Proceedings of 20th International Congress on Acoustics, ICA 2010

23–27 August 2010, Sydney, Australia

Psychoacoustical experiments on loudness perception in wave field synthesis

Florian Völk, Martin Straubinger, and Hugo Fastl

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

florian.voelk@mytum.de

PACS: 43.66.Cb

ABSTRACT

Wave field synthesis is an audio reproduction procedure aiming to produce a sound field completely correct at least within the so called listening area. In theory, the exact synthesis is possible under certain circumstances based on the Kirchhoff-Helmholtz integral equation. For practical realizations, usually closed loudspeaker boxes are employed. However, using closed loudspeaker boxes leads to a deviation of the synthesized from the intended wave field and requires approximations in the derivation of the loudspeaker driving signals, since preliminaries of the Kirchhoff-Helmholtz integral equation are not fulfilled anymore. If wave field synthesis is employed in a typical listening environment, room influences further disturb the resulting sound field. Whether auditory perceptions generated by such a system equal those occurring in the reference scene is unclear. Within this contribution, loudness adjustments of narrow band noises presented by wave field synthesis to the free field case are shown. The results of these experiments under different room acoustical conditions are discussed and compared to a baseline, where a single loudspeaker in the same reproduction room was adjusted to the free field situation. Further, the acquired data are compared to the prognoses of a typical loudness model according to DIN 45 631/A1.

INTRODUCTION

Wave field synthesis (WFS, Berkhout et al. 1993) is an audio playback technology aiming to reproduce the sound field physically present in the listening situation to be simulated (the primary sound field) within the so called listening volume V (cf. Berkhout 1988, Vogel 1993, Boone et al. 1995). To achieve this goal, the two dimensional boundary area ∂V surrounding V theoretically has to be completely covered with a continuous monopole and dipole source distribution (secondary source distribution, cf. Start 1997).

For realizations, closed dynamic loudspeaker boxes are typically used instead of the monopole point sources (cf. Theile and Wittek 2004). This causes errors in the resulting wave field, since loudspeaker boxes deviate from monopole characteristics, especially at high frequencies (Zollner and Zwicker 1993). This can to a certain extent be accounted for by modifying the loudspeaker signals (for example de Vries 1996, Ahrens and Spors 2009b). The dipole sources are usually neglected, which has to be accounted for in the derivation of the loudspeaker input signals (cf. Verheijen 1997, Hulsebos 2004). In typical use cases, this accounting can be done only approximately (cf. Spors 2007, Melchior et al. 2008), causing further artifacts in the wave field produced. Using loudspeaker boxes with a physical extent of at least some centimeters means a spatial sampling of the continuous secondary source distribution with an alias frequency within the audible frequency range (cf. Wittek and Augustin 2005, Wittek et al. 2007, Wittek 2007, Corteel et al. 2008). In addition, due to the enormous hardware effort required, the listening volume is usually reduced to a listening area with a one dimensional boundary ∂V (cf. Spors et al. 2008). As a consequence, the number of loudspeakers required decreases remarkably, but additional errors are introduced in the synthesized wave field (cf. e. g. Ahrens and Spors 2008).

The fundamental concern of Psychoacoustics is to establish relationships between physical stimulus properties and corresponding hearing sensations (see for example Stevens and Davis 1938, Feldtkeller and Zwicker 1956). Traditionally, a frontally incident plane wave, propagating in a free sound field is regarded as the most simple stimulus for psychoacoustic studies (Zwicker and Feldtkeller 1967), since it can easily be described by closed mathematical equations. For studies concerning directional hearing, more complex scenarios, consisting of one or more point sources are typically used (cf. Blauert 1997), and the stimulus (the spatio-temporal acoustical wave field) becomes more complicated. In every case, the physical properties of the considered stimuli (in directional hearing especially the source positions and extensions) are mathematically clearly describable, which allows for definition of psychoacoustical relationships.

When dealing with wave field synthesis, the situation becomes more complicated. For wave field synthesis rendering, physically existing secondary sound sources are used to create a sound field that would originally be created by one or more virtual (physically not existing) primary sources. Conducting a psychoacoustical experiment in wave field synthesis usually is done by varying a physical property of the physically not present primary source by adjusting the sound field rendering algorithm and inspecting related changes in hearing sensations. In other words, the virtual primary source's properties are regarded as the stimulus magnitudes. This procedure is insofar in line with traditional psychoacoustical experiments, as relations between changes in physical stimuli and corresponding changes in auditory sensations are assessed. However, a major difference to traditional experiments is that the stimulus changes are caused indirectly by changing the rendering equations. This results in a considerable restriction of generality, since the specific wave field synthesis setup's characteristics (signal processing and hardware influences) are inevitably included in the results. For that reason, in addition to the traditional description of the experimental setup, detailed knowledge about the synthesis algorithm (including its implementation) and about the complete electro-acoustical signal processing chain has to be associated with results of psychoacoustical experiments concerning wave field synthesis, to allow for meaningful discussion of the results.

One could argue that psychoacoustical studies concerning wave field synthesis are redundant, since the aim of the WFS procedure is to synthesize the sound field created by one or more primary sources, so the psychoacoustical data known for real sources can be employed to describe a virtual scenario. This argumentation is not accurate, since in wave field synthesis both, the reduction to a two dimensional listening area as well as the employment of physically extended loudspeaker boxes, whose radiation patterns deviate especially at high frequencies from monopole characteristics, neglecting the dipole sources required, introduce errors in the synthesized sound field. When reproducing in a reflective environment, listening room influences further disturb the resulting wave field (cf. Betlehem and Abhayapala 2005). Therefore, it is not to be expected that the auditory perceptions evoked by the secondary equal those evoked by the primary field and it is inadmissible to consult psychoacoustical relationships determined for real (physically present) sound sources for wave field synthesis without verifying their validity for the considered case (e. g. Völk 2010).

It should be mentioned here that it is possible to partially compensate for undesired room influences on wave field synthesis (cf. e. g. Spors et al. 2003a,b, Corteel and Nicol 2003, Spors et al. 2004, Spors 2005, Corteel 2006, Spors et al. 2007, Gauthier and Berry 2006, 2007, 2008a,b). The paper on hand focuses on wave field synthesis without reproduction room compensation, since the perceptual consequences of room influences are currently widely unclear (cf. e. g. Spors 2005).

FUNDAMENTALS AND INITIAL SITUATION

At first, the most relevant facts on loudness perception and wave field synthesis, as well as the ideas of loudness modeling and free field equalization are summed up here. Based on these data, the aim of the study on hand is formulated. Within this paper, the terms critical band rate (Feldtkeller 1955, Zwicker 1961) and frequency are used interchangeably, thinking in both cases of an auditory adequately spaced frequency scale.

Loudness

Loudness is the psychoacoustical magnitude that describes how loud a stimulus is perceived (cf. Zwicker 1959). "The sensation that corresponds most closely to the sound intensity of the stimulus is loudness" (cf. Fastl and Zwicker 2007, p. 205). For that reason, loudness, if assessed with narrow band stimuli at different critical band rates, is assumed to be suited to reveal unexpected frequency dependencies of a reproduction system's absolute transfer function, for example of a WFS setup. Since loudness depends not only on the intensity of a test sound, but also on its frequency content, bandwidth, or duration, the choice of an appropriate stimulus for the assessment of reproduction system influences by loudness experiments is of essential importance for the results' applicability (cf. section *Methods and stimuli*). In addition, a baseline scenario has to be defined for the system under consideration to be compared against.

Free field equalization

The loudness perception produced by a plane monochromatic wave under anechoic conditions can also be elicited by free field equalized headphones (cf. Zwicker and Maiwald 1963, Villchur 1969). For measurement of the free field equalization (FFE), subjects have to adjust the voltage at the input terminals of the headphones used, so that the headphone presentation elicits the same loudness as presentation in the free field (cf. DIN 45 619). The procedure will be summed up in the following: In turn a sound from the headphones and the same sound in the free field are presented. In between the two, the subject's task is to remove the headphones and to compare the loudness of the presentations. Afterwards, the subject indicates whether the headphones produced more or less loudness than the free field and puts on the headphones again. The voltage at the headphones is adjusted with respect to the subject's judgment and the procedure is repeated until a certain limit is reached. Usually, narrow band signals such as tones (cf. e. g. Fastl and Fleischer 1978) or narrow band noises (DIN 45 619) are employed as test sounds. If during the free field equalization process a calibration is determined also, following situation is achieved: narrow band signals presented at a certain so called free field equivalent level elicit the same loudness perception as if they were presented as plane propagating waves at the mentioned level in the free field.

Primary sources in wave field synthesis

Typical primary sources to be simulated by means of wave field synthesis are spherical and plane waves under anechoic conditions. When WFS is set up in a typical reflective listening environment, the resulting sound field deviates from the intended due to approximations in the derivation of the WFS equations as well as undesired room influences. For that reason, it is unclear whether the auditory perceptions elicited by WFS equal those produced by FFE.

Modeling loudness perception

Loudness models are intended to predict by computation the loudness to be expected for a certain stimulus and distinct listening conditions based on the stimulus' physical properties (cf. e. g. Chalupper and Fastl 2002). There exist different standards for calculating the loudness for stationary (e. g. DIN 45 631, ANSI S 3.4) or time-varying sounds (e. g. DIN 45 631/A1), reflecting the perceived loudness for different sounds with differing accuracy (Fastl et al. 2009, Rennies et al. 2010).

Aim of the present work

The aim of the experiments presented in this contribution is to compare the loudness perceptions elicited by wave field synthesis and free field equalized headphones. For that purpose, subjects adjusted the level of critical band wide narrow band noises presented using WFS in a reflective laboratory environment by an adaptive forced choice procedure, so that the noises elicited the same loudness as if they were presented by FFE at 65 dB SPL. This procedure was repeated for a different orientation of the WFS loudspeaker array and the listener towards the room, to assess room influences. Further, for comparison to traditional monophonic reproduction, the procedure was repeated for single loudspeaker reproduction in the same laboratory. From these experiments, the following achievements are to be expected:

1. for wave field synthesis and loudspeaker reproduction the frequency dependent level necessary at the listening point to produce the same loudness as critical band wide narrow band noise presented as plane wave with 65 dB SPL
2. the frequency dependent level difference at equal loudness between loudspeaker and WFS reproduction in the reproduction room under consideration
3. the frequency dependent level difference at equal loudness between two orientations of the WFS system in the reproduction room under consideration

With additional consideration of predictions by a loudness model, the following additional subjects can be discussed:

4. verification of the predicted loudness (according to DIN 45 631) of narrow band noises for loudspeaker reproduction in the room considered by comparison to the predictions for the equally loud free field situation
5. verification of the predicted loudness for wave field synthesis reproduction by comparison to the predictions for the equally loud free field situation

These results are expected to help in understanding loudness perception in wave field synthesis reproduction and to verify model predictions for this situation and for loudspeaker reproduction in a reverberant environment. Further, artifacts resulting from the approximations in the WFS derivation and having influence on the loudness of narrow band noises can be identified.

SETUP AND PROCEDURE

Before the experimental procedure and the selection of the stimuli used is discussed, the hardware employed and the listening room environment is introduced in this section. In addition, details on the playback methods are given.

Reproduction room and geometrical setup

All experiments presented within this paper were conducted in a laboratory room (6.8 m × 3.9 m × 3.3 m) at Lehrstuhl für Mensch-Maschine-Kommunikation of Technische Universität München. Figure 1 shows its reverberation time averaged over measurements from each involved array loudspeaker at its position during the experiment to two microphones at the listening point (midpoint of the loudspeaker array) and 10 cm aside.

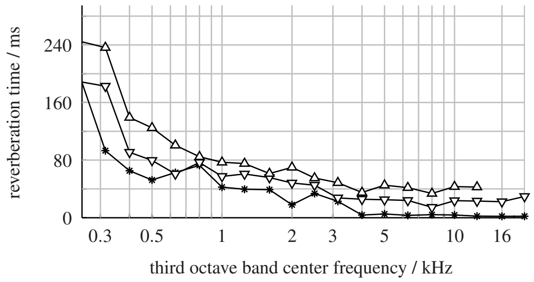

Figure 1: Reverberation time of the laboratory room where all listening experiments presented here took place. Early decay time (stars), T_{20} (downward pointing triangles), and T_{30} (upward pointing triangles).

The experiments are carried out in complete darkness. The laboratory is darkened after the subjects are instructed and positioned on a chair in such a way that the midpoint of their head was positioned at the reference point. The position is adjusted using a head tracking system, ensuring an accuracy in the plane defined by the loudspeaker array of ±3 cm as well as an accuracy in height of ±1 cm.

Figure 2 illustrates the geometrical arrangement of the the hardware and the subject within the reproduction room for the different loudspeaker and WFS playback conditions. The subject's position and viewing direction is indicated by the position and point of a triangle. In addition, the loudspeaker and the circular wave field synthesis array is shown receptively, as well as the virtual plane waves' normal vectors.

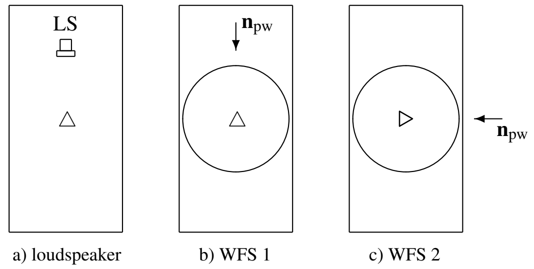

Figure 2: Geometrical configurations. The listening room is indicated as rectangle, subject positions and viewing directions as triangles. In addition, the loudspeaker, the WFS arrays and the normal vectors of the primary plane waves are shown.

Free field equalized diotic presentation

For diotic playback, Beyer DT 48 A headphones with a free field equalizer according to Zwicker and Maiwald (1963) are used (cf. also Fastl and Fleischer 1978, Fastl and Zwicker 2007). This setup allows for calibrated playback, since the voltage at the equalizer's input terminal that is necessary for a 1 kHz tone to elicit the same loudness as a frontally incident plane propagating wave at 80 dB SPL in the free sound field is known.

Wave field synthesis rendering

For the wave field synthesis rendering, a circular loudspeaker array centered at the laboratory room's midpoint is used. The array consists of 96 broadband loudspeaker boxes (Bose Freespace 3 Satellit, loudspeaker spacing $\Delta x = 8.5$ cm, array radius $r = 129.9$ cm). Each loudspeaker's free-field response is equalized using an individually designed FIR-filter with the aim of a frequency independent absolute transfer function on the symmetry axis. The equalized loudspeakers are calibrated so that every single loudspeaker produces a level deviating less than ±0.1 dB from all others when reproducing broad band pink noise.

Wave field synthesis is capable of generating a sound field at a distinct listening position (reference point $\mathbf{x}_{\mathrm{ref}}$) within a limited frequency range. $\hat{P}(\omega)$ denotes the Fourier-spectrum of the sound pressure time signal radiated from the primary source, $\mathbf{x}_0$ the position of the considered secondary source, Δx_0 the distance to the next secondary source, and $\mathbf{n}(\mathbf{x}_0)$ the normal vector on ∂V pointing from $\mathbf{x}_0$ into V (nomenclature adjusted from Spors et al. 2008). Using these conventions, it is possible to describe the loudspeaker signals' Fourier-spectra (the so called driving functions) for the synthesis of plane wave fronts (index pw, synthesis adapted from Spors et al. 2008) as follows:

$$D_{\mathrm{2.5D,pw}}(\mathbf{x}_0,\omega) = a_{\mathrm{pw}}(\mathbf{x}_0)\sqrt{jk}\sqrt{2\pi|\mathbf{x}_{\mathrm{ref}}-\mathbf{x}_0|}\cdot \\ \cdot\, \mathbf{n}_{\mathrm{pw}}^{\mathrm{T}}\mathbf{n}(\mathbf{x}_0)\hat{P}(\omega)e^{-jk\mathbf{n}_{\mathrm{pw}}^{\mathrm{T}}\mathbf{x}_0}\Delta x_0 \tag{1}$$

Here, $\mathbf{n}_{\mathrm{pw}}$ represents the normal vector of the wave front to be synthesized in propagation direction (cf. figure 2), and the following equation applies:

$$a_{\mathrm{pw}}(\mathbf{x}_0) = \begin{cases} 1 & \text{for} \quad \mathbf{n}_{\mathrm{pw}}^{\mathrm{T}}\mathbf{n}(\mathbf{x}_0) > 0, \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

In addition, 20% of the currently active loudspeakers are truncated to both sides using a cosine-shaped window.

Method and stimuli

The method used within this study is adapted from the procedure described in section *Free field equalization* for measurement of the free field equalization (cf. also DIN 45 619, Teil 1). Within the experiments described here, the subjects listen in turn to loudspeaker presentation / wave field synthesis and to free field equalized headphones. In contrast to measurements of the free field equalization, the voltage at the headphone input terminals is kept constant so that a free field equivalent level of 65 dB SPL is presented. The level of the loudspeaker / the wave field synthesis system is adjusted according to the subject's judgments. For selection of the adjustment's step size, the PEST-algorithm (Parameter Estimation by Sequential Testing, cf. Gelfand 1990) is employed with 1 dB minimal step size.

As sound stimuli, critical band wide uniform exciting noises (UEN, ac. to Fastl and Zwicker 2007) centered at critical band rates distributed in steps of 1 Bark between 3.5 and 24.5 Bark are used. The stimulus duration is selected to 1.5 s, with 20 ms Gaussian gating. These stimuli are selected since they represent an equally spaced distribution on an auditory adequate frequency scale. In other words, every stimulus excites the hearing system with the same perceptual weight. The influences of transmission errors in a reproduction system under consideration are therefore assumed to be weighted in a perceptually comparable manner.

After the first signal is played back over the headphones, the subjects have to take off the headphones within a pause of 1.5 s, then the second signal is presented with the loudspeaker system. Within a second pause of 3 s, the subjects indicate their judgment by pressing a foot switch if the loudness elicited by the second stimulus is higher than that elicited by the first, and doing nothing in the other case. Afterwards, the subjects put the headphones back on and the procedure is repeated until the PEST-algorithm indicates no further remarkable change in the adjusted result (in detail: the last four judgments have to be within 4 dB and the step size 1 dB or less). The control processes for the narrow band noises of different center frequencies is carried out in an interleaved manner to prevent recognition effects. The individual result for every control process is computed as median over the last four reversal points.

RESULTS

Within this section, the results of the loudness adjustments are presented. The results are defined as the levels at the listening position in absence of the listener, in the condition adjusted by the respective subject. The following figures show the inter-individual medians and inter-quartile ranges of the individual results from eight normal hearing subjects in age from 22 to 29 years (mean 25.5 years). These same eight subjects participated in all experiments presented.

Loudspeaker reproduction

For the baseline experiment, the loudspeaker (Klein & Hummel O98) was positioned at a distance of 2.45 m from the listening position (cf. figure 2). The mean trial duration is about 33.3 minutes (divided in two sessions separated by at least 15 minutes). In figure 3, the results are given. The median over the medians for all critical band rates resulted to 64.3 dB SPL, none of the individual results show effects that could not be discussed alongside the inter-individual results.

Figure 3 shows in addition to the results from the listening experiment a solid curve. This curve indicates the attenuation a_D, necessary to produce the same equal loudness curves (cf. Fastl et al. 1990) for narrow band noises in a diffuse and in a free sound field (according to Fastl and Zwicker 2007, p. 205), shifted towards higher levels by 65 dB. The line therefore represents the level to be expected for narrow band noises adjusted in the diffuse field to the loudness of free field equalized headphones at 65 dB free field equivalent level. This expectation is motivated as follows: Measurement of equal loudness curves

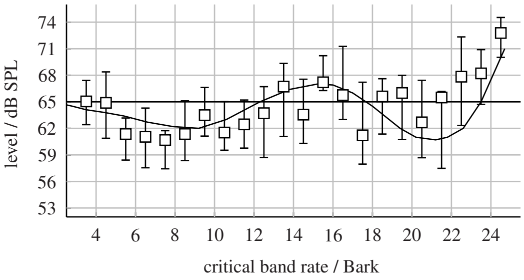

Figure 3: Level of critical band wide narrow band noise (medians as squares with inter-quartile ranges) reproduced by a loudspeaker under reflective listening room conditions, adjusted in loudness to presentation with free field equalized headphones at 65 dB free field equivalent level. The solid line represents the attenuation a_D necessary to produce the same equal loudness curves in a diffuse and free field (ac. to Fastl and Zwicker 2007), shifted upwards by 65 dB.

means adjustment of test sounds (tones or narrow band noises) of different frequencies to another sound (usually a reference stimulus at 1 kHz, that is 8.5 Bark, of the same kind as the test sound, cf. DIN ISO 226). If this adjustment is carried out in a reflective environment, the results differ (independently of the absolute level) by a_D from those acquired under anechoic conditions (cf. Zwicker 1959). In other words: the frequency dependence of loudness differs in free and diffuse sound fields. For example tones between 2 and 4 kHz (13 and 17.5 Bark) have to be presented with a higher level in a diffuse than in a free field, to produce the same loudness as the reference stimulus.

The experiments discussed within this paper were carried out in a damped but still reflective listening environment (cf. figure 1). The sound field occurring in that situation is more likely described as diffuse than free, whereas the free field equalization used as baseline is intended to simulate free field conditions. Therefore, a difference similar to a_D is expected to occur between the level produced by the loudspeaker and the free field equivalent level at the headphones (in this case 65 dB constant for all experiments) in the loudness equality case. Since no loudspeaker equalization was employed and the sound field can not be regarded as perfectly diffuse, similarity but not equality is to be expected.

The good agreement of the results and the shifted attenuation a_D supports the validity of the procedure used (highly significant correlation, r=0.54, 0.4 dB mean deviation of the medians from a_D, and a_D lies within the inter-quartile ranges for 81.8% of all stimuli).

Wave field synthesis situation 1

Figure 4 shows the results for reproduction with wave field synthesis in situation 1 (mean duration 34.3 minutes in two sessions). In addition, the attenuation a_D, according to Fastl and Zwicker (2007, p. 205) shifted upwards by 65 dB is depicted.

The results for wave field synthesis show the unexpected tendency that lower reproduction levels than in the loudspeaker case lead to the same target loudness. This is especially reflected in a by 2.5 dB decreased median over all frequencies (about 61.8 dB SPL), compared to the loudspeaker reproduction. Further, compared to a_D, more irregular frequency characteristics are visible (highly significant correlation, r=0.56, but -1.7 dB mean deviation of the medians and a_D in 72.7% within the inter-quartile ranges).

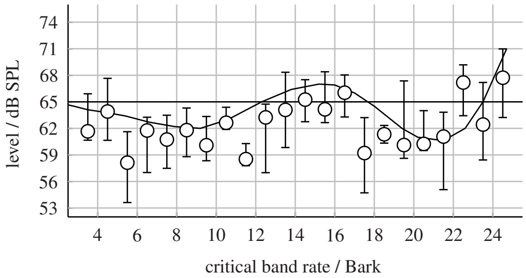

Figure 4: Level of critical band wide narrow band noise (medians as circles with inter-quartile ranges) reproduced by wave field synthesis (situation 1) under reflective listening room conditions, adjusted in loudness to presentation with free field equalized headphones at 65 dB free field equivalent level. The solid line represents the attenuation a_D necessary to produce the same equal loudness curves in a diffuse and free field (ac. to Fastl and Zwicker 2007), shifted upwards by 65 dB.

Wave field synthesis situation 2

Figure 5 shows the results for reproduction with wave field synthesis (situation 2, mean duration 31.1 minutes in two sessions) and the attenuation a_D, shifted upwards by 65 dB.

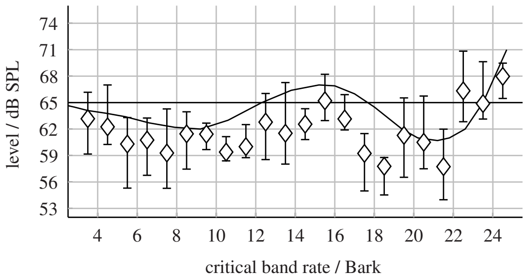

Figure 5: Level of critical band wide narrow band noise (medians as diamonds with inter-quartile ranges) reproduced by wave field synthesis (situation 2) under reflective listening room conditions, adjusted in loudness to presentation with free field equalized headphones. The solid line represents the attenuation a_D necessary to produce the same equal loudness curves in a diffuse and free field, shifted upwards by 65 dB.

The median of the inter-individual medians over all frequencies is computed to 61.4 dB SPL. Comparison to a_D reveals highly significant correlation (r=0.58) with -2.3 dB mean deviation of the medians. In 59.1% of all stimuli, the sifted a_D lies within the results' inter-quartile ranges.

Comparison and discussion

Since the subjects' task remained unchanged for all experiments, it is possible to compare the data for wave field synthesis (figures 4 and 5) directly to those acquired with loudspeaker reproduction (shown in figure 3). All graphics represent the level that, for the considered reproduction method, elicits the same loudness as the free field equalized headphones.

Two factorial analysis of variance (ANOVA) indicates highly significant main effects of the factors presentation method [$F(2,14) = 10.2$, $p = 0.002$] and critical band rate [$F(21,147) = 11.66$, $p < 0.001$]. Further, highly significant interaction between the factors occurs [$F(42,294) = 1.96$, $p < 0.001$]. Post hoc comparison according to Scheffé reveals that the results acquired with wave field synthesis presentation differ for both situations significantly from the results for loudspeaker reproduction. The well known dependence of loudness on frequency is also reflected in the post hoc comparison.

The highly significant interaction between the factors indicates that, in addition to the difference of loudness perception in wave field synthesis from loudness with common loudspeaker reproduction, also the frequency dependence of loudness changes with the different playback conditions discussed in this paper.

Figures 6 and 7 show the differences between the medians of the levels adjusted for loudspeaker and wave field synthesis reproduction in the different situations considered within this paper.

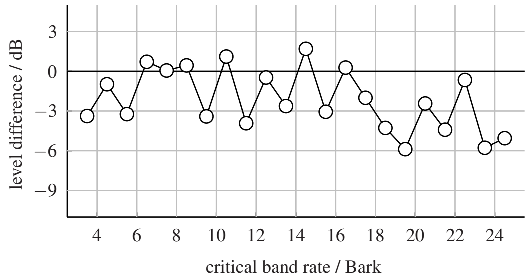

Figure 6: Difference between the medians of the levels adjusted for loudspeaker and wave field synthesis reproduction (situation 1) to 65 dB SPL free field equivalent level.

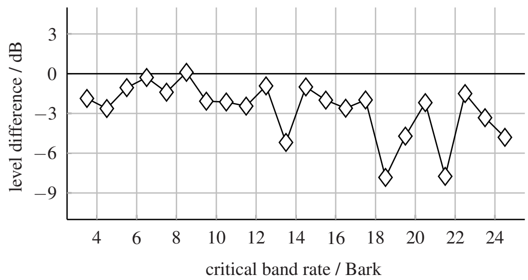

Figure 7: Difference between the medians of the levels adjusted for loudspeaker and wave field synthesis reproduction (situation 2) to 65 dB SPL free field equivalent level.

Having a closer look at the differences occurring, a general tendency for wave field synthesis is visible to elicit the target loudness at lower levels than loudspeaker reproduction or free field equalized headphone listening. This confirms well the authors' impression when listening to wave field synthesis, which basically motivated the work currently presented. A possible reason for this at first unexpected effect may be room influences. Wave field synthesis under anechoic conditions is intended to generate the desired sound field at least at the listening position. This goal can be reached in the frequency range below the spatial aliasing frequency (see next paragraph). Therefore, the same loudness perception as elicited by the wave field to be synthesized is to be expected. Under reflective conditions as considered here, additional contributions to the wave field in the listening area arise, since each secondary source radiates sound intensity not only into the listening area, but also in other directions. Since this undesired intensity is reflected at the listening room boundaries, it propagates back into the listening area from directions different from the primary plane wave's propagation direction. This might result in an increase of the perceived loudness.

Wave field synthesis reproduction causes so called spatial aliasing artifacts (due to the spatial discretization of the loudspeaker array, cf. e. g. Spors and Rabenstein 2006, Wittek et al. 2007, Spors 2006, 2008, Spors and Ahrens 2009a,b, Ahrens and Spors 2009a,c). For the system and setup used for the experiments described here, these artifacts are measurable in the frequency range above about 3.5 kHz (around 17 Bark). Comparison of the results for WFS and loudspeaker reproduction (figures 6 and 7) reveals a slight tendency for the wave field synthesis to reach the target loudness at somewhat lower levels for frequencies higher than about 17 Bark. In other words: aliasing artifacts seem to have no pronounced influence on the loudness of narrow band noise, but show a tendency to further increase the perceived loudness. Figure 8 shows the differences between the levels resulting for the experiments with wave field synthesis in the two situations considered.

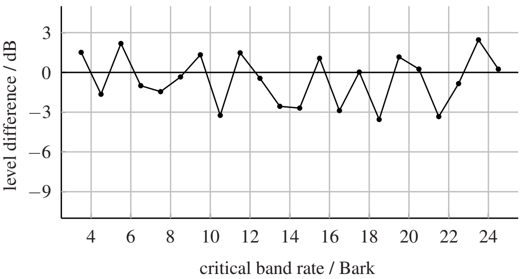

Figure 8: Difference between the medians of the levels adjusted for wave field synthesis reproduction in situations 1 and 2 to 65 dB SPL free field equivalent level.

These differences can be assumed to reflect different listening room influences on the adjusted levels, since the only component that is different in the two WFS situations is the orientation of the WFS setup (including the listener) with respect to the listening room. A detailed comparison of the results' fine structure for the two wave field synthesis situations as given in figure 8 manifests a tendency of the reproduction room influencing loudness judgments at distinct critical band rates. This tendency visible in the fine structure could not be confirmed to be significant by ANOVA.

COMPARISON TO MODEL PREDICTIONS

The aim of the experiments presented in the preceding chapter is to assess loudness perception for different playback conditions. The results given in figures 3, 4, and 5 are sound pressure levels at different critical band rates. The median of the levels at each critical band rate elicits for a typical subject the same loudness in all playback situations considered and in a free sound field, since the subjects' task was to adjust the loudness to that caused by free field equalized headphones. Therefore, common loudness models, for example a so called Zwicker-type model, should be capable of reproducing the experimental results. In other words, the loudness predicted for the narrow band noises at the levels resulting from the listening experiments (N_{ls} or N_{wfs}) are expected for the wave field synthesis and the loudspeaker reproduction cases to be equal to the loudness predictions N_{hp} for narrow band noises at 65 dB SPL in the free sound field. Since the sensation-stimulus relation of loudness is usually measured by relative comparison to the loudness of a reference stimulus, a meaningful way to indicate deviations in loudness is the relative difference in percent. Within this chapter, the relative differences of the resulting loudness predictions are presented and discussed.

Loudspeaker reproduction

Loudness computations according to DIN 45 631 require the user to select the sound field the recordings were taken in from the two options *diffuse* or *free* field. While the loudness computations for the sounds presented by free field equalized headphones without doubt requires the free field mode, the appropriate mode for assessing loudspeaker presentation in the reproduction room under consideration is not completely clear in advance. On the one hand, experience indicates that the free field mode is the appropriate choice for many practical situations (cf. e. g. Fastl et al. 2006). On the other hand, one might argue the listening room where the experiments took place does not exhibit free field characteristics (cf. fig. 1). For that reason, both methods' results are discussed here. In figure 9, the ratio of the instrumental loudness predictions for loudspeaker and headphone reproduction, normalized to N_{hp}, is shown in percent.

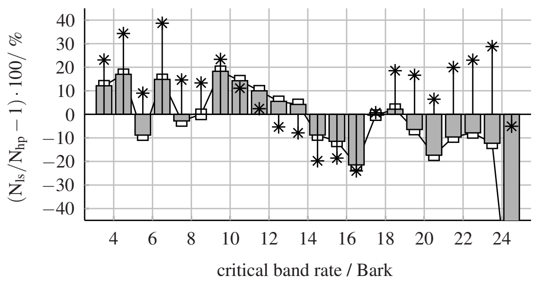

Figure 9: Relative differences between the loudness predictions (ac. to DIN 45 631) for loudspeaker (reflective conditions) and free field equalized headphone playback of critical band wide narrow band noises. 65 dB SPL free field equivalent headphone level, loudspeaker level adjusted to elicit the same loudness. Bars with terminating squares indicate the free field model for the loudspeaker case, stars the diffuse field model.

Stars indicate computation of N_{ls} using the diffuse field mode of DIN 45 631, bars with terminating squares the results acquired using the free field mode. The deviation's median over all critical bands for the diffuse field mode is computed to about 12%, the median of the deviation's absolute magnitude to about 18%. For the free field mode, considerably lower values of -2% and 10% occur. This indicates that the free field mode of DIN 45 631 is better suited to predict the actual listening experiment's results than the diffuse field mode. For that reason, all the following discussions are based on the free field mode of DIN 45 631. This computation is for the stationary sounds considered identical to the free field mode of DIN 45 631/A1. The median deviation of about 2% indicates that on average the instrumental prediction fits the listening experiment's results quite well. Nevertheless, a tendency is visible for the loudness elicited by the loudspeaker to be overestimated at critical band rates below about 14 Bark (about 2.3 kHz) and underestimated for higher frequencies. This might be due to room influences.

Wave field synthesis situation 1

Figure 10 shows for wave field synthesis situation 1 the ratio of the instrumental loudness predictions, normalized to N_{hp}, in percent (N_{wfs} computed using the free field mode of DIN 45 631). The median over the deviations for all critical band rates results to about -10%, the median of the deviations' absolute magnitude to 12%.

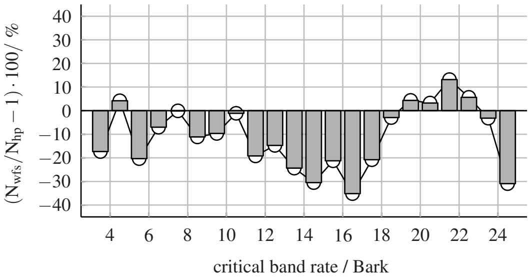

Figure 10: Relative differences between the loudness predictions for WFS (reflective cond., sit. 1) and free field equalized headphone playback of critical band wide narrow band noises.

Wave field synthesis situation 2

In figure 11, for WFS situation 2, the ratio of the loudness predictions, normalized to N_{hp}, (N_{wfs} computed using the free field mode of DIN 45 631) is given in percent. The median over the deviations for all critical band rates results to about -11%, the median of the deviations' absolute magnitude to 17%.

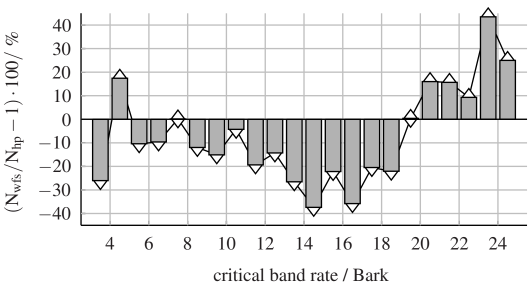

Figure 11: Relative differences between the loudness predictions for WFS (reflective cond., sit. 2) and free field equalized headphone playback of critical band wide narrow band noises.

Discussion

A comparison of the experimental results to the model predictions indicates different model behavior for common loudspeaker and WFS reproduction. While the loudness of critical band wide narrow band noise of different center frequencies presented over a loudspeaker in a reflective listening environment is predicted rather well (about -2% deviation on average), the loudness of the same stimuli presented using WFS is systematically underestimated (median deviation about -10%).

SUMMARY

The results of the experiments presented indicate to some extent considerable differences in loudness perception between

1. a plane wave under anechoic conditions (simulated by free field equalized headphones) and a plane wave as primary source rendered by wave field synthesis without room correction in a reflective listening environment,
2. conventional sound reproduction systems and wave field synthesis,
3. wave field synthesis systems under different listening room conditions.

There is a general tendency visible for the loudness of the wave field synthesis system considered to exceed the loudness of common loudspeaker or headphone reproduction at the same level. This might be due to reproduction room influences. The differences occurring are only partly reflected in predictions by typical loudness models, possibly due to binaural effects.

ACKNOWLEDGMENTS

Parts of this work were supported by grant FA 140/4 of the Deutsche Forschungsgemeinschaft (DFG). The authors acknowledge the kind support of Bose GmbH, Audio AG/RME, and Sound Service GmbH/Samson Technologies Corp. in the setup of the wave field synthesis system.

REFERENCES

DIN 45 619, *Kopfhörer: Bestimmung des Freifeldübertragungsmaßes durch Lautstärkevergleich mit einer fortschreitenden Schallwelle (Headphones: determination of the free-field sensitivity level by loudness comparison with a progressive sound wave), Beuth Verlag, Berlin* (Juni 1975)

DIN 45 631, *Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum. Verfahren nach E. Zwicker (Procedure for calculating loudness level and loudness after E. Zwicker), Beuth Verlag, Berlin* (März 1991)

DIN ISO 226, *Akustik – Normalkurven gleicher Lautstärkepegel (ISO 226:2003) (Acoustics – Normal equal-loudness-level contours (ISO 226:2003)), Beuth Verlag, Berlin* (April 2006)

ANSI S3.4-2007, *Procedure fo the Computation of Loudness of Steady Sounds, American National Standards Institute, Inc.* (May 2007)

DIN 45 631/A1, *Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum – Verfahren nach E. Zwicker – Änderung 1: Berechnung der Lautheit zeitvarianter Geräusche; mit CD-ROM (Calculation of loudness level and loudness from the sound spectrum – Zwicker method – Amendment 1: Calculation of the loudness of time-invariant sound; with CD-ROM, Beuth Verlag, Berlin* (März 2010)

Ahrens J., S. Spors: An Analytical Approach to Sound Field Reproduction Using Circular and Spherical Loudspeaker Distributions. ACTA ACUSTICA UNITED WITH ACUSTICA **94**, 988–999 (2008)

Ahrens J., S. Spors: Alterations of the Temporal Spectrum in High-Resolution Sound Field Reproduction of Different Spatial Bandwidths. In *126th AES Convention* (2009a)

Ahrens J., S. Spors: Sound Field Reproduction Employing Non-Omnidirectional Loudspeakers. In *126th AES Convention* (2009b)

Ahrens J., S. Spors: Timbral Coloration in High Resolution Sound Field Reproduction Due to Spatial Bandwidth Limitation. In *Fortschritte der Akustik, DAGA 2009* (2009c)

Berkhout A. J.: A Holographic Approach to Acoustic Control. J. Audio Eng. Soc. **36**, 977–995 (1988)

Berkhout A. J., D. de Vries, P. Vogel: Acoustic control by wave field synthesis. J. Acoust. Soc. Am. **93**, 2764–2778 (1993)

Betlehem T., T. D. Abhayapala: Theory and design of sound field reproduction in revberberant rooms. J. Acoust. Soc. Am. **117**, 2100–2111 (2005)

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

Boone M. M., E. N. G. Verheijen, P. F. van Tol: Spatial Sound-Field Reproduction by Wave-Field Synthesis. J. Audio Eng. Soc. **43**, 1003–1012 (1995)

Chalupper J., H. Fastl: Dynamic Loudness Model (DLM) for Normal and Hearing-Impaired Listeners. ACTA ACUSTICA UNITED WITH ACUSTICA **88**, 378–386 (2002)

Corteel E.: Equalization in an Extended Area Using Multichannel Inversion and Wave Field Synthesis. J. Audio Eng. Soc. **54**, 1140–1161 (2006)

Corteel E., R. Nicol: Listening room compensation for Wave Field Synthesis. What can be done? In *23rd AES International Conference* (2003)

Corteel E., R. Pellegrini, C. Kuhn-Rahloff: Wave Field Synthesis with increased aliasing frequency. In *124th AES*

Convention (2008) (Convention Paper 7362)

Fastl H., A. Jaroszewski, E. Schorer, E. Zwicker: Equal loudness contours between 100 and 1000 Hz for 30, 50, and 70 phon. ACUSTICA **70**, 197–201 (1990)

Fastl H., H. Fleischer: Freifeldübertragungsmaße verschiedener elektrodynamischer und elektrostatischer Kopfhörer (Free-field transfer functions of different electrodynamic and electrostatic headphones). ACUSTICA **39**, 182–187 (1978)

Fastl H., D. Menzel, W. Maier: Entwicklung und Verifikation eines Lautheits-Thermometers (Design and verification of a loudness thermometer). In *Fortschritte der Akustik, DAGA '06* (2006)

Fastl H., F. Völk, M. Straubinger: Standards for calculating loudness of stationary or time-varying sounds. In *Proc. of Inter-Noise 2009* (2009)

Fastl H., E. Zwicker: *Psychoacoustics - Facts and Models*. 3rd ed. (Springer, Berlin Heidelberg, 2007)

Feldtkeller R.: Über die Zerlegung des Schallspektrums in Frequenzgruppen durch das Gehör (On the decomposition of the sound spectrum into critical bands by the hearing system). Elektron. Rundschau **9**, 387–389 (1955)

Feldtkeller R., E. Zwicker: *Das Ohr als Nachrichtenempfänger (The ear as a communication receiver)* (S. Hirzel Verlag Stuttgart, 1956)

Gauthier P.-A., A. Berry: Adaptive wave field synthesis with independent radiation mode control for active sound field reproduction: Theory. J. Acoust. Soc. Am. **119**, 2721–2737 (2006)

Gauthier P.-A., A. Berry: Adaptive Wave Field Synthesis for Sound Field Reproduction: Theory, Experiments, and Future Perspectives. J. Audio Eng. Soc. **55**, 1107–1124 (2007)

Gauthier P.-A., A. Berry: Adaptive wave field synthesis for active sound field reproduction: Experimental results. J. Acoust. Soc. Am. **123**, 1991–2002 (2008a)

Gauthier P.-A., A. Berry: Adaptive wave field synthesis for broadband active sound field reproduction: Signal processing. J. Acoust. Soc. Am. **123**, 2003–2016 (2008b)

Gelfand S. A.: *Hearing. An introduction to psychological and physiological acoustics*. Second ed. (Marcel Dekker Inc., New York, 1990)

Hulsebos E.: *Auralization using Wave Field Synthesis*, PhD thesis, Technische Universiteit Delft (2004)

Melchior F., C. Sladeczek, D. de Vries, B. Fröhlich: User-dependent optimization of wave field synthesis reproduction for directive sound fields. In *124th AES Convention* (2008) (Convention Paper 7376)

Rennies J., J. L. Verhey, H. Fastl: Comparison of loudness models for time-varying sounds. ACTA ACUSTICA UNITED WITH ACUSTICA **96**, 383–396 (2010)

Spors S., A. Kuntz, R. Rabenstein: An approach to Listening Room compensation with Wave Field Synthesis. In *24th AES International Conference* (2003a)

Spors S., A. Kuntz, R. Rabenstein: Listening Room Compensation for Wave Field Synthesis. In *IEEE International Conference on Multimedia and Expo* (2003b)

Spors S.: *Active Listening Room Compensation for Spatial Sound Reproduction Systems*, PhD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg (2005)

Spors S.: Spatial aliasing artifacts produced by linear loudspeaker arrays used for wave field synthesis. In *Second IEEE-EURASIP International Symposium on Control, Communications, and Signal Processing* (2006)

Spors S.: Extension of an Analytic Secondary Source Selection Criterion for Wave Field Synthesis. In *123rd AES Convention* (2007) (Convention Paper 7299)

Spors S.: Investigation of Spatial Aliasing Artifacts of Wave Field Synthesis in the Temporal Domain. In *Fortschritte der Akustik, DAGA 2008* (Dt. Gesell. für Akustik e. V., Berlin, 2008)

Spors S., J. Ahrens: Spatial Sampling Artifacts of Focused Sources in Wave Field Synthesis. In *Fortschritte der Akustik, DAGA 2009* (Dt. Gesell. für Akustik e. V., Berlin, 2009a)

Spors S., J. Ahrens: Spatial Sampling Artifacts of Wave Field Synthesis for the Reproduction of Virtual Point Sources. In *126th AES Convention* (2009b)

Spors S., H. Buchner, R. Rabenstein: Adaptive Listening Room Compensation for Spatial Audio Systems. In *Proceedings of the 12th European Signal Processing Conference (EUSIPCO)* (2004)

Spors S., H. Buchner, R. Rabenstein, W. Herbordt: Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering. J. Acoust. Soc. Am. **122**, 354–369 (2007)

Spors S., R. Rabenstein: Spatial Aliasing Artifacts Produced by Linear and Cirular Loudspeaker Arrays used for Wave Field Synthesis. In *120th AES Convention* (2006) (Convention Paper 6711)

Spors S., R. Rabenstein, J. Ahrens: The Theory of Wave Field Synthesis Revisited. In *124th AES Convention* (2008) (Convention Paper 7358)

Start E.: *Direct Sound Enhancement by Wave Field Synthesis*, PhD thesis, Technische Universiteit Delft (1997)

Stevens S. S., H. Davis: *Hearing - Its Psychology and Physiology* (John Wiley & Sons, Inc., New York, London, Sydney, 1938) (Seventh Printing, March 1966)

Theile G., H. Wittek: Wave field synthesis: A promising spatial audio rendering concept. Acoust. Sci. & Tech. **25**, 393–399 (2004)

Verheijen E.: *Sound Reproduction by Wave Field Synthesis*, PhD thesis, Technische Universiteit Delft (1997)

Villchur E.: Free-Field Calibration of Earphones. J. Acoust. Soc. Am. **46**, 1527–1534 (1969)

Vogel P.: *Application of Wavefield Synthesis in Room Acoustics*, PhD thesis, Technische Universiteit Delft (1993)

Völk F.: Psychoakustische Experimente zur Distanz mittels Wellenfeldsynthese erzeugter Hörereignisse (Psychoacoustical experiments on the distance of auditory events created by wave field synthesis). In *Fortschritte der Akustik, DAGA 2010* (Dt. Gesell. für Akustik e. V., Berlin, 2010)

Vries D. de: Sound Reinforcement by Wavefield Synthesis: Adaptation of the Synthesis Operator to the Loudspeaker Directivity Characteristics. J. Audio Eng. Soc. **44**, 1120–1131 (1996)

Wittek H.: *Perceptual differences between wavefield synthesis and stereophony*, PhD thesis, University of Surrey (October 2007)

Wittek H., T. Augustin: Räumliche Wahrnehmung von Wellenfeldsynthese: Der Einfluss von Alias-Effekten auf die Klangfarbe. In *Fortschritte der Akustik, DAGA '05*, 199–200 (Dt. Gesell. für Akustik e. V., Berlin, 2005)

Wittek H., F. Rumsey, G. Theile: Perceptual Enhancement of Wavefield Synthesis by Stereophonic Means. J. Audio Eng. Soc. **55**, 723–751 (2007)

Zollner M., E. Zwicker: *Elektroakustik (Electroacoustics)*. Third ed. (Springer, Berlin, 1993)

Zwicker E.: Lautstärke und Lautheit (Loudness level and loudness). In *Proceedings 3rd International Congress on Acoustics*, ed. by L. Cremer (Elsevier Publishing Company, Amsterdam, 1959)

Zwicker E., D. Maiwald: Über das Freifeldübertragungsmaß des Kopfhörers DT 48 (On the free-field transfer function of the headphone DT 48). Acustica **13**, 181–182 (1963)

Zwicker E.: Subdivision of the audible frequency range into critical bands (Frequenzgruppen). J. Acoust. Soc. Am. **33**, 248 (1961)

Zwicker E., R. Feldtkeller: *Das Ohr als Nachrichtenempfänger (The ear as a communication receiver)*. 2nd ed. (S. Hirzel Verlag Stuttgart, 1967)