathrm{ind}_{\mathrm{rec}_\mathrm{mic}}(\vec{x}_\mathrm{h}, \vec{x}_\mathrm{ls}, \vec{\mathbf{x}}_\mathrm{mic})$ describe the connection between the digital sample sequence s_ls to be played back by the loudspeaker and the input sequences stemming from the microphones $\mathbf{s}^\mathrm{ind}_\mathrm{mic}(\vec{x}_{\mathrm{h}_\mathrm{rec}}, \vec{x}_{\mathrm{ls}_\mathrm{rec}}, \vec{\mathbf{x}}_{\mathrm{mic}_\mathrm{rec}})$ in the recording situation:

$$\begin{aligned} \mathbf{H}^\mathrm{ind}_{\mathrm{rec}_\mathrm{mic}}(\vec{x}_{\mathrm{h}_\mathrm{rec}}, \vec{x}_{\mathrm{ls}_\mathrm{rec}}, \vec{\mathbf{x}}_{\mathrm{mic}_\mathrm{rec}}) &= \frac{\mathbf{S}^\mathrm{ind}_\mathrm{mic}(\vec{x}_{\mathrm{h}_\mathrm{rec}}, \vec{x}_{\mathrm{ls}_\mathrm{rec}}, \vec{\mathbf{x}}_{\mathrm{mic}_\mathrm{rec}})}{S_\mathrm{ls}} \\ &= H_{\mathrm{o}_\mathrm{rec}} \cdot \mathbf{H}^\mathrm{ind}_\mathrm{ls,mic}(\vec{x}_{\mathrm{h}_\mathrm{rec}}, \vec{x}_{\mathrm{ls}_\mathrm{rec}}, \vec{\mathbf{x}}_{\mathrm{mic}_\mathrm{rec}}) \cdot \mathbf{H}_{\mathrm{i}_\mathrm{rec}} \end{aligned} \tag{12}$$

Playback scene for normal hearing listeners

In the standard BRS situation with normal hearing subjects, which is called normal synthesis (BRS) in the following, the signals encoded in $\mathbf{s}_\mathrm{hp}$ are presented by headphones (*hp*), which are located on the subject's head at the positions $\vec{\mathbf{x}}_\mathrm{hp}$ and driven with the voltages $\mathbf{u}_\mathrm{hp}$ (at their input terminals).

Thereby, the transfer characteristics of the headphone playback paths $\mathbf{H}^\mathrm{ind,h}_\mathrm{play}(\vec{\mathbf{x}}_{\mathrm{hp}_\mathrm{play}})$ 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. 2009a), but is not taken into account here.

When using headphones for playback, the path descriptions consist of the transfer characteristics of the headphone output equipment, described by

$$\mathbf{H}_{\mathrm{o_{hp}}} = \frac{\mathbf{U}_{\mathrm{hp}}}{\mathbf{S}_{\mathrm{hp}}}, \tag{13}$$

as well as of the sound propagation paths between the headphone input voltages and the sound pressure signals at the eardrums (under the headphones, indicated by the additional superscript *h*):

$$\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hp,e}}(\vec{\mathbf{x}}_{\mathrm{hp}}) = \frac{\mathbf{P}^{\mathrm{ind,h}}_{\mathrm{e}}(\vec{\mathbf{x}}_{\mathrm{hp}})}{\mathbf{U}_{\mathrm{hp}}} \tag{14}$$

The system called headphone output equipment consists of the audio device output channels used for conversion of the digital sample sequences $\mathbf{S}_{\mathrm{hp}}$ to the voltages $\mathbf{U}_{\mathrm{da}}$, described by their transfer functions

$$\mathbf{H}_{\mathrm{da}} = \frac{\mathbf{U}_{\mathrm{da}}}{\mathbf{S}_{\mathrm{hp}}} \tag{15}$$

as well as of the headphone amplifiers with the input voltages $\mathbf{U}_{\mathrm{da}}$, stemming from the audio device:

$$\mathbf{H}_{\mathrm{ahp}} = \frac{\mathbf{U}_{\mathrm{hp}}}{\mathbf{U}_{\mathrm{da}}} \tag{16}$$

Therefore, the transfer functions of the headphone output equipment are given by the following equation:

$$\mathbf{H}_{\mathrm{o_{hp}}} = \mathbf{H}_{\mathrm{da}} \cdot \mathbf{H}_{\mathrm{ahp}} \tag{17}$$

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

$$\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{play}}(\vec{\mathbf{x}}_{\mathrm{hp_{play}}}) = \frac{\mathbf{P}^{\mathrm{h}}_{\mathrm{e}}(\vec{\mathbf{x}}_{\mathrm{hp_{play}}})}{\mathbf{S}_{\mathrm{hp}}} = \mathbf{H}_{\mathrm{o_{hp}}} \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hp,e}}(\vec{\mathbf{x}}_{\mathrm{hp_{play}}}) \tag{18}$$

Headphone impulse responses

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

$$\begin{aligned}\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hp}}(\vec{\mathbf{x}}_{\mathrm{hp}}, \vec{\mathbf{x}}_{\mathrm{mic}}) &= \frac{\mathbf{S}^{\mathrm{ind,h}}_{\mathrm{mic}}(\vec{\mathbf{x}}_{\mathrm{hp}}, \vec{\mathbf{x}}_{\mathrm{mic}})}{\mathbf{S}_{\mathrm{hp}}} \\ &= \mathbf{H}_{\mathrm{o_{hp}}} \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hp,mic}}(\vec{\mathbf{x}}_{\mathrm{hp}}, \vec{\mathbf{x}}_{\mathrm{mic}}) \cdot \mathbf{H}_{\mathrm{mic}} \cdot \mathbf{H}_{\mathrm{i}}\end{aligned} \tag{19}$$

Their time domain representations are usually referred to as headphone impulse responses (HPIRs).

Non-equalized binaural room synthesis situation for normal hearing listeners

Convolving an audio signal with the impulse responses recorded according to section *Recording scene for normal hearing listeners* and presenting the convolution products to a listener by headphones as described in section *Playback scene for normal hearing listeners* is here referred to as the (temporary) unequalized BRS scene (*tmp*). Using equations 12 and 18, the ear signals occurring in this unequalized BRS situation can be computed. Assuming $\vec{x}_{\mathrm{h_{rec}}} = \vec{x}_{\mathrm{h_{ref}}}$ and $\vec{x}_{\mathrm{ls_{rec}}} = \vec{x}_{\mathrm{ls_{ref}}}$, the following ear signal spectra occur:

$$\begin{aligned}&\mathbf{P}^{\mathrm{ind,h}}_{\mathrm{e_{tmp}}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}}, \vec{\mathbf{x}}_{\mathrm{mic_{rec}}}, \vec{\mathbf{x}}_{\mathrm{hp_{play}}}) = \\ &\quad S_{\mathrm{ls}} \cdot \mathbf{H}^{\mathrm{ind}}_{\mathrm{rec_{mic}}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}}, \vec{\mathbf{x}}_{\mathrm{mic_{rec}}}) \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathrm{play}}(\vec{\mathbf{x}}_{\mathrm{hp_{play}}})\end{aligned} \tag{20}$$

The transfer functions

$$\begin{aligned}&\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{tmp}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}}, \vec{\mathbf{x}}_{\mathrm{mic_{rec}}}, \vec{\mathbf{x}}_{\mathrm{hp_{play}}}) = \\ &\mathbf{H}^{\mathrm{ind}}_{\mathrm{rec_{mic}}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}}, \vec{\mathbf{x}}_{\mathrm{mic_{rec}}}) \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathrm{play}}(\vec{\mathbf{x}}_{\mathrm{hp_{play}}}) = \\ &H_{\mathrm{o_{rec}}} \cdot \mathbf{H}^{\mathrm{ind}}_{\mathrm{ls,mic}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}}, \vec{\mathbf{x}}_{\mathrm{mic_{rec}}}) \cdot \mathbf{H}_{\mathrm{i_{rec}}} \cdot \\ &\quad \mathbf{H}_{\mathrm{o_{hp}}} \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hp,e}}(\vec{\mathbf{x}}_{\mathrm{hp_{play}}})\end{aligned} \tag{21}$$

describe the connections between the source signal and the ear signal spectra in the binaural room synthesis situation.

Equalization requirements for binaural room synthesis with normal hearing listeners

The ear signals in the BRS situation are intended to equal the ear signals of the reference scene:

$$\mathbf{p}^{\mathrm{ind,h}}_{\mathrm{e_{BRS}}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}}, \vec{\mathbf{x}}_{\mathrm{mic_{rec}}}, \vec{\mathbf{x}}_{\mathrm{hp_{play}}}) \stackrel{!}{=} \mathbf{p}^{\mathrm{ind}}_{\mathrm{e}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}}) \tag{22}$$

To reach this goal, equalization filters $\mathbf{h}^{\mathrm{ind}}_{\mathrm{eq}}(\vec{\mathbf{x}}_{\mathrm{hp_{play}}}, \vec{\mathbf{x}}_{\mathrm{mic_{rec}}})$ have to be applied so that following equation holds:

$$\begin{aligned}\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{tmp}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}}, \vec{\mathbf{x}}_{\mathrm{mic_{rec}}}, \vec{\mathbf{x}}_{\mathrm{hp_{play}}}) \cdot \mathbf{H}^{\mathrm{ind}}_{\mathrm{eq}}(\vec{\mathbf{x}}_{\mathrm{hp_{play}}}, \vec{\mathbf{x}}_{\mathrm{mic_{rec}}}) &\stackrel{!}{=} \\ \mathbf{H}^{\mathrm{ind}}_{\mathrm{ref}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}})&\end{aligned} \tag{23}$$

Assuming invertibility of $\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{tmp}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}}, \vec{\mathbf{x}}_{\mathrm{mic_{rec}}}, \vec{\mathbf{x}}_{\mathrm{hp_{play}}})$, the required equalization filters are given as follows:

$$\begin{aligned}\mathbf{H}^{\mathrm{ind}}_{\mathrm{eq}}(\vec{\mathbf{x}}_{\mathrm{hp_{play}}}, \vec{\mathbf{x}}_{\mathrm{mic_{rec}}}) &= \frac{\mathbf{H}^{\mathrm{ind}}_{\mathrm{ref}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}})}{\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{tmp}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}}, \vec{\mathbf{x}}_{\mathrm{mic_{rec}}}, \vec{\mathbf{x}}_{\mathrm{hp_{play}}})} \\ &= \frac{H_{\mathrm{o_{ref}}} \cdot \mathbf{H}^{\mathrm{ind}}_{\mathrm{ls,e}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}})}{H_{\mathrm{o_{rec}}} \cdot \mathbf{H}^{\mathrm{ind}}_{\mathrm{ls,mic}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}}, \vec{\mathbf{x}}_{\mathrm{mic_{rec}}})} \cdot \\ &\quad \frac{1}{\mathbf{H}_{\mathrm{i_{rec}}} \cdot \mathbf{H}_{\mathrm{o_{hp}}} \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hp,e}}(\vec{\mathbf{x}}_{\mathrm{hp_{play}}})}\end{aligned} \tag{24}$$

If probe microphones are used and the sound pressure signals at the microphones are assumed to approximate the ear signals, the following approximation is valid:

$$\mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}^{\mathrm{ind}}_{\mathrm{ls,e}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}}) \approx \mathbf{H}^{\mathrm{ind}}_{\mathrm{ls,pm}}(\vec{x}_{\mathrm{h_{ref}}}, \vec{x}_{\mathrm{ls_{ref}}}, \vec{\mathbf{x}}_{\mathrm{pm_{rec}}}). \tag{25}$$

Eventually, if the same output equipment as in the reference situation is used ($H_{\mathrm{o_{rec}}} = H_{\mathrm{o_{ref}}}$), equation 24 can be simplified (for probe microphones) as follows:

$$\mathbf{H}^{\mathrm{ind}}_{\mathrm{eq_{pm}}}(\vec{\mathbf{x}}_{\mathrm{hp_{play}}}, \vec{\mathbf{x}}_{\mathrm{pm_{rec}}}) \approx \frac{1}{\mathbf{H}_{\mathrm{pm}} \cdot \mathbf{H}_{\mathrm{i_{pm}}} \cdot \mathbf{H}_{\mathrm{o_{hp}}} \cdot \mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hp,e}}(\vec{\mathbf{x}}_{\mathrm{hp_{play}}})} \tag{26}$$

The relation of this equalization filters to the HPTFs is of interest, since a practical acquirement of equalization filters is commonly done by recording distinct HPIRs with the recording situation's output system $\mathbf{H}_{\mathrm{o_{hp}}}$ and input system $\mathbf{H}_{\mathrm{i_{pm}}}$, and inverting them. As a consequence, comparison to equation 19 reveals:

$$\mathbf{H}^{\mathrm{ind}}_{\mathrm{eq_{pm}}}(\vec{\mathbf{x}}_{\mathrm{hp_{play}}}, \vec{\mathbf{x}}_{\mathrm{pm_{rec}}}) \approx \frac{1}{\mathbf{H}^{\mathrm{ind,h}}_{\mathrm{hp_{pm}}}(\vec{\mathbf{x}}_{\mathrm{hp}}, \vec{\mathbf{x}}_{\mathrm{pm}})} \tag{27}$$

To sum up: If the sound pressure signals at the probe microphones right in front of the eardrums are assumed to represent the ear signals, and if the probe microphone and headphone positions are kept constant for headphone impulse response measurement, recording, and playback situation, it is possible to equalize a binaural room synthesis system with the inverses of the headphone impulse responses so that the synthesis equals the reference scene.

BINAURAL ROOM SYNTHESIS WITH HEARING IMPAIRED LISTENERS

For most hearing aid fittings, the time varying sound pressure signals at the eardrums, called ear signals $\mathbf{p}_{\mathrm{e}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}})$ for normal hearing listeners, are involved in the hearing process also for hearing impaired listeners. These time varying sound pressure signals at the eardrums for hearing impaired listeners using a hearing aid, regarded while wearing the hearing aid, are referred to as $\mathbf{p}_{\mathrm{e}}^{\mathrm{ind,ha}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}},\vec{\mathbf{x}}_{\mathrm{hals}})$ in the following. Possible contributions to those signals may either come from the hearing aid (*ha*) or directly stem from the sound field (*sf*) the listener is located in. The contributions are referred to as $\mathbf{p}_{\mathrm{e_{ha}}}^{\mathrm{ind,ha}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}},\vec{\mathbf{x}}_{\mathrm{hals}})$ and $\mathbf{p}_{\mathrm{e_{sf}}}^{\mathrm{ind,ha}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}},\vec{\mathbf{x}}_{\mathrm{hals}})$, respectively. The following equation describes the ear signals for hearing impaired listeners using a hearing aid:

$$\begin{aligned}\mathbf{p}_{\mathrm{e}}^{\mathrm{ind,ha}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}},\vec{\mathbf{x}}_{\mathrm{hals}}) &= \mathbf{p}_{\mathrm{e_{ha}}}^{\mathrm{ind,ha}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}},\vec{\mathbf{x}}_{\mathrm{hals}})\\ &\quad+\mathbf{p}_{\mathrm{e_{sf}}}^{\mathrm{ind,ha}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}},\vec{\mathbf{x}}_{\mathrm{hals}})\end{aligned} \tag{28}$$

Propagation paths to a hearing aid

It is possible to regard either the time varying sound pressure signals $\mathbf{p}_{\mathrm{ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}})$ at the hearing aid microphones (*ham*) or these microphones' output signals as the input signals to the hearing aid. Either one of these points of view opens a different approach to binaural synthesis and will therefore be discussed within this paper. The microphones' output signals can be represented either as voltage signals $\mathbf{u}_{\mathrm{ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}})$ or (after analog-digital conversion) as digital sample sequences $\mathbf{s}_{\mathrm{ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}})$. Within the paper on hand, the latter approach is discussed, since it is the procedure commonly used. A comparable discussion can be done for analog hearing aids, if necessary for a specific application.

With the definitions given within this section, the discussion of this paper holds true for hearing aids with one or more microphones. For simplicity, the plural form is used here. If more than one microphone is used per side, their input signals are summed up per side in vectors and combined in $\mathbf{p}_{\mathrm{ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}})$, which then becomes a matrix. With this definition, the transfer functions of the hearing aid microphones can be defined:

$$\mathbf{H}_{\mathrm{ham}} = \frac{\mathbf{U}_{\mathrm{ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}})}{\mathbf{P}_{\mathrm{ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}})} \tag{29}$$

The transfer functions describing the propagation paths from the loudspeaker input voltage to the sound pressure signals at the hearing aid microphones are given as follows:

$$\begin{aligned}\mathbf{H}_{\mathrm{ls,ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}}) &= \frac{\mathbf{P}_{\mathrm{ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}})}{U_{\mathrm{ls}}}\\ &= \frac{\mathbf{U}_{\mathrm{ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}})}{\mathbf{H}_{\mathrm{ham}}\cdot U_{\mathrm{ls}}}\end{aligned} \tag{30}$$

The following equation describes the transfer functions from the loudspeaker input voltage to the voltages at the hearing aid microphones' outputs:

$$\begin{aligned}\mathbf{H}_{\mathrm{ls,ha}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}}) &= \frac{\mathbf{U}_{\mathrm{ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}})}{U_{\mathrm{ls}}}\\ &= \mathbf{H}_{\mathrm{ls,ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}})\cdot\mathbf{H}_{\mathrm{ham}}\end{aligned} \tag{31}$$

Hearing aid transfer functions for acoustic output

Since it is not the aim of this paper to discuss certain properties of hearing aids, the hearing aids are modeled as linear systems without considering their specific properties, and are therefore described by their transfer functions. The inputs of conventional hearing aids are typically signals stemming from one or more microphones. Here, the sequences $\mathbf{S}_{\mathrm{ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}})$ after analog/digital conversion are regarded as hearing aid input signals $\mathbf{s}_{\mathrm{ha}}$. Being $\vec{\mathbf{x}}_{\mathrm{hals}}$ the hearing aid loudspeaker positions, the hearing aid's acoustical output signals are defined as the ear signals $\mathbf{P}_{\mathrm{e_{ha}}}^{\mathrm{ind,ha}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}},\vec{\mathbf{x}}_{\mathrm{hals}})$ (under the hearing aids). With these definitions, the hearing aid transfer functions for acoustic stimulation (AS) are given as follows:

$$\mathbf{H}_{\mathrm{ha,AS}}^{\mathrm{ind}}(\vec{\mathbf{x}}_{\mathrm{hals}}) = \frac{\mathbf{P}_{\mathrm{e_{ha}}}^{\mathrm{ind,ha}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}},\vec{\mathbf{x}}_{\mathrm{hals}})}{\mathbf{S}_{\mathrm{ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}})} \tag{32}$$

Categorization of hearing aids

The applications of binaural synthesis with hearing impaired listeners can be divided in two basically different groups, based on the fact whether the subject is able to use the contributions to the ear signals $\mathbf{p}_{\mathrm{e_{sf}}}^{\mathrm{ind,ha}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}},\vec{\mathbf{x}}_{\mathrm{hals}})$ not being processed by the hearing aid for auditory perception or not. If the contribution $\mathbf{p}_{\mathrm{e_{sf}}}^{\mathrm{ind,ha}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}},\vec{\mathbf{x}}_{\mathrm{hals}})$ at the considered ear is involved in the hearing process, it is necessary to synthesize $\mathbf{p}_{\mathrm{e_{sf}}}^{\mathrm{ind,ha}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}},\vec{\mathbf{x}}_{\mathrm{hals}})$ and the input signal or signals to the hearing aid (combined synthesis, CBRS). In case the subject is no longer able to use the ear signals for the hearing process or if the hearing aid occludes the auditory canal and provides adequate sound insulation, it is sufficient to synthesize the hearing aid input (aid only synthesis, AOBRS). With these definitions, in ear as well as behind the ear hearing aids are covered by the given theory.

In the remainder of this paper, binaural room synthesis for both of the situations introduced is discussed in detail. Finally, possible applications are given.

AID ONLY SYNTHESIS

Typical aid only synthesis cases are in ear hearing aids which occlude the auditory canal or traditional cochlear implant systems (where the listener's auditory system is not able to process the sound pressure at the eardrum). In the AOBRS case, either the microphone output signals or the sound pressure at the microphones can be regarded as the input to the hearing aid. Both situations are considered in the following.

If the hearing aid offers the possibility to include external signals (analog or digital), this way is more simple to use. In case the hearing is not capable of processing external input signals, the sound pressure at the microphones has to be simulated. Both situations are discussed in this section and therefore denoted according to following definitions:

1. *Electronic playback:* The synthesis output signals are delivered to the hearing aid inputs electronically.
2. *Acoustic playback:* Headphones are used to play back the synthesis results.

Reference scene for aid only synthesis

In the aid only synthesis case, it is sufficient to define the transfer functions of the reference situation to the hearing aid inputs, since they represent the only inputs to the hearing system. Combining equations 3, 7, and 31, the following transfer functions can be derived, which describe for both playback methods the reference situation in the aid only synthesis case:

$$\begin{aligned}\mathbf{H}_{\mathrm{ref_{AOBRS}}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}}) &= \frac{\mathbf{S}_{\mathrm{ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}})}{S_{\mathrm{ls}}}\\ &= H_{\mathrm{o_{ref}}}\cdot\mathbf{H}_{\mathrm{ls,ha}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}})\cdot\mathbf{H}_{\mathrm{i_{ha}}}\\ &= H_{\mathrm{o_{ref}}}\cdot\mathbf{H}_{\mathrm{ls,ham}}^{\mathrm{ind}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}})\cdot\mathbf{H}_{\mathrm{ham}}\cdot\mathbf{H}_{\mathrm{i_{ha}}}\end{aligned} \tag{33}$$

Recording scene for aid only synthesis

Independent of the selected playback method, the impulse responses of the transfer paths from the digital signal s_{ls} to be played back by the loudspeaker to the hearing aid microphones' output sequences are recorded. The corresponding transfer functions are given as follows:

$$\begin{aligned}\mathbf{H}^{\text{ind}}_{\text{rec}_{\text{AOBRS}}}(\vec{x}_{\text{h}_{\text{rec}}},\vec{x}_{\text{ls}_{\text{rec}}},\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}}) &= \frac{\mathbf{S}^{\text{ind}}_{\text{ham}}(\vec{x}_{\text{h}_{\text{rec}}},\vec{x}_{\text{ls}_{\text{rec}}},\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}})}{S_{\text{ls}}} \\ &= H_{\text{o}_{\text{rec}}}\cdot\mathbf{H}^{\text{ind}}_{\text{ls,ha}}(\vec{x}_{\text{h}_{\text{rec}}},\vec{x}_{\text{ls}_{\text{rec}}},\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}})\cdot\mathbf{H}_{\text{i}_{\text{ha}}} \\ &= H_{\text{o}_{\text{rec}}}\cdot\mathbf{H}^{\text{ind}}_{\text{ls,ham}}(\vec{x}_{\text{h}_{\text{rec}}},\vec{x}_{\text{ls}_{\text{rec}}},\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}})\cdot\mathbf{H}_{\text{ham}}\cdot\mathbf{H}_{\text{i}_{\text{ha}}}\end{aligned} \tag{34}$$

Playback scene for aid only synthesis with electronic playback

Since the hearing aid's digital inputs are assumed to be used here, it is possible to send the digital signals to be played back ($\mathbf{s}_{\text{hp}}$) directly to the hearing aid inputs. If equal sample rates and word lengths are assumed (which could easily be assured, if necessary), and $\mathbf{s}_{\text{ha}}$ denotes the hearing aid input signals, $\mathbf{S}_{\text{ha}} = \mathbf{S}_{\text{hp}}$ holds true. The transfer functions accounting for the playback procedure in the aid only synthesis case with electronic playback are therefore given as follows:

$$\mathbf{H}^{\text{ind}}_{\text{play}_{\text{AOBRS}}}(\vec{\mathbf{x}}_{\text{ha}_{\text{play}}}) = \frac{\mathbf{S}_{\text{ha}}}{\mathbf{S}_{\text{hp}}} = 1 \tag{35}$$

Playback scene for aid only synthesis with acoustic playback

If the sound pressure signals at the hearing aids are excited by headphones, which are driven by the voltages $\mathbf{u}_{\text{hp}}$, the following transfer functions describe their relation:

$$\mathbf{H}^{\text{ind,h}}_{\text{hp,ham}}(\vec{\mathbf{x}}_{\text{ha}},\vec{\mathbf{x}}_{\text{hp}}) = \frac{\mathbf{P}^{\text{ind,h}}_{\text{ham}}(\vec{\mathbf{x}}_{\text{hp}},\vec{\mathbf{x}}_{\text{ha}})}{\mathbf{U}_{\text{hp}}} \tag{36}$$

The transfer functions relating the headphone input voltages to the hearing aid microphone output voltages $\mathbf{u}^{\text{ind,h}}_{\text{ham}}(\vec{\mathbf{x}}_{\text{hp}},\vec{\mathbf{x}}_{\text{ha}})$ are defined by:

$$\begin{aligned}\mathbf{H}^{\text{ind,h}}_{\text{hp,ha}}(\vec{\mathbf{x}}_{\text{ha}},\vec{\mathbf{x}}_{\text{hp}}) &= \frac{\mathbf{U}^{\text{ind,h}}_{\text{ham}}(\vec{\mathbf{x}}_{\text{hp}},\vec{\mathbf{x}}_{\text{ha}})}{\mathbf{U}_{\text{hp}}} \\ &= \mathbf{H}^{\text{ind,h}}_{\text{hp,ham}}(\vec{\mathbf{x}}_{\text{ha}},\vec{\mathbf{x}}_{\text{hp}})\cdot\mathbf{H}_{\text{ham}}\end{aligned} \tag{37}$$

Using the definitions given above, the following transfer functions describe the playback situation from the sequences $\mathbf{s}_{\text{hp}}$ to the sequences $\mathbf{s}^{\text{ind,h}}_{\text{ham}}(\vec{\mathbf{x}}_{\text{hp}_{\text{play}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}})$ encoding the sound pressure signals at the hearing aid microphones:

$$\begin{aligned}\mathbf{H}^{\text{ind,h}}_{\text{play}_{\text{AOBRS}}}(\vec{\mathbf{x}}_{\text{hp}_{\text{play}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}}) &= \frac{\mathbf{S}^{\text{ind,h}}_{\text{ham}}(\vec{\mathbf{x}}_{\text{hp}_{\text{play}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}})}{\mathbf{S}_{\text{hp}}} \\ &= \mathbf{H}_{\text{O}_{\text{hp}}}\cdot\mathbf{H}^{\text{ind,h}}_{\text{hp,ham}}(\vec{\mathbf{x}}_{\text{hp}_{\text{play}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}})\cdot\mathbf{H}_{\text{ham}}\cdot\mathbf{H}_{\text{i}_{\text{ha}}}\end{aligned} \tag{38}$$

The additional index h indicates consideration under the headphones, and therefore the acoustic playback procedure.

Non-equalized aid only synthesis with electronic playback

As for normal hearing listeners, the convolution of an audio signal with the recorded impulse responses and following presentation of the convolution products to a listener is referred to as the (temporary) unequalized aid only synthesis scene (index *tmp*). The transfer functions describing this situation for AOBRS with electronic playback can be given using equation 34 and equation 35 by:

$$\begin{aligned}&\mathbf{H}^{\text{ind}}_{\text{tmp}_{\text{AOBRS}}}(\vec{x}_{\text{h}_{\text{ref}}},\vec{x}_{\text{ls}_{\text{ref}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}},\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}}) = \\ &= \mathbf{H}^{\text{ind}}_{\text{rec}_{\text{AOBRS}}}(\vec{x}_{\text{h}_{\text{ref}}},\vec{x}_{\text{ls}_{\text{ref}}},\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}})\cdot\mathbf{H}^{\text{ind}}_{\text{play}_{\text{AOBRS}}}(\vec{\mathbf{x}}_{\text{ha}_{\text{play}}}) \\ &= H_{\text{o}_{\text{rec}}}\cdot\mathbf{H}^{\text{ind}}_{\text{ls,ham}}(\vec{x}_{\text{h}_{\text{ref}}},\vec{x}_{\text{ls}_{\text{ref}}},\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}})\cdot\mathbf{H}_{\text{ham}}\cdot\mathbf{H}_{\text{i}_{\text{ha}}}\end{aligned} \tag{39}$$

Non-equalized aid only synthesis with acoustic playback

The temporary unequalized AOBRS scene for acoustic playback can be described using equation 34 and equation 38 by its transfer functions:

$$\begin{aligned}&\mathbf{H}^{\text{ind,h}}_{\text{tmp}_{\text{AOBRS}}}(\vec{x}_{\text{h}_{\text{ref}}},\vec{x}_{\text{ls}_{\text{ref}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}},\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}},\vec{\mathbf{x}}_{\text{hp}_{\text{play}}}) = \\ &= \mathbf{H}^{\text{ind}}_{\text{rec}_{\text{AOBRS}}}(\vec{x}_{\text{h}_{\text{ref}}},\vec{x}_{\text{ls}_{\text{ref}}},\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}})\cdot \\ &\quad\cdot\mathbf{H}^{\text{ind,h}}_{\text{play}_{\text{AOBRS}}}(\vec{\mathbf{x}}_{\text{hp}_{\text{play}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}}) \\ &= H_{\text{o}_{\text{rec}}}\cdot\mathbf{H}^{\text{ind}}_{\text{ls,ham}}(\vec{x}_{\text{h}_{\text{ref}}},\vec{x}_{\text{ls}_{\text{ref}}},\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}})\cdot\mathbf{H}_{\text{ham}}\cdot\mathbf{H}_{\text{i}_{\text{ha}}}\cdot \\ &\quad\cdot\mathbf{H}_{\text{O}_{\text{hp}}}\cdot\mathbf{H}^{\text{ind,h}}_{\text{hp,ham}}(\vec{\mathbf{x}}_{\text{hp}_{\text{play}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}})\cdot\mathbf{H}_{\text{ham}}\cdot\mathbf{H}_{\text{i}_{\text{ha}}}\end{aligned} \tag{40}$$

Equalization requirements for aid only synthesis with electronic playback

As in section *Equalization requirements for binaural room synthesis with normal hearing listeners*, assuming invertibility of $\mathbf{H}^{\text{ind}}_{\text{tmp}_{\text{AOBRS}}}(\vec{x}_{\text{h}_{\text{ref}}},\vec{x}_{\text{ls}_{\text{ref}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}},\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}})$, the equalization requirements for AOBRS with electronic playback can be defined using equation 33 and 39 by following transfer function:

$$\begin{aligned}&\mathbf{H}^{\text{ind}}_{\text{eq}_{\text{AOBRS}}}(\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}}) = \\ &= \frac{\mathbf{H}^{\text{ind}}_{\text{ref}_{\text{AOBRS}}}(\vec{x}_{\text{h}_{\text{ref}}},\vec{x}_{\text{ls}_{\text{ref}}},\vec{\mathbf{x}}_{\text{ha}_{\text{ref}}})}{\mathbf{H}^{\text{ind}}_{\text{tmp}_{\text{AOBRS}}}(\vec{x}_{\text{h}_{\text{ref}}},\vec{x}_{\text{ls}_{\text{ref}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}},\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}})} \\ &= \frac{H_{\text{o}_{\text{ref}}}\cdot\mathbf{H}^{\text{ind}}_{\text{ls,ham}}(\vec{x}_{\text{h}_{\text{ref}}},\vec{x}_{\text{ls}_{\text{ref}}},\vec{\mathbf{x}}_{\text{ha}_{\text{ref}}})\cdot\mathbf{H}_{\text{ham}}\cdot\mathbf{H}_{\text{i}_{\text{ha}}}}{H_{\text{o}_{\text{rec}}}\cdot\mathbf{H}^{\text{ind}}_{\text{ls,ham}}(\vec{x}_{\text{h}_{\text{ref}}},\vec{x}_{\text{ls}_{\text{ref}}},\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}})\cdot\mathbf{H}_{\text{ham}}\cdot\mathbf{H}_{\text{i}_{\text{ha}}}}\end{aligned} \tag{41}$$

Assuming further that the reference scene's output equipment is used in the recording situation, i. e. $H_{\text{o}_{\text{rec}}} = H_{\text{o}_{\text{ref}}}$, and that the head, loudspeaker, and hearing aid positions are identical in reference, recording, and playback situation ($\vec{x}_{\text{h}_{\text{rec}}} = \vec{x}_{\text{h}_{\text{ref}}}$, $\vec{x}_{\text{ls}_{\text{rec}}} = \vec{x}_{\text{ls}_{\text{ref}}}$, and $\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}} = \vec{\mathbf{x}}_{\text{ha}_{\text{ref}}}$), equation 41 can be simplified as follows:

$$\mathbf{H}^{\text{ind}}_{\text{eq}_{\text{AOBRS}}}(\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}}) = 1 \tag{42}$$

In other words: If the hearing aid's digital inputs are used for playback of the synthesis results, under the given assumptions, no equalization is required in the aid only synthesis case.

Equalization requirements for aid only synthesis with acoustic playback

Assuming again that the reference scene's output equipment is used in the recording situation, i. e. $H_{\text{o}_{\text{rec}}} = H_{\text{o}_{\text{ref}}}$, and that the head, loudspeaker, and hearing aid positions are identical in reference, recording, and playback situation ($\vec{x}_{\text{h}_{\text{rec}}} = \vec{x}_{\text{h}_{\text{ref}}}$, $\vec{x}_{\text{ls}_{\text{rec}}} = \vec{x}_{\text{ls}_{\text{ref}}}$, and $\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}} = \vec{\mathbf{x}}_{\text{ha}_{\text{ref}}}$), the equalization requirements for AOBRS with acoustic playback are given using the equations 33 and 40 as follows:

$$\begin{aligned}&\mathbf{H}^{\text{ind,h}}_{\text{eq}_{\text{AOBRS}}}(\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}},\vec{\mathbf{x}}_{\text{hp}_{\text{play}}}) = \\ &= \frac{\mathbf{H}^{\text{ind}}_{\text{ref}_{\text{AOBRS}}}(\vec{x}_{\text{h}_{\text{ref}}},\vec{x}_{\text{ls}_{\text{ref}}},\vec{\mathbf{x}}_{\text{ha}_{\text{ref}}})}{\mathbf{H}^{\text{ind,h}}_{\text{tmp}_{\text{AOBRS}}}(\vec{x}_{\text{h}_{\text{ref}}},\vec{x}_{\text{ls}_{\text{ref}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}},\vec{\mathbf{x}}_{\text{ha}_{\text{rec}}},\vec{\mathbf{x}}_{\text{hp}_{\text{play}}})} \\ &= \frac{1}{\mathbf{H}_{\text{O}_{\text{hp}}}\cdot\mathbf{H}^{\text{ind,h}}_{\text{hp,ham}}(\vec{\mathbf{x}}_{\text{hp}_{\text{play}}},\vec{\mathbf{x}}_{\text{ha}_{\text{play}}})\cdot\mathbf{H}_{\text{ham}}\cdot\mathbf{H}_{\text{i}_{\text{ha}}}}\end{aligned} \tag{43}$$

If the sound pressure at a hearing aid is generated by means of headphones, comparison to equation 19 reveals that it is possible to equalize an aid only synthesis system with the inverses of the headphone impulse responses recorded at the hearing aid microphones according to equation 19 so that the synthesis equals the reference scene. Necessary premises are constant hearing aid microphone and headphone positions for the headphone impulse response measurement, the recording, and the playback situation.

COMBINED SYNTHESIS

Most conventional behind the ear models belong to the class of hearing aids for which combined synthesis reveals the appropriate procedure. Further, in ear models not occluding the auditory canal and electro-acoustical stimulation have to be supplied with CBRS. For combined synthesis, the situation is somewhat more complicated than with aid only synthesis.

Reference scene for combined synthesis

Since different contributions have to be taken into account that provide input to the auditory system, it is not possible to use the hearing aid's input signals as reference, as it is possible for aid only synthesis. Otherwise, one would neglect the contributions to the ear signals caused directly from the sound field. Using the output signals of probe microphones right in front of the subject's eardrums, as for normal hearing subjects, could be useful for hearing aids with exclusive acoustical stimulation, but would include a specific hearing aid's characteristics in the synthesis procedure and is not applicable to electro-acoustical stimulation. Therefore, it is not considered for the more general point of view discussed within this paper. Nevertheless, the presented theoretical framework can easily be extended to cover this reference, too.

A reference situation that provides coverage of possible application scenarios can be deduced from the ear signals under the hearing aids. The transfer functions describing the relation between the loudspeaker driving signal and a hearing impaired listener's ear signals are given using equation 28 and equation 32 as follows:

$$\begin{aligned}
&\mathbf{H}^{\mathrm{ind}}_{\mathrm{reftmp_{CBRS}}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}},\vec{\mathbf{x}}_{\mathrm{hals_{ref}}}) = \\
&= \frac{\mathbf{P}^{\mathrm{ind,ha}}_{\mathrm{e}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}},\vec{\mathbf{x}}_{\mathrm{hals_{ref}}})}{S_{\mathrm{ls}}} \\
&= \frac{\mathbf{P}^{\mathrm{ind,ha}}_{\mathrm{e_{ha}}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}},\vec{\mathbf{x}}_{\mathrm{hals_{ref}}})}{S_{\mathrm{ls}}} \\
&\quad + \frac{\mathbf{P}^{\mathrm{ind,ha}}_{\mathrm{e_{sf}}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}},\vec{\mathbf{x}}_{\mathrm{hals_{ref}}})}{S_{\mathrm{ls}}} \\
&= \mathbf{H}^{\mathrm{ind}}_{\mathrm{ha,AS}}(\vec{\mathbf{x}}_{\mathrm{hals_{ref}}}) \cdot \frac{\mathbf{S}^{\mathrm{ind}}_{\mathrm{ham}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}})}{S_{\mathrm{ls}}} \\
&\quad + \frac{\mathbf{P}^{\mathrm{ind,ha}}_{\mathrm{e_{sf}}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}},\vec{\mathbf{x}}_{\mathrm{hals_{ref}}})}{S_{\mathrm{ls}}}
\end{aligned} \tag{44}$$

The hearing aid transfer function does not necessarily have to be included in the synthesis, in fact, as mentioned above, a more flexible synthesis system arises if the sound pressure $\mathbf{P}^{\mathrm{ind}}_{\mathrm{ham}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}})$ at the hearing aid microphones or the digitized microphone output signal $\mathbf{S}^{\mathrm{ind}}_{\mathrm{ham}}(\vec{x}_{\mathrm{h}},\vec{x}_{\mathrm{ls}},\vec{\mathbf{x}}_{\mathrm{ha}})$ is taken as reference for the partial path through the hearing aid. The latter procedure results in the following reference situation for CBRS:

$$\begin{aligned}
\mathbf{H}^{\mathrm{ind}}_{\mathrm{ref_{CBRS}}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}}) &= \frac{\mathbf{S}^{\mathrm{ind}}_{\mathrm{ham}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}})}{S_{\mathrm{ls}}} \\
&+ \frac{\mathbf{P}^{\mathrm{ind,ha}}_{\mathrm{e_{sf}}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}},\vec{\mathbf{x}}_{\mathrm{hals_{ref}}})}{S_{\mathrm{ls}}}
\end{aligned} \tag{45}$$

Regarding this reference, two different possible approaches for signal playback in combined synthesis become evident, which have to be considered for selection of the recording scene, too:

1. *Acoustic playback:* Headphones are used to play back the synthesis' output signals. This way, the hearing aid microphones as well as the listener's eardrum at the respective ear are stimulated.
2. *Electro-acoustic playback:* Headphones are used to stimulate the listener's eardrum, but the hearing aid inputs are stimulated electronically.

Combined synthesis with acoustic playback

The acoustic playback situation can be regarded as a combination of one part comparable to the headphone playback for normal hearing listeners and another part similar to the acoustic playback with aid only synthesis. Spoken descriptively, this requires two different equalization processes at one time. In theory, it is possible to perform such an equalization procedure, for example using physically separated transducers for both channels or a so called crosstalk cancellation approach. From a practical point of view, problems like time variance or providing sufficient sound insulation make the realization of such systems rather difficult. This is why the acoustic playback for combined synthesis is not discussed within this paper. However, it might be worth the effort to implement combined synthesis with acoustic playback for specific application scenarios. The theoretical framework introduced here is fully capable to describe this situation, too.

Reference scene for combined synthesis with electro-acoustic playback

The electro-acoustic playback situation can be regarded as a combination of two contributions: One part comparable to the headphone playback for normal hearing listeners and one part comparable to the electronic playback for aid only synthesis. This way, both contributions can be equalized separately, since crosstalk between the channels can not occur. An additional advantage arising from this procedure is the possibility to define a reference scene description for each contribution independently of each other. The only requirement is that the same situation with the same input signal is regarded in both cases.

The electronic contribution's reference situation can be given as follows:

$$\begin{aligned}
\mathbf{H}^{\mathrm{ind}}_{\mathrm{ref_{CBRS,el}}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}}) &= \frac{\mathbf{S}^{\mathrm{ind}}_{\mathrm{ham}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}})}{S_{\mathrm{ls}}} \\
&= \mathbf{H}^{\mathrm{ind}}_{\mathrm{ref_{AOBRS}}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}})
\end{aligned} \tag{46}$$

Therefore, the complete synthesis procedure from the electronic playback with aid only synthesis can be adapted.

The acoustic contribution's reference is defined by:

$$\begin{aligned}
&\mathbf{H}^{\mathrm{ind,ha}}_{\mathrm{ref_{CBRS,ac}}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}},\vec{\mathbf{x}}_{\mathrm{hals_{ref}}}) = \\
&= \frac{\mathbf{P}^{\mathrm{ind,ha}}_{\mathrm{e_{sf}}}(\vec{x}_{\mathrm{h_{ref}}},\vec{x}_{\mathrm{ls_{ref}}},\vec{\mathbf{x}}_{\mathrm{ha_{ref}}},\vec{\mathbf{x}}_{\mathrm{hals_{ref}}})}{S_{\mathrm{ls}}}
\end{aligned} \tag{47}$$

This reference indicates that the acoustic part is comparable to binaural room synthesis for normal hearing listeners (cf. eq. 10). The only difference is that all involved measurements have to be done with the probe microphone under the hearing aid.

Employing this procedure, combined synthesis can be used to synthesize a target acoustical scenario by means of the physical properties involved in the hearing process as it was in the intended reference situation.

POSSIBLE APPLICATIONS

Using the presented theory, binaural room synthesis is capable of synthetically generating the physical signals involved in the hearing process within a specific listening situation for hearing impaired listeners using a hearing aid. In detail, these signals are the time varying hearing aid input signals and the ear signals. Possibly, a highly detailed synthesis is not necessary, since hearing impaired listeners usually use just a few localization cues. For that reason, it could be sufficient to simulate only the most prominent localization cues. However, the aim of the paper on hand is not to achieve good localization results, but to synthetically generate the physical conditions of a specific listening situation as correctly as possible, to allow for simulation of realistic conditions.

Typical applications of this procedure are research and fitting purposes in the field of hearing aids or cochlear implants. In this context, binaural synthesis can save effort and time by providing arbitrary acoustical environments in the laboratory. One major advantage compared to traditional procedures is the high reproducibility of a completely controllable acoustical scenario. This is especially interesting since not only typical stimulus properties like the level or spectral content are intended to be reproduced, but also the perceived location of the sound sources is to be preserved. This results in a freely arrangeable but controllable and repeatable spatial distribution of discrete sound sources with relatively low effort. The most important hardware required is a computer with a sound interface and a head tracking device. Dependent on the actual implementation, headphones might be necessary in addition. The remaining items required are the hearing aid or a hearing aid dummy and a specialized software algorithm for the necessary computations.

A combination with a hearing aid simulation is easily possible using a hearing aid dummy, which opens a broad field of applications, for example in the evaluation of hearing aid algorithms (cf. e. g. Kayser et al. 2009). This procedure is especially interesting in the combined synthesis case, since with the different propagation paths consideration of different delays introduced by the signal processing within the hearing aid is possible, causing out of phase signals at the eardrums, which is a common problem in hearing aid fitting.

SUMMARY

Within this paper, an adaption of the traditional binaural synthesis theory for normal hearing listeners to hearing impaired subjects using hearing aids is presented. It is shown that exclusive acoustic stimulation is only applicable when using a hearing aid that completely blocks the auditory canal. In every other case, to achieve a theory applicable in general, a combined system has to be used, which stimulates the hearing aid input electronically while stimulating the eardrum acoustically.

ACKNOWLEDGMENTS

Parts of this work were supported by grant FA 140/4 of the Deutsche Forschungsgemeinschaft (DFG).

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 ed. (The MIT Press, Cambridge, Massachusetts, London, England, 1997)

Blauert J., H. Lehnert, J. Sahrhage, H. Strauss: An Interactive Virtual-Environment Generator for Psychoacoustic Research. I: Architecture and Implementation. ACTA ACUSTICA UNITED WITH 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. ACTA ACUSTICA UNITED WITH ACUSTICA **86**, 1046–1053 (2000)

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: Methods for Binaural Recording and Reproduction. ACTA ACUSTICA UNITED WITH ACUSTICA **88**, 303–311 (2002)

Kayser H., S. D. Ewert, J. Anemüller, T. Rohdenburg, V. Hohmann, B. Kollmeier: Database of Multichannel In-Ear and Behind-the-Ear Head-Related and Binaural Room Impulse Responses. EURASIP Journal on Advances in Signal Processing **2009**, Article ID 298605 (2009)

Mackensen P., U. Felderhoff, G. Theile, U. Horbach, R. Pellegrini: Binaural Room Scanning - A new Tool for Acoustic and Psychoacoustic Research. ACTA ACUSTICA UNITED WITH ACUSTICA **85**, 417 (1999)

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

Oppenheim A. V., A. S. Wilskey: *Signals and Systems* (Prentice-Hall Inc., 1983)

Völk F.: Externalization in data-based binaural synthesis: effects of impulse response length. In *Proc. of Intern. Conf. on Acoustics NAG/DAGA 2009* (Dt. Gesell. für Akustik e. V., Berlin, 2009)

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

Völk F., E. Faccinelli, H. Fastl: Überlegungen zu Möglichkeiten und Grenzen virtueller Wellenfeldsynthese (Considerations of possibilities and limitations of virtual wave field synthesis). In *Fortschritte der Akustik, DAGA 2010* (Dt. Gesell. für Akustik e. V., Berlin, 2010)

Völk F., F. Heinemann, H. Fastl: Externalization in binaural synthesis: effects of recording environment and measurement procedure. In *Proceedings of Acoustics '08*, 6419–6424 (2008a)

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 laboratory). In *Fortschritte der Akustik, DAGA 2007* (Dt. Gesell. für Akustik e. V., Berlin, 2007)

Völk F., J. Konradl, H. Fastl: Simulation of wave field synthesis. In *Proc. of Acoustics '08*, 1165–1170 (2008b)

Völk F., T. Musialik, H. Fastl: Crosstalk cancellation between phantom sources. In *126th AES Convention* (2009a) (Convention Paper 7722)

Völk F., M. Straubinger, L. Roalter, H. Fastl: Measurement of head related impulse responses for psychoacoustic research. In *Proc. of Intern. Conf. on Acoustics NAG/DAGA 2009* (Dt. Gesell. für Akustik e. V., Berlin, 2009b)

Wilkens H.: Kopfbezügliche Stereofonie - ein Hilfsmittel für Vergleich und Beurteilung verschiedener Raumeindrücke (Head related stereophonic reproduction - an aid in comparison and assessment of different room impressions). 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)*. Third ed. (Springer, Berlin, 1993)