Scientific Article

Precedence effect for specular and diffuse reflections

Florian Wendt[*] and Robert Höldrich

University of Music and Performing Arts Graz, Austria, Institute of Electronic Music and Acoustics,
Inffeldgasse 10/3, 8010 Graz, Austria

Abstract – Studies on the precedence effect are typically conducted by presenting two identical sounds simulating direct sound and specular reflection. However, when a sound is reflected from irregular surface, it is redirect into many directions resulting in directional and temporal diffusion. This contribution introduces a simulation of Lambertian diffusing reflections. The perceptual influences of diffusion are studied in a listening experiment; echo thresholds and masked thresholds of specular and diffuse reflections are measured. Results show that diffusion makes the reflections more easily detectable than specular reflections of the same total energy. Indications are found that this mainly due to temporal diffusion, while the directional diffusion has little effect. Accordingly, the modeling of the echo thresholds is achieved by a temporal alignment of the experimental data based on the energy centroid of reflection responses. For the modeling of masked threshold the temporal masking pattern for forward masking is taken into account.

1 Introduction

The precedence effect refers to a group of perceptual phenomena which allow us to localize sound sources in challenging reverberant environments. Research has shown that the auditory system differently weights the spatial cues over the stimulus duration; the sound source direction is determined mainly by the onset of the stimulus, where localization cues of the direct sound prevail. Corresponding studies typically utilize a two-source paradigm, representing a leading direct sound and a lagging specular reflection at the same intensity. The parameter of interest for this setup is the lead-lag delay. For very short delays between direct sound and reflection below 1 ms, summing localization occurs and the two sounds appear as a single fused image located between the sound instances. As the delay slightly exceeds the limits for summing localization the precedence effect becomes active. In this range the auditory system suppresses localization cues carried by reflections and the leading direct sound dominates the perceived location of a single auditory image. Another increase of the delay yields widening of the auditory image until it breaks apart at the upper limit of echo suppression. At this transition, the so-called *echo threshold* the reflection becomes audible as a second auditory image. This threshold is especially relevant because it can be used to determine the strength of

precedence; increased echo threshold delays imply a stronger echo suppression and consequently a stronger precedence effect.

Many investigations have contributed to our understanding of the precedence effect by measuring the echo threshold with reflections that are identical to the direct sound (see [1, 2] for reviews). However, such specular reflections represent rather artificial conditions, which require an infinitely large, flat, rigid reflecting surface [3]. When the sound encounters a real-life wall, it is inevitably redirected also into angles other than the specular reflection angle. There are only a handful of contributions which studied the precedence effect in more realistic scenarios using wall panels for creating the reflection, e.g. [4–6].

This contribution examines how reflection properties of a wall influence the echo suppression for pulsed noise signals. First, an overview on the relevant literature is given in Section 2. Subsequently, Section 3 presents a simulation for diffuse reflection responses based Lambertian reflection. The simulation considers both temporal and directional diffusion of the reflected sound by incorporating the scattering coefficient. The perceptual influences of diffuse reflections are studied in a listening experiment. Section 4 outlines the experimental design for measuring the echo threshold and the masked threshold. The experimental results are discussed in Section 5. Modeling attempts of the results of both thresholds are presented in Section 6. Section 7 relates the findings to previous studies and Section 8 summarizes the contribution.

*Corresponding author: wendt@iem.at

2 Echo suppression for non-specular reflections

Specular reflections represent one of two extreme conditions which have been identified for rigid surfaces. When the direct sound of a source is reflected specularly, both magnitude spectrum and phase spectrum are maintained and the sound bounces off the surface with the same angle as it encountered the surface. The other extreme condition is the *diffuse* reflection that occurs when the reflected energy is scattered [3, 7]. It yields a reflected sound that spreads directionally and temporally.

In room acoustics, reflections on rigid walls are typically approximated by dividing the reflected energy into two components, specular and diffuse, with the scattering coefficient s defining the ratio of the scattered energy to the total energy reflected by the surface. Thus, the scattering coefficient is defined in the interval $s = [0, 1]$ and a value of 1 corresponds to a fully diffuse reflection with no specular reflection component.

While many psychoacoustic experiments have contributed to our understanding of the precedence effect with specular reflections, knowledge on the influence of spatio-temporal diffuse reflections is still rough and partly contradictory. The temporal diffusion for example yields spectral colorations of the reflection. According to Blauert and Dinveyi [8] suppression occurs for any natural reflection where the spectrum of the lag contains exclusively those regions that are present in the spectrum of the lead. However, for reflections differing in magnitude spectrum and phase spectrum, the suppression was found to be generally weaker compared to a specular reflection, e.g. [8–10].

Robinson et al. [11] studied the influence of temporal diffusion on the echo threshold delay by modeling visual observations of measured reflection responses from diffusors. Their simulation approximates the impulse response of diffuse reflections by multiplying Gaussian white noise with the envelope describing the probability density function of a gamma distribution. Resulting echo threshold delays, defined by the temporal energy centroid of impulse responses, revealed mostly no difference between the suppression of compact and a temporal diffuse reflections.

Grosse et al. [12] studied the perceptual influence of directional diffusion of either direct sound or reflection by presenting the sound from a single loudspeaker or by presenting mutually uncorrelated versions of similar sounds from nine adjacent loudspeakers. Although no pure specular lead/lag condition was tested in their experiments, there is evidence that the directional diffusion of the reflection might reinforce the echo suppression and for both speech and noise bursts, the echo threshold delay is longer if diffuse reflections are used.

Visentin et al. [13] recently studied the influence of an early reflection on spatial attributes by employing real flat and diffusive wall panels. They reported that the presence of a single diffuse reflection reduces the perceived distance of a frontal speech source, makes it clearer, and increases its intelligibility.

Lokki et al. [14] studied the qualitative influence of the temporal diffusion of reflections and found significant perceptual differences. In contradiction to [12] and [13], they assumed that the clear and open acoustics reported for specular reflections is due to stronger precedence, compared to temporal diffuse reflections that render the sound weak and muddy.

3 Simulating diffuse reflections

A *specular* reflection occurs when a plane wave is reflected on an infinitely large, smooth, and rigid wall. Snell's law for the specular reflection states that angle of reflection is equal to the angle of incidence. The impulse response of such a reflection is described by a Dirac delta distribution $\delta(t - T_r)$ with the delay T_r calculated from the distance between the corresponding image source and the receiver.

For simulating a diffuse reflection, suppose the point source S in Figure 1 emits a short power impulse at time $t = 0$, represented by a Dirac delta function $\delta(t)$ with unity energy. The sound power $\mathrm{d}P_W(t)$ reaching the wall element $\mathrm{d}W$ by the angle of incidence θ_S and distance r_S is then defined by,

$$\mathrm{d}P_W(t) = \frac{\cos \theta_S}{4\pi \, r_S^2} \, \delta(t - t_S) \, \mathrm{d}W, \tag{1}$$

with $r_S = \sqrt{x^2 + z^2 + d_S^2}$, time $t_S = r_S/c$, and speed of sound c. Diffuse reflections are typically interpreted by assuming Lambertian reflection in which the scattered sound power from the element $\mathrm{d}W$ is proportional to the cosine of the angle of reflection $\cos \theta_R$ [3]. The intensity $\mathrm{d}I_W$ reflected by the wall element $\mathrm{d}W$ of a perfectly reflective wall without absorption that reaches the receiver R is defined by,

$$\mathrm{d}I_W(t) = \frac{\cos \theta_S \, \cos \theta_R}{4\pi \, r_S^2 \, r_R^2} \, \delta(t - t_{SR}) \mathrm{d}W, \tag{2}$$

with $r_R = \sqrt{x^2 + z^2 + d_R^2}$ and $t_{SR} = (r_S + r_R)/c$. Assuming the reflected sounds of wall elements $\mathrm{d}W$ to be incoherent, the overall intensity I_{diff} reaching the receiver R is obtained by the integration over both dimensions of the wall,
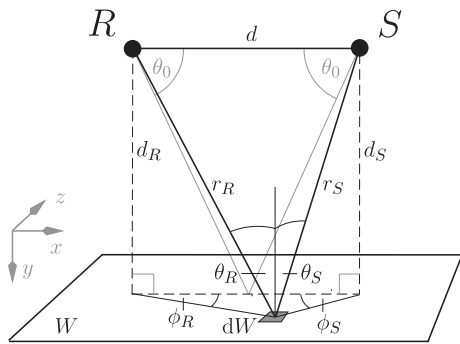
$$I_{\mathrm{diff}}(t) = \iint\limits_{W} \frac{\cos \theta_S \, \cos \theta_R}{4\pi \, r_S^2 \, r_R^2} \, \delta(t - t_{SR}) \, \mathrm{d}x \mathrm{d}z. \tag{3}$$

The two cosines can be replaced by distance relations $\cos \theta_S = d_S/r_S$ and $\cos \theta_R = d_R/r_R$, yielding the elliptical integral,
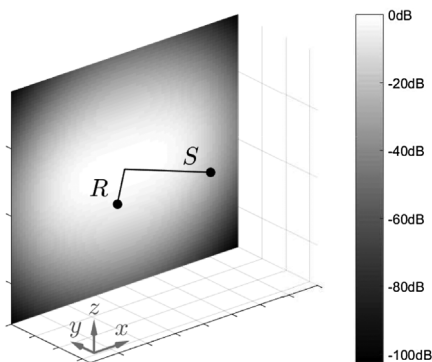
$$I_{\mathrm{diff}}(t) = \iint\limits_{W} \frac{d_S \, d_R}{4\pi \, r_S^3(x,z) \, r_R^3(x,z)} \, \delta(t - t_{SR}(x,z)) \, \mathrm{d}x \mathrm{d}z, \tag{4}$$

which we solve numerically.

Figure 2 shows the reflected intensities $\mathrm{d}I_W$ reaching the receiver R and directionally spreading around the specular
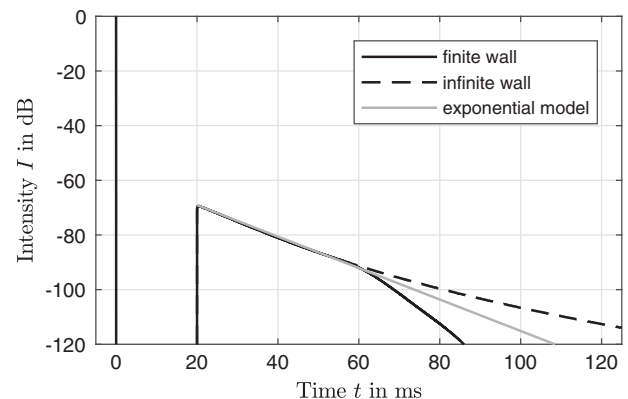
Figure 1. Schematic geometry of a reflection on the xz-plane. The basic constellation consists of a source S, a receiver R, and a wall W.

Figure 2. Directional spread of diffusely reflected intensities dI_W normalized to the maximum and coded in gray-scale. The sound propagation path of a specular reflection is indicated as solid line.

Figure 3. Intensity envelopes of the impulse response of condition A_{20} consisting of direct sound an diffuse reflection. The temporal spreads for diffusely reflected intensities I_{diff} for an infinite wall and the finite-sized wall W are normalized to the direct sound. The exponential model proposed in [15] is fitted to the early decay.

direction. It is easy to see that the size of the wall influences the amount of the reflected energy. The overall intensity I_{diff} for two different wall sizes is depicted in Figure 3 (setup A, cf. Table 1) and represents the temporal spread. The dimension of the rectangular finite-sized wall is chosen in a way that compared to an infinite wall approximately 90% of the energy is reflected. For a better comparability, intensity envelopes are shifted in time by $T_d = d/c$. In this way, the direct sound of all conditions reaches the receiver at $t = 0$ ms, whereas corresponding reflections start at $\Delta T = T_r - T_d = 20$ ms.

In [15] the temporal structure of diffuse reflections is modeled based on the rough surface theory [16, 17]. The energy envelope of their simulated diffuse reflection is an exponential function $e^{-t/\tau}$ with the decay coefficient τ found by fitting a line to the Schroeder-integrated energy curve corresponding to the measured impulse response. Fitting exponential functions to our simulation with decay coefficients based on the early decay yields an envelope decay of $e^{-t/\tau}$ with $\tau = \Delta T/1.33$. Figure 3 shows a neat overlapping of the early parts of the reflective impulse responses and the exponential functions comprise about 95% of the overall energy reflected by an infinite wall.

The simulated reflection response of our diffusely-reflecting wall has an instant onset at ΔT followed by an exponential decay. Although its shape complies with other simulations, it is different from measured impulse responses of conventional diffusers. The impulse response of a Schroeder diffuser [18] for example exhibits a damped onset and reaches its peak amplitude several milliseconds later, e. g. [19]. This is because of its specific surface structure, which in addition to the diffusion influences also the delay ΔT as the specular path is not necessarily the shortest. The aim of this study is to reveal perceptual differences between specular and diffuse reflections rather than the simulation of a real-world diffuser, which is why it considers only reflections simulated with a plane wall where delay times ΔT are well defined.

4 Experimental setup

In a constellation of sound source, receiver, and reflective wall the distance traveled by the reflection determines not just the delay ΔT, but also the inverse-square law intensity difference ΔL of the reflection compared to the direct sound. As these parameters are closely linked together, alternatively to the lead-lag delay ΔT, some studies define the echo threshold as level difference ΔL between leading direct sound and lagging specular reflection, e.g. [20, 21]. Thus, a stronger echo suppression is implied by a longer echo threshold delay or a higher echo threshold level.

In contrast to specular refections, the intensity envelope of a diffuse reflection response obtained by our simulation varies with the setup and thus also with the delay. Therefore, varying the delay alone is not straightforward, which is why the experiment examines the reflection's level for a fixed delay. The echo threshold level is defined as the minimum level ΔL_E of the lagging sound compared to the direct sound at which it is possible to detect a second auditory event. Even after the level of a lagging sound is

Table 1. Setup parameters of the conditions under investigation.

| Condition	θ_0 / $^\circ$	θ_W / $^\circ$	ΔT / ms	d / m	d_R / m	W / m^2	w / m^2
A_5	45	90	5	4.1	2.0	10×8.5	0.01
A_{10}	45	90	10	8.3	4.1	20×17	0.04
A_{20}	45	90	20	16.5	8.3	40×34	0.16
A_{30}	45	90	30	24.8	16.6	60×51	0.64
B_{20}	68	90	20	4.0	5.1	20×18	0.04
C_{20}	0	0	20	4.0	7.4	18×18	0.04

reduced to the point where an echo is no longer perceptible, the presence of the reflection is still audible. In addition to the echo threshold level, the experiment determines the masked threshold level defined as the minimum level ΔL_{M} at which it is possible to detect that a lagging sound is present at all.

4.1 Stimuli and procedure

The excitation signal for all conditions consisted of 50-ms Gaussian noise bursts with instant onset and offset.

The echo threshold was measured by a method of adjustment [21]. The actual stimulus was looped with 200-ms-long silence between each repetition. The listener was given control over the level of the reflection with the instruction to reduce its level down to the point where it is as faint as possible while still remaining audible as a second auditory event. To ensure that the reflection was audible as an echo at the beginning of each trial, the level of the lag was 10 dB above that of the direct sound. Since the echo suppression requires a buildup time to become fully effective [22], the amplitude of the ongoing sequence (consisting of direct sound and reflection) was linearly faded in by 2 s at the beginning of each trial. This time constant was determined by the authors by informal listening. Control of the level of the lag was enabled after this time.

The masked threshold was measured by a two-alternative forced choice adaptive procedure. Two intervals with a pause of 1 s between the intervals were presented consecutively. The direct sound of each interval consisted of four noise bursts with 200-ms-long silence between the bursts. In one randomly chosen interval the reflection was added, whereas in the other interval it was absent. The listeners' task was to specify the interval containing the reflection. Possible loudness cues were removed by roving the level of both intervals, and feedback was given after each response. During a run, the level of the reflection was adjusted in a 3-down 1-up rule [23] estimating the 79.7% point on the psychometric function. At the beginning of each run direct sound and reflection were equally loud with a step size of 10 dB. After two reversals the step size was decreased to 5 dB and set to a final value of 2 dB after another two reversals. Using the 2 dB step size 6 more reversals were obtained, and according to [24] the threshold value is calculated by averaging over all levels obtained by final step size.
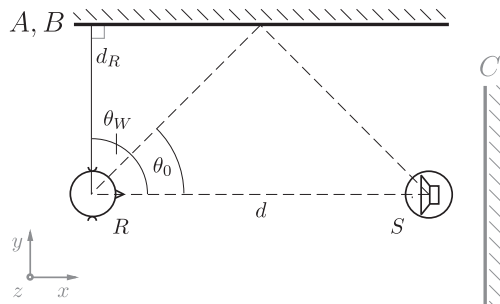
Figure 4. Schematic representation of the constellation of the reflective wall for setups that were examined in the experiment. For lateral-wall setups A and B (black) sound paths of direct sound and specular reflection, and specular reflection angle θ_0 and wall angle θ_W are indicated. For setup C (gray) both direct sound and reflection are perpendicular to the wall.

4.2 Conditions

Directional and temporal distribution of diffuse reflections vary with the setup of the geometric constellation, i.e., source-to-receiver distance and respective distances to the wall. In the experiment three different setups A, B, C were tested; Figure 4 schematically shows the lateral wall simulated with setup A and B, and the frontal wall of setup C. Respective conditions derived from the setups are listed in Table 1 with the receiver R facing always the point source S at equal height z. Note that the conditions' subscript numbers indicate the corresponding delay ΔT in ms and the superscripts indicate the scattering coefficient s. To study the influence of the delay ΔT on specular ($s = 0$) and diffuse ($s = 1$) reflections while keeping the angle $\theta_0 = 45^\circ$ constant, conditions $A^0_{\Delta T}$ and $A^1_{\Delta T}$ with $\Delta T = (5, 10, 20, 30)$ ms were tested for both suppression thresholds. The influence of the angle of reflection θ_0 on suppression thresholds was tested by conditions A^1, B^1, C^1 while keeping the delay $\Delta T = 20$ ms constant. Moreover, to study the influence of the scattering coefficient, condition A^s_{20} with $s = 0.5$ was included in the test. The condition $A^\star_{20}$ is a hybrid of specular and diffuse reflection, and examined the influence of the directional spread of the diffuse reflection. Compared to fully-diffuse A^1_{20}, the diffuseness of condition $A^\star_{20}$ considers only the monaural-temporal spread but not the directional spread. Instead, similar to A^0_{20}, the reflection was played back only from the (specular) direction θ_0.

For the numerical solution of Equation (3) the size of the wall W was chosen in a way that approximately 90% of the energy reflected by an infinite wall is taken into account by a uniform grid consisting of quadratic patches w. Computed intensities dI_W were encoded into Ambisonics (order $N = 17$) and decoded on a spherical t-design (degree $t = 35$) using max-r_E weighting [25] resulting in 632 envelope signals describing the energy distribution of a diffuse reflection. The corresponding impulse responses were obtained by multiplying white Gaussian noise with the square roots of intensity envelopes. An iterative whitening procedure similar to the Hilbert transform approach described by [26] was applied to each reflection waveform to avoid coloration of the diffusely reflected sound. Finally, impulse responses were normalized to the total energy. Impulse responses of specular reflections were represented by a Dirac delta distribution, and thus specular reflections are an exact copy of the direct sound.

All testing was conducted over headphones. The excitation signal (i.e., the noise burst) was convolved with respective impulse responses and binaural stimuli were created by convolving them with corresponding HRTF measurements[1] of the Neumann KU100 dummy head. For constructing the stimuli Matlab was used. Playback employed PureData and the output was fed through a M-Audio MobilePre sound card to Beyerdynamics DT770 headphones. The playback level of the direct sound was fixed at 70 dB(A).

To prevent listeners fatigue, the listening experiment was performed in three runs. In the first run all echo conditions $A^1_{(5,10,20,30)} + A^0_{(5,10,20,30)} + A^{0.5}_{20} + B^1_{20} + C^1_{20} + A^{\star}_{20} = 12$ were tested twice yielding 24 adjustment tasks. The second run included masked conditions with $\Delta T = 20$ ms, which were tested once yielding $A^0_{20} + A^{0.5}_{20} + A^1_{20} + B^1_{20} + C^1_{20} + A^{\star}_{20} = 5$ adaptive tasks. The last run consisted of masked conditions for specular and diffuse reflections $A^0_{(5,10,20,30)} + A^1_{(5,10,20,30)} = 8$. In this way masked conditions A^0_{20} and A^1_{20} were tested twice, once in the context of the refection properties (second run), and once in the context of the delay ΔT (third run). Conditions within each run were tested in individual random order.

5 Experimental results

Twelve listeners (age 26–55 years) participated in the experiment. All of them were experienced listeners and reported normal hearing acuity.

5.1 Influence of scattering

Figure 5 shows echo threshold levels ΔL_E and masked threshold levels ΔL_M examined with setup A as a function of the delay ΔT. Unsurprisingly, masked threshold levels are below echo threshold levels with differences agreeing with literature [21]. Both levels decrease progressively with increasing ΔT with a progression of means resembling those obtained by similar studies, e.g. [27]. The size of
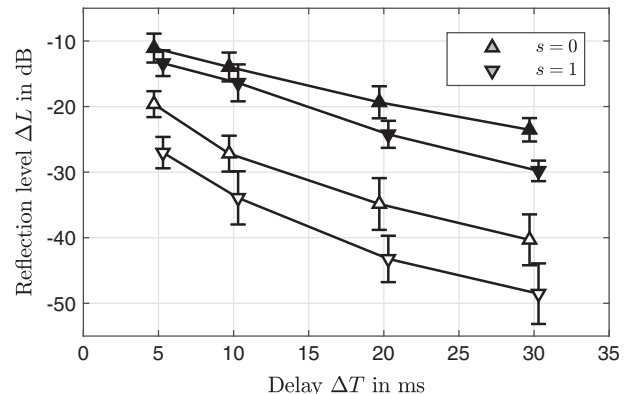
Figure 5. Means and 95% confidence intervals of echo threshold levels (filled symbols) and masked threshold levels (open symbols) for specular reflections ($s = 0$) and diffuse reflections ($s = 1$) over delays ΔT tested with setup A.

corresponding 95% confidence intervals provides evidence that listeners performed similarly for both reflection types of each experiment. A reason for the overall smaller confidence intervals of the echo threshold is due the fact that corresponding conditions were tested twice.

Statistics using the two-way repeated measures analysis of variance (RM ANOVA) reveal the threshold type, the delay, and the reflection type to be significant parameters ($p \leq 0.05$). Interestingly, diffusion weakens the echo suppression and diffuse reflections are more easily detectable. For all conditions the diffuse threshold level is below the corresponding specular threshold level. A Tukey HSD post hoc analysis of both reflections types shows conditions $A_{(20,30)}$ of the echo threshold and all conditions of the masked threshold to be significantly different ($p \leq 0.05$) with effect sizes expressed as Cohen's $d \geq 0.9$, i.e., *large* effects [28].

In addition to specular ($s = 0$) and diffuse ($s = 1$) reflections, condition A_{20} was tested with a scattering coefficient of $s = 0.5$. Respective results are given in Figure 6. For both suppression types, mean levels of $A^{0.5}$ are in between corresponding means of extreme conditions A^0 and A^1 and a two-way RM ANOVA reveals the scattering coefficient to be a significant parameter ($p \leq 0.05$). Note that masked threshold levels, obtained in the second run, are consistently higher than respective levels from the third run given in Figure 5 ($p > 0.05$).

5.2 Influence of directional spread and spatial separation

The influence of directional spread of the diffuse reflection on the echo suppression is examined with the hybrid condition $A^{\star}$, cf. Figure 6. This condition combines the monaural-temporal characteristics of a diffuse reflection with the binaural-temporal characteristics of a specular reflection. Mean thresholds levels are below mean levels of the corresponding directionally-spreading condition A^1 and paired sample t-tests reveal significant differences of $A^{\star}$ and A^1 at least for the echo threshold ($p \leq 0.05$;
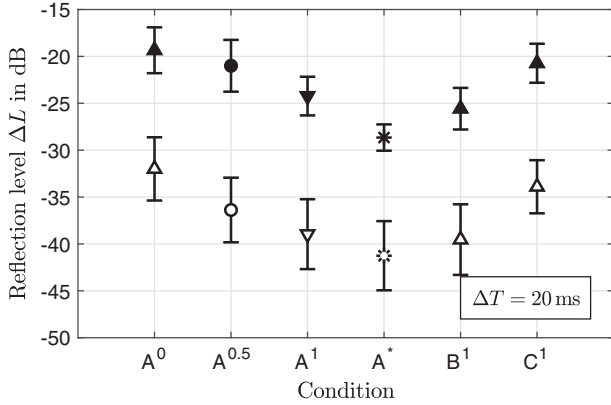
Figure 6. Results of the echo threshold level (filled symbols) and the masked threshold level (open symbols) given as means and 95% confidence intervals. The influence of the scattering coefficient s is examined with conditions A^s, the influence of the angle θ_0 is examined with conditions A^1, B^1, C^1, and the influence of the directional spread is examined with the hybrid condition $A^\star$, which combines the temporal characteristics of a diffuse reflection with the directional characteristics of a specular reflection. For all conditions the delay is constant with $\Delta T = 20$ ms.

Cohen's $d = 0.99$, *large* effect [28]). It thus can be concluded that the temporal and directional spread have an opposite effect on the perception; compared to fully diffuse reflections, reflections spreading only temporally but not directionally are more easily detectable and thus weaken the echo suppression. Conversely, this means that the directional diffusion increases the suppression, making reflections more difficult detectable.

For specular reflections studies could show that the suppression is higher if direct sound and reflection arise from similar directions than when they are spatially separated [11, 27, 29]. The influence of spatial separation is examined by diffuse conditions A_{20}, B_{20}, C_{20}. Echo and masked threshold levels of condition B_{20} are not different from A_{20} ($p > 0.05$, Tukey HSD post hoc analysis) suggesting that the influence of spatial separation is not applicable for diffuse reflections that spread directionally and temporally. However, corresponding reflection responses depicted in Figure 7 reveal a higher temporal spread of condition A_{20}, which might compensate any effect of spatial separation. In contrast, conditions B_{20} and C_{20} exhibit a similar temporal spread, and significant increases of echo and masked thresholds levels are seen for the decreased spatial separation of direct sound and reflection of condition C_{20} ($p \leq 0.05$, Tukey HSD post hoc analysis; Cohen's $d \geq 1.0$).

6 Modeling the results

6.1 Echo threshold

Following the approach of Rakerd et al. [21] a linear model is used to describe the echo threshold level ΔL_E as a function of the delay ΔT,
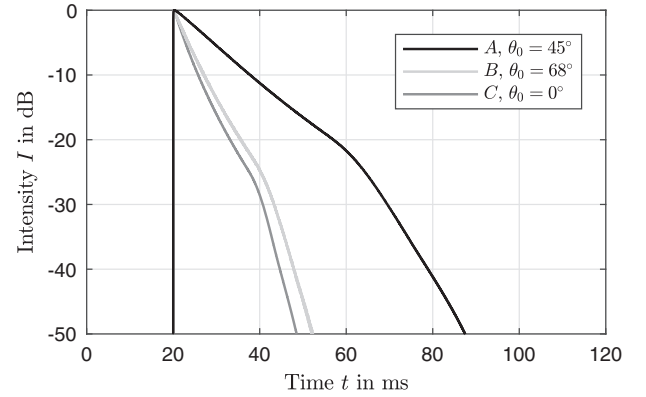
Figure 7. Diffusely reflected intensity I_{diff} arriving at the receiver R of conditions A^1, B^1, and C^1 with the finite sized wall W, normalized to their maximums and shifted in time by T_d.

$$\Delta L_E = \alpha + \beta \cdot \Delta T, \tag{5}$$

with the intercept α in dB and the slope β in dB/ms. Modeling individual echo thresholds with the regression fit given in Equation (5) reveals the differences of obtained slopes of both reflection types to be significantly different form zero (t-test: $p \leq 0.05$). In other words, the slope of the specular regression line is different from the slope of the diffuse regression line. The differences of individual intercepts of specular and diffuse threshold on the other hand are not significant ($p > 0.05$), and we conjecture that the temporal spread of diffuse reflections results in an increase of the effective delay.
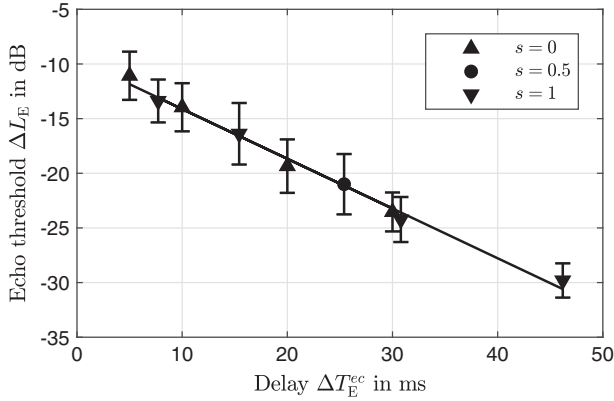
To account for different scattering coefficients s in a single model, we introduce a simple predictor of the effective delay ΔT_E replacing ΔT in Equation (5) by,

$$\Delta T_E = \Delta T \left(1 + s \cdot k_E\right). \tag{6}$$

This yields a temporal alignment of the results with the alignment parameter k_E. The optimal alignment parameter is obtained by pooling the data of setup A for $s = (0, 0.5, 1)$ with corresponding effective delays of Equation (6), and we choose the alignment parameter k_E to maximize the coefficient of determination of the model in Equation (5). The optimal parameter is found at $k_E^{\text{opt}} = 0.48$ yielding $\alpha = -9.5$ dB and $\beta = -0.47$ dB/ms with a coefficient of determination $R^2 = 0.99$ of the regression with corresponding means (95% confidence intervals: $k_E = [0.38, 0.55]$, $\alpha = [-10.8, -8.1]$ dB; $\beta = [-0.53, -0.41]$ dB/ms).

In [11] the echo threshold of specular and temporally diffuse reflections is examined by a variation of the delay. Their definition of the delay, i.e., the time between the direct sound and the centroid of energy of the reflection response, did not yield any perceptual differences between specular and temporally diffuse reflections for most conditions (5 out of 6) with speech and music signals. Thus, their definition of the delay could serve as model to predict the echo threshold of conditions with arbitrary scattering coefficients.

Delays ΔT_E^{ec} based on the centroid of energy of diffuse reflections of setup A are calculated using Equation (6)

Figure 8. Means and 95% confidence intervals of echo threshold levels ΔL_{E} for different scattering coefficients s of setup A. The regression line is fitted with Equation (5) and delays $\Delta T_{\mathrm{E}}^{ec}$ are calculated from the temporal energy centroid of corresponding reflection responses using Equation (6) with $k_{\mathrm{E}}^{ec} = 0.54$.

and we obtain $k_{\mathrm{E}}^{ec} = 0.54$. This value falls within the 95% confidence interval of the optimal parameter $k_{\mathrm{E}}^{\mathrm{opt}}$. Figure 8 shows the temporally aligned data with $k_{E}^{ec} = 0.54$ and the corresponding regression line, which highly correlates with mean echo thresholds ($R^2 = 0.99$, $\alpha = -9.6$ dB, $\beta = -0.45$ dB/ms with 95% confidence intervals $\alpha = [-10.7, -8.5]$ dB/ms, $\beta = [-0.50, -0.41]$ dB/ms).

6.2 Masked threshold

The temporal masking effect can be interpreted as a mixture of *simultaneous masking* and *forward masking*. Fastl and Zwicker [30] examined simultaneous masking of noise bursts by uniform masking noise, and found threshold levels that were not more than 10 dB below the masker level. In our experiments thresholds are sharply lower than the masker (i.e., the direct sound), and we assume that masked levels are mainly caused by forward masking. Following temporal masking theory a signal is masked as long as its excitation pattern is below the temporal masking pattern evoked by the masker. According to [30] (p. 83, Fig. 4.22) temporal masking patterns for forward masking of noise bursts do not exhibit an exponential decay, but thresholds can be well approximated by a logarithmic function of the delay time, i.e., $\Delta L_{\mathrm{M}} = \gamma + \epsilon \cdot \ln(\Delta T / \mathrm{ms})$ with $\epsilon < 0$ for $\Delta T = 10, \ldots, 100$ ms.

The results of our experiments indicate that diffuse reflections linearly lower the masked threshold ΔL_{M} by a factor δL in dB compared to specular reflections, cf. Figure 5. Thus, in order to model masked thresholds of arbitrary reflections, we extend the logarithmic function by the scattering coefficient yielding,

$$\Delta L_{\mathrm{M}} = \gamma + s \cdot \delta L + \epsilon \cdot \ln(\Delta T / \mathrm{ms}). \qquad (7)$$

Figure 9 shows the forward masking function from [30] together with envelopes of specular and diffuse reflections. For a specular condition, the reflection amplitude exceeds the temporal masking pattern for the first time at the
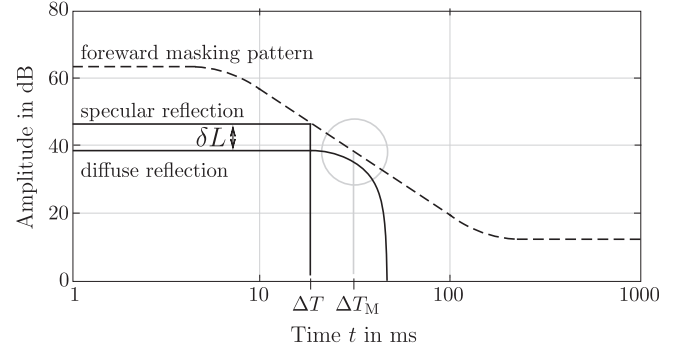
Figure 9. Schematic representation of the forward masking pattern (dashed) from [30] (p. 83, Fig. 4.22) as a function of the delay time $\log(\Delta T)$ together with nominal amplitudes of specular and diffuse reflection (solid) as a function of time t. The masker (direct sound) ends at $t = 0$ ms. The masked level is achieved, if reflection's envelope peaks exceed the post masking pattern. For the specular reflection this when the reflection ends at ΔT, whereas envelope fluctuations of the diffuse reflection exceed the level of masking in the time range of ΔT_{M}.
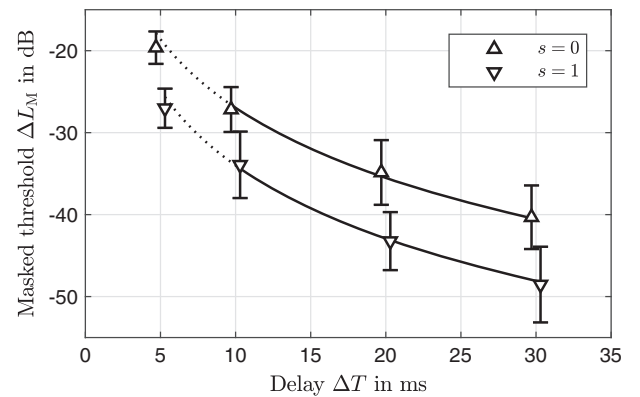
Figure 10. Means and 95% confidence intervals of masked threshold levels for different scattering coefficients of setup A. The regression lines are fitted in the interval 10 ms $\leq \Delta T \leq 30$ ms using Equation (7).

instant offset $t = \Delta T$. For a diffuse condition on the other hand, the decay of the reflection amplitude is flatter than the logarithmic function given in Equation (7). Accordingly, peaks of the reflection envelope exceed the masking pattern at ΔT_{M} at a lower level, cf. Figure 9. Fitting parameters for specular and diffuse masked levels of conditions $A_{(10,20,30)}$ are calculated using Equation (7) and are $\gamma = 2.4$ dB, $\epsilon = -12.6$ dB, and $\delta L = -7.8$ dB (corresponding 95% confidence intervals: $\gamma = [-6.7, 11.5]$ dB; $\epsilon = [-15.7, -9.5]$ dB; $\delta L = [-14.0, -3.3]$ dB). This yields a coefficient of determination $R^2 = 0.99$ with mean masked thresholds.

Figure 10 shows masked threshold levels of specular and diffuse conditions of setup A with the corresponding logarithmic fit. For $\Delta T < 10$ ms the forward masking function given in Figure 9 saturates. Accordingly, thresholds of conditions A_5^0 and A_5^1 are overestimated by the logarithmic model of Equation (7).

In contrast to our logarithmic model, in [21] a linear fit is used for masked threshold levels with delays $20\,\mathrm{ms} \le \Delta T \le 80\,\mathrm{ms}$. Similarly, for our results in this range a linear model would be sufficient. For short delays $\Delta T < 20\,\mathrm{ms}$, however, Figure 10 clearly shows the advantage of the logarithmic approach.

7 Discussion

The results of the experiment indicate that diffusion makes reflections more easily perceivable as masked threshold levels are below the corresponding levels of specular reflections. Accordingly, diffuse reflections similarly weaken the precedence effect and less level is needed to perceive a diffuse echo compared to specular echoes. This finding agrees with the observations made in [14] but contradicts the findings from [13], where reflections form a Schroeder diffuser were reported to be more focused compared to specular reflections of the same energy. As their diffuser had a higher absorption coefficient in the frequency range 250 kHz–1 kHz, it is conceivable that it altered the spectrum of the scattered speech signal in such a way that it was less distractive.

By removing the spatial spread of our diffuse reflections, the suppression further decreases and obtained thresholds levels of temporally diffuse refections are below corresponding levels of fully diffuse conditions. Thus, we conclude that directional and temporal diffuseness have an opposite effect on the echo suppression. This finding dissolves the apparent contradiction of literature. In agreement to [12], directional diffusion increases the suppression. However, if reflections spread also temporally this effect vanishes, cf. [14].

The effect of spatial separation between direct sound and reflection, comprehensively investigated with specular reflections in [2, 27], and approved for temporally diffuse reflections in [11], applies also to fully diffuse reflections and the echo suppression is stronger when direct sound and reflection are close together.

In contrast to our results, in [11] no differences between specular and temporally diffuse reflections where obtained in most conditions. However, their definition of the delay between direct sound and reflection differs from ours as it is derived from the temporal centroid of energy of corresponding impulse responses. This finding is used to establish a linear model for the echo threshold that allows the prediction of arbitrary reflections by aligning their echo thresholds using temporal energy centroids of corresponding reflection responses.

8 Conclusions

Diffuse reflections were simulated based on Lambert's cosine law. The simulated reflections yield a directional and temporal smearing of the reflected sound field, which is influenced by the geometrical setup of sound source, receiver, and reflective wall. The early decay of the temporal spread resembles an exponential function and agrees with other simulations found in the available literature. The echo threshold and the masked threshold were measured in a listening experiment to investigate the influence of a diffuse reflection on the precedence effect. The main findings of the study can be summarized:

1. The comparison of both reflection types reveals the echo suppression to be weaker for diffuse reflections than for specular reflections of the same total energy. In other words, if the reflected sound is scattered, less level is necessary to hear it as an echo and less level to notice that a reflection is present at all.
2. The weakening of the precedence effect is mainly due to the temporal diffusion. Spatial diffusion has a reverse effect and increases the suppression. However, as the latter hypothesis is based on the results from a single condition, it will need a more comprehensive validation.
3. The spatial release from echo suppression was determined for diffuse reflections and echo threshold levels are lower if direct sound and diffuse reflection are directionally separated than when they arise from similar directions.
4. Temporally aligning the echo threshold levels to energy centroids of corresponding reflection responses yields highly correlated curves and allows the modeling of the echo suppression of arbitrary scattering coefficients.
5. Masked thresholds of specular reflections exhibit a logarithmic relation between the reflection's delay and level which is linearly lowered if the refection is scattered. This relation can be explained by the temporal masking pattern for forward masking.

Acknowledgments

References

1. A.D. Brown, G. Christopher Stecker, D.J. Tollin: The precedence effect in sound localization. Journal of the Association for Research in Otolaryngology 16, 1 (2015) 1–28.
2. R. Litovsky, B. Shinn-Cunningham: Investigation of the relationship among three common measures of precedence: Fusion, localization dominance, and discrimination suppression. The Journal of the Acoustical Society of America 109, 1 (2001) 346–358.
3. H. Kuttruff: Room Acoustics. Taylor & Francis, 5th ed., 2009.
4. R. Guski: Auditory localization: Effects of reflecting surfaces. Perception 19, 6 (1990) 819–830.
5. B. Rakerd, W.M. Hartmann: Localization of sound in rooms, II: The effects of a single reflecting surface. The Journal of the Acoustical Society of America 78 (1985) 524–533.

6. B. Rakerd, W.M. Hartmann: Localization of sound in rooms, III: Onset and duration effects. The Journal of the Acoustical Society of America 80, 6 (1986) 1695–1706.

7. T.J. Cox, B.-I.L. Dalenback, P. D'Antonio, J.J. Embrechts, J.Y. Jeon, E. Mommertz, M. Vorländer: A tutorial on scattering and diffusion coefficients for room acoustic surfaces. Acta Acustica United with Acustica 92 (2006) 1–15.

8. Jens Blauert, P.L. Divenyi: Spectral selectivity in binaural contralateral inhibition. Acta Acustica United with Acustica 66, 5 (1988) 267–274.

9. D.R. Perrott, T. Strybel, C. Manligas: Conditions under which the Haas precedence effect may or may not occur. The Journal of Auditory Research 27 (1987) 59–72.

10. A. Walther, P. Robinson, O. Santala: Effect of spectral overlap on the echo suppression threshold for single reflection conditions. The Journal of the Acoustical Society of America 134, 2 (2013) EL158–EL164.

11. P.W. Robinson, A. Walther, C. Faller, J. Braasch: Echo thresholds for reflections from acoustically diffusive architectural surfaces. The Journal of the Acoustical Society of America 134, 4 (2013) 2755.

12. J. Grosse, S. van de Par, C. Trahiotis: Stimulus coherence influences sound-field localization and fusion/segregation of leading and lagging sounds. The Journal of the Acoustical Society of America 141, 4 (2017) 2673–2680.

13. C. Visentin, M. Pellegatti, N. Prodi: Effect of a single lateral diffuse reflection on spatial percepts and speech intelligibility. The Journal of the Acoustical Society of America 148, 1 (2020) 122–140.

14. T. Lokki, J. Pätynen, S. Tervo, S. Siltanen, L. Savioja: Engaging concert hall acoustics is made up of temporal envelope preserving reflections. The Journal of the Acoustical Society of America 129, 6 (2011) EL223–EL228.

15. S. Siltanen, T. Lokki, S. Tervo, L. Savioja: Modeling incoherent reflections from rough room surfaces with image sources. The Journal of the Acoustical Society of America 131, 6 (2012) 4606–4614.

16. M.A. Biot: Reflection on a rough surface from an acoustic point source. The Journal of the Acoustical Society of America 29, 11 (1957) 1193–1200.

17. M.A. Biot: On the reflection of acoustic waves on a rough surface. The Journal of the Acoustical Society of America 30, 5 (1958) 479–480.

18. M.R. Schroeder: Binaural dissimilarity and optimum ceilings for concert halls: More lateral sound diffusion. The Journal of the Acoustical Society of America 65 (1979) 958–963.

19. T.J. Cox, P. D'Antonio: Acoustic Absorber and Diffusers. 2005.

20. L. Dietsch, W. Kraak: Ein objektives Kriterium zur Erfassung von Echostörungen bei Musik- und Sprachdarbietungen. Acta Acustica United with Acustica 60, 3 (1986) 205–216.

21. B. Rakerd, W.M. Hartmann, J. Hsu: Echo suppression in the horizontal and median saggittal planes. Journal of the Acoustical Society of America 107, 2 (2000) 1061–1064.

22. R.L. Freyman, R.K. Clifton, R.Y. Litovsky: Dynamic processes in the precedence effect. The Journal of the Acoustical Society of America 90, 2 (1991) 874–884.

23. H. Levitt: Transformed up-down methods in psychoacoustics. The Journal of the Acoustical Society of America 49, 2B (1971) 467–477.

24. S.A. Klein: Measuring, estimating, and understanding the psychometric function: A commentary. Perception & Psychophysics 63, 8 (2001) 1421–1455.

25. J. Daniel: Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia. PhD thesis, Université Paris 6, 2001.

26. A. Kohlrausch, R. Kortekaas, M. van der Heijden, S. van de Par, A.J. Oxenham, D. Püschel: Detection of tones in low-noise noise: Further evidence for the role of envelope fluctuations. Acta Acustica United with Acustica 83 (1997) 659–669.

27. H.P. Sepharim: Über die Wahrnehmbarkeit mehrerer Rückwurfe von Sprachschall. Acustica 11, 2 (1961) 80–91.

28. S.S. Sawilowsky: New effect size rules of thumb, The Journal of Modern Applied Statistical Methods 8, 2 (2009) 597–599.

29. B.G. Shinn-Cunningham, P.M. Zurek, N.I. Durlach: Adjustment and discrimination measurements of the precedence effect. The Journal of the Acoustical Society of America 93 (1993) 2923–2932.

30. H. Fastl, E. Zwicker: Psychoacoustics – Facts and Models. 3rd ed., Hugo Fastl, Eberhard Zwicker, 2006.