A Physically Motivated Approach for Binaural Simulation of Moving Sound Sources and Receivers

Christoph URBANIETZ[(1)], Gerald ENZNER[(2)]

[(1)]Ruhr University Bochum, Institute of Communication Acoustics, Germany, christoph.urbanietz@rub.de

[(2)]Ruhr University Bochum, Institute of Communication Acoustics, Germany, gerald.enzner@rub.de

Abstract

In this paper we derive a physically motivated approach to handle the simulation of moving sound sources and receivers including phenomena such as the Doppler effect in audio signal processing by using the concept of retarded time in a very rigorous way. After recapitulating the physical basis of moving point-like sound sources and receivers we derive an alternative formulation of the received signal based on superposition of Green's functions and transfer it to a system theory point of view. A first proof of the validity of this formulation is shown for continuous-time systems. Afterwards we consider a discrete-time implementation and reveal the strengths and the limits of this approach. Finally a binaural receiver is added to the model. The proposed formulation is anchored on a physical basis and therefore it handles physical effects inherently. Beside the correct reproduction of the Doppler frequency shift it also includes the physically attested amplitude shift and further phenomena. Nevertheless the specific implementation for discrete-time signal processing can be challenging. We present a possible implementation and verify some simulation examples against physically predictable features. It turns out that the results of the presented approach and the physical prediction match with high precision.

Keywords: Dynamic Rendering, Binaural, Doppler Effect

1 INTRODUCTION

Virtual acoustics is a key feature in virtual reality and multimedia applications [1] and has been investigated in the past years in various ways. For the directional listening with headphones the use of head-related impulse responses (HRIRs) is the state-of-the-art technique [2] although there is still much ongoing research in this field. The theory is well known for static scenes, where sound source and receiver have static positions and orientations and, therefore, the distance between sound source and receiver can be described as a delay-line and amplitude attenuation. Furthermore, many approaches have been presented how to handle motion of source and receiver. The dynamic or time-varying acoustic environment, however, exhibits a number of sound features that require particular care, e.g., Doppler shift, amplitude modulation, or spatial mislocalization of sound sources.

To address dynamic virtual acoustic environments there are phenomenological approaches which try to mimic physical features without calculating the whole true physical process and, on the other side, physical simulations of the whole acoustic field which suffer from a very high computational effort. The first kind calculates expected features such as Doppler shift or time delays and mimics it, e.g., by fractional delay lines that cover only a part of the physical truth [10, 11, 3, 5, 4]. The second kind uses comprehensive simulations of the acoustic field, e.g., by boundary elements methods [7] but is limited due to high computational effort.

The presented approach is positioned inbetween as it derives a point-to-point description for moving sound sources and receivers from a deep physical starting point. In contrast to work such as [8] where an analytically calculated transfer function based on the physical basis is used for the inverse process, our approach starts at a more basic level of physical description. The desired physical phenomena, such as the Doppler shift, amplitude shift, or further small nonlinearities, are thus generated inherently. Basically the derived description is grounded on a rigorous utilization of the concept of retarded time. In [12] it is shown that already pure azimuthal binaural rendering with constant distances benefits from this treatment of retarded time. Nevertheless, in this paper, the concept is derived from a more fundamental physical basis and also radial velocities are considered.
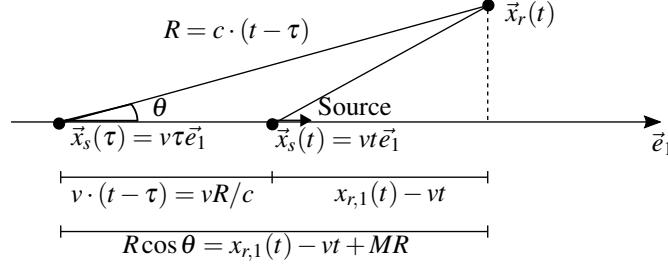
Figure 1. Sound source moving along the $\vec{e}_1$ axis, according to [9].

2 SOUND FIELD OF MOVING SOUND SOURCE AT MOVING RECEIVING POINT

In a first step we review the physical description of a moving sound source. In particular we will constrain ourselves to point-like sound sources. We start with static source and receiver positions and follow [9] defining a velocity potential ψ by $\frac{\partial \psi}{\partial x_i} = -\rho_0 v_i$ where ρ_0 is the static pressure, and v_i is the sound particle velocity in the i-th spatial direction. Thus, according to linearized acoustic Euler equation,

$$p = \frac{\partial \psi}{\partial t} \tag{1}$$

is the sound pressure. We define the sound source by a rate of mass flux $q(t)$. In a scene with a statically positioned source $q(t)$ at the origin of the coordinate system we get the non-homogeneous wave equation

$$\left(\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \psi = -q(t) \delta(x_1) \delta(x_2) \delta(x_3) \tag{2}$$

where c is the speed of sound and $\delta(x_i)$ is the spatial delta distribution. The solution is given by

$$\psi = \frac{q(t - r/c)}{4\pi r} \quad , \quad r = \sqrt{x_1^2 + x_2^2 + x_3^2}. \tag{3}$$

Now we move the source with the velocity v along the $\vec{e}_1$-axis as visualized in Fig. 1. Therefore we get

$$\left(\Delta - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \psi = -q(t) \delta(x_1 - vt) \delta(x_2) \delta(x_3). \tag{4}$$

Morse and Ingard [9] obtained, using Lorentz transform and multiple substitutions, the solution

$$\psi = \frac{q(t - R/c)}{4\pi R (1 - M \cos \theta)} \tag{5}$$

where R is the distance between emission point $\vec{x}_s(\tau)$ at retarded time $\tau = t - R/c$ and receiving point $\vec{x}_r(t)$ of the sound received at time t. The angle θ is visualized in Fig. 1 and R has been found to be

$$R = \frac{M(x_1 - vt) + \sqrt{(x_1 - vt)^2 + (x_2^2 + x_3^2)(1 - M^2)}}{1 - M^2} \quad , \quad M = \frac{v}{c}. \tag{6}$$

2.1 Alternative way of the solution

For a more general trajectory of the source (and receiver) we now choose another description. The solution to the non-homogeneous wave equation for acoustic waves

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \Delta \right) \psi = \delta(\vec{x}) \delta(t) \tag{7}$$

with a given point-like source in space and time is Green's function

$$\psi = G(\vec{x},t) = \frac{1}{4\pi}\frac{\delta(t-|\vec{x}|/c)}{|\vec{x}|}.$$ (8)

For a source distribution with only a single point-like source $q(t)$ positioned at $\vec{x}_s(t)$

$$Q(\vec{x},t) = q(t)\delta(\vec{x}-\vec{x}_s(t))$$ (9)

we can get the velocity potential at receiving point $\vec{x}_r$ as superposition of elementary solutions by

$$\psi(\vec{x}_r,t) = \iint G(\vec{x}_r-\vec{x},t-\tau)Q(\vec{x},\tau)\,\mathrm{d}^3\vec{x}\,\mathrm{d}\tau = \frac{1}{4\pi}\iint \frac{q(\tau)\delta(\vec{x}-\vec{x}_s(\tau))}{|\vec{x}_r-\vec{x}|}\delta(t-\tau-|\vec{x}_r-\vec{x}|/c)\,\mathrm{d}^3\vec{x}\,\mathrm{d}\tau$$ (10)

$$= \frac{1}{4\pi}\int \frac{q(\tau)}{|\vec{x}_r-\vec{x}_s(\tau)|}\delta(t-\tau-|\vec{x}_r-\vec{x}_s(\tau)|/c)\,\mathrm{d}\tau.$$ (11)

This new description potentially can be used with arbitrary trajectories. For first verification we apply $\vec{x}_s(\tau) = v\tau\vec{e}_1$ and compare the result to (5). The argument of the delta distribution is $t-\tau-\sqrt{(x_{r,1}-v\tau)^2+x_{r,2}^2+x_{r,3}^2}/c$ which has its root at $\tau = t - \frac{R}{c}$ with R as defined in (6). Due to the derivative of the argument of the delta distribution $\frac{\partial}{\partial\tau}(t-\tau-\sqrt{(x_{r,1}-v\tau)^2+x_{r,2}^2+x_{r,3}^2}/c) = 1-M\cos(\theta)$ we get during the integration an additional factor of $1/(1-M\cos\theta)$ and therefore the result

$$\psi(\vec{x}_r,t) = \frac{1}{4\pi}\frac{q(t-R/c)}{R(1-M\cos\theta)}$$ (12)

is indeed equal to (5).

3 SYSTEM THEORY POINT OF VIEW

In the next step we transfer (11) into a system theory context. We absorb all constants in r_0 and therefore

$$\psi(t) = \int \frac{r_0}{|\vec{x}_r(t)-\vec{x}_s(\tau)|}\delta(t-\tau-|\vec{x}_r(t)-\vec{x}_s(\tau)|/c)q(\tau)\,\mathrm{d}\tau$$ (13)

where we consider now $q(t)$ to be an input signal to our system and $\psi(t)$ to be the output signal at position $\vec{x}_r(t)$. Thus, we can identify a time varying, i.e., t-dependent, transfer function as

$$h(\vec{x}_r,\vec{x}_s,t,\tau) = \frac{r_0}{|\vec{x}_r(t)-\vec{x}_s(t-\tau)|}\delta(\tau-|\vec{x}_r(t)-\vec{x}_s(t-\tau)|/c).$$ (14)

When we start with a sound pressure signal $s(t)$ which was measured with a microphone in a stationary case and, therefore, was achieved by differentiating a velocity potential by time, we have vice versa to integrate $s(t)$ over time to get a mass flux equivalent $q(t)$. The resulting sound pressure $p(t)$ is achieved by a final differentiation by

$$q(t) = \int_0^t s(\tau)\,\mathrm{d}\tau \quad , \quad p(t) = \frac{\mathrm{d}\psi}{\mathrm{d}t}.$$ (15)

3.1 Proof on specific example

We shall confirm that the proposed approach leads to reasonable results, especially that this approach reproduces Doppler effect and amplitude modulation inherently. We therefore do an explicit calculation of a moving source
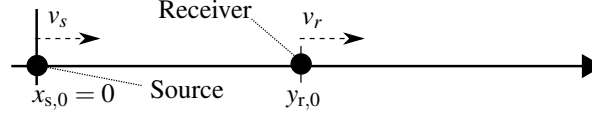
Figure 2. Visualization of the tested scene.

and receiver along the first axis of the global coordinate system. The sound source is given by $s(t) = \sin(\omega_0 t)$. The receiver moves with velocity v_r starting at $x_{r,0}$ and the source moves with velocity v_s starting at $x_{s,0} = 0$ as visualized in Fig. 2. We further restrict to $x_s(t) = v_s t < x_r(t) = x_{r,0} + v_r t$. The output signal received at x_r is

$$p(t) = \frac{\mathrm{d}}{\mathrm{d}t} \int \frac{r_0}{|x_r(t) - x_s(\tau)|} \delta\left(t - \tau - \frac{|x_r(t) - x_s(\tau)|}{c}\right) \int_0^\tau s(\tau') \, \mathrm{d}\tau' \, \mathrm{d}\tau \tag{16}$$

$$= \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{c}{c - v_s} \frac{r_0}{|x_r(t) - x_s(\tau_t)|} \frac{1}{\omega_0}\left(1 - \cos(\omega_0 \tau_t)\right)\bigg|_{\tau_t = \left(t\frac{c-v_r}{c} - \frac{x_{r,0}}{c}\right)\frac{c}{c-v_s}}\right) \tag{17}$$

$$= \frac{c(c - v_r)}{(c - v_s)^2} \frac{r_0}{|x_r(t) - x_s(\tau_t)|} \sin(\omega_0 \tau_t) - \frac{c^2(v_r - v_s)}{(c - v_s)^2} \frac{r_0}{|x_r(t) - x_s(\tau_t)|^2} \frac{1}{\omega_0}(1 - \cos(\omega_0 \tau_t)). \tag{18}$$

The root of the delta distribution argument $\tau_t = \left(t\frac{c-v_r}{c} - \frac{x_{r,0}}{c}\right)\frac{c}{c-v_s}$, called "retarded time", is the emission time of the sound received at time t and therefore $|x_r(t) - x_s(\tau_t)|$ is the distance between emission and receiving point of this sound part. The first of the two terms in (18) equals the results in [9] if we choose $v_r = 0$. The second term is decreasing with $1/R^2$ and rather small and therefore not further considered. Nevertheless the presented approach also recovers these non-linear effects. This second term differs a bit from [9] which can be explained by a mistake found in [9]. In [9] the second term of 11.2.15 has to be $\frac{q}{4\pi}\frac{(\cos\theta - M)V}{R^2(1 - M\cos\theta)^3}$ due to the outer derivation, and this coincides with (18).

Having a deeper look at the first term of (18), we see the sound pressure is by a factor $(c + v_s)^2/(c - v_s)^2$ larger for a sound source moving towards the receiver compared to a sound source moving away from the receiver. This amplitude factor is known in literature [9] and reproduced by the presented calculation. An even more commonly known phenomenon is the Doppler-shift. If we differentiate the phase of the sine in (18) we get

$$\frac{\mathrm{d}}{\mathrm{d}t}\omega_0 \tau_t = \omega_0 \frac{c - v_r}{c - v_s} \tag{19}$$

which is the Doppler shift as expected.

3.2 With binaural receiver

As the human ears are no point-like receivers we have to consider their characteristics. Usually this is done by an HRIR $\hat{h}_{l,r}(\theta, \phi, \tau)$ depending on the relative angular position, given by azimuth ϕ and elevation θ, of the source in the head coordinate system. For a scene with the position $\vec{x}_r$ and orientation (e.g, given by a triple of Euler angles $\vec{\alpha}_r$) of the receiving head and the position $\vec{x}_s$ of the point-like source we can calculate the relative angles ϕ and θ. Instead of denoting the HRIR with relative angles we can also denote the relative orientation of the source by absolute positions and orientations with $\hat{h}_{l,r}(\theta, \phi, \tau) = \check{h}_{l,r}(\vec{x}_r, \vec{\alpha}_r, \vec{x}_s, \tau)$. As the HRIR describes the difference between a fictional point-like microphone in the center of the head and the signals at the ears we can describe the whole transfer function from source to binaural receiver as

$$h_{l,r}(\tau) = \hat{h}_{l,r}(\theta, \phi, \tau) * h(\vec{x}_r, \vec{x}_s, \tau) = \check{h}_{l,r}(\vec{x}_r, \vec{\alpha}_r, \vec{x}_s, \tau) * h(\vec{x}_r, \vec{x}_s, \tau). \tag{20}$$

For treating time-varying scenes we associate the position and orientation of the receiver (i.e., $\vec{x}_r, \vec{\alpha}_r$) with the receiving time t and the position of the source (i.e., $\vec{x}_s$) with the emission time τ, in analogy to (13).

This approach has been presented and validated in [12], although we have to admit that the difference between velocity potential and sound pressure was not taken into account in [12]. The binaural received velocity potential is then given by

$$
\begin{aligned}
\psi_{l,r}(t) &= \iint h(\vec{x}_r(t), \vec{x}_s(\tau), \tau_1) \check{h}_{l,r}(\vec{x}_r(t), \vec{\alpha}_r(t), \vec{x}_s(\tau), (t-\tau) - \tau_1) q(\tau) \, d\tau_1 \, d\tau \\
&= \iint \frac{1}{|\vec{x}_r(t) - \vec{x}_s(\tau)|/r_0} \delta(\tau_1 - |\vec{x}_r(t) - \vec{x}_s(\tau)|/c) \check{h}_{l,r}(\vec{x}_r(t), \vec{\alpha}_r(t), \vec{x}_s(\tau), (t-\tau) - \tau_1) q(\tau) \, d\tau_1 \, d\tau \\
&= \int \frac{1}{|\vec{x}_r(t) - \vec{x}_s(\tau)|/r_0} \check{h}_{l,r}(\vec{x}_r(t), \vec{\alpha}_r(t), \vec{x}_s(\tau), (t-\tau) - |\vec{x}_r(t) - \vec{x}_s(\tau)|/c) q(\tau) \, d\tau.
\end{aligned}
\tag{21}
$$

By choosing a unity transfer function for the HRIR, i.e., $\hat{h}_{l,r}(\theta, \phi, \tau) = \delta(\tau)$, it reverts to (13).

4 DISCRETE SOUND SCENE RENDERING

The next step is to transform (21) into a discrete form to use it for digital sound rendering. We denote discrete sampled signals with square brackets like $\psi[k] = \psi(kT)$ with discrete-time index k and sampling time T. Thus

$$
\psi[k] = \sum_{\kappa} \frac{r_0}{|\vec{x}_r[k] - \vec{y}_s[\kappa]|} \check{h}_{l,r}(\vec{x}_r[k], \vec{\alpha}_r[k], \vec{y}_s[\kappa], (k-\kappa)T - |\vec{x}_r[k] - \vec{y}_s[\kappa]|/c) q[\kappa].
\tag{22}
$$

There are now some obstacles to overcome:

- The limits for κ range from $-\infty$ to ∞, which means infinite computational effort.

- $(k-\kappa)T - |\vec{x}_r[k] - \vec{y}_s[\kappa]|/c$ is in general not an integer multiple of T, so that we cannot simply make use of a discrete sampled version of $\check{h}_{l,r}$.

To overcome the second issue we have to temporally interpolate the discrete HRIR for a given orientation $\check{h}_{l,r}[\vec{x}_r[k], \vec{\alpha}_r[k], \vec{y}_s[\kappa], \cdot]$, shortly denoted by $h_{l/r}[\cdot]$ or $h_{l/r}(\cdot)$, respectively. As an optimal interpolation we use

$$
h_{l/r}(\tau) = \sum_{\kappa=0}^{N-1} h_{l/r}[\kappa] \operatorname{si}(\kappa - \tau/T) \quad \text{with} \quad \operatorname{si}(x) = \frac{\sin(\pi x)}{\pi x}
\tag{23}
$$

where N is the order of the discrete HRIR-filter. As a result we get

$$
\psi[k] = \sum_{\kappa} \sum_{m=0}^{N-1} \frac{r_0}{|\vec{x}_r[k] - \vec{y}_s[\kappa]|} \check{h}_{l,r}\left[\vec{x}_r[k], \vec{\alpha}_r[k], \vec{y}_s[\kappa], m\right] \operatorname{si}\left(m - \left((k-\kappa) - \frac{|\vec{x}_r[k] - \vec{y}_s[\kappa]|}{Tc}\right)\right) q[\kappa]
\tag{24}
$$

At this stage we would have to sum over all possible κ, meaning over the whole input signal. Especially we would have to do this N times for every output sample. This could be very time consuming and would result in an infeasible problem for continuous audio rendering. One approach to overcome this is to chose an interpolator with finite size which we will later denote by $\operatorname{I}(\cdot)$ instead of $\operatorname{si}(\cdot)$.

4.1 More efficient concatenated processing

For more tangible processing we rely on (21) and have a closer look at (14). The impulse response is only non-zero if the argument of the delta distribution is zero and, thus, we can find a unique mapping from t to τ and vise versa as long as the velocities are lower than the speed of sound. In other words, this is a mapping between emission time τ and receiving time t of a distinct sound event. In (17) this mapping from t to τ for example is done by the definition of τ_t. We denote the general mapping from t to τ as $\tau = o(t)$ where o stands
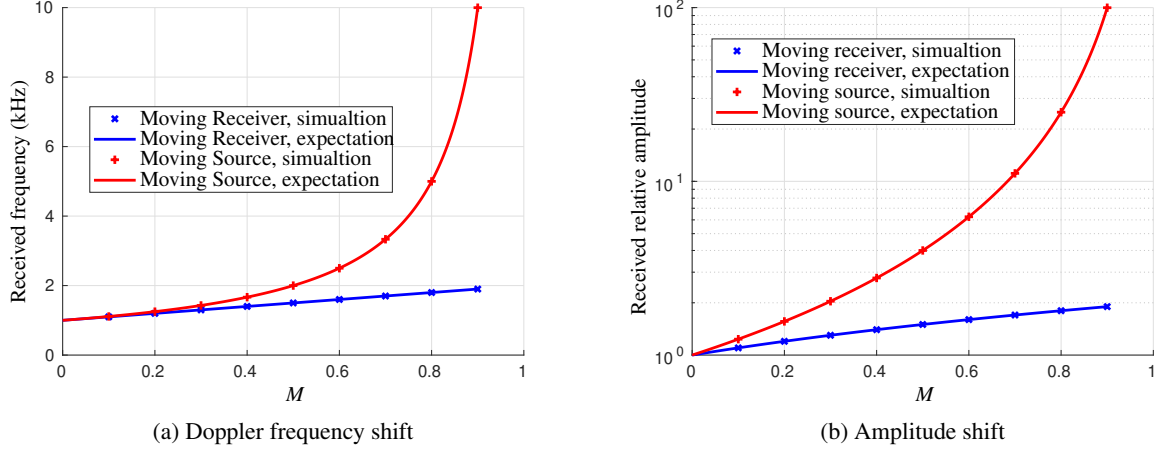
Figure 3. Measured amplitude and frequency of received signals, simulation vs. expectation, $w = 100$.

for originating. Therefore we can write (21) as

$$\psi_{l,r}(t) = \iint h(\vec{x}_r(t), \vec{y}_s(\tau), (t-\tau) - \tau_1) \check{h}_{l,r}(\vec{x}_r(t), \vec{\alpha}_r(t), \vec{y}_s(\tau), \tau_1) q(\tau) \, \mathrm{d}\tau_1 \, \mathrm{d}\tau$$

$$= \int \left[\int h(\vec{x}_r(t), \vec{y}_s(\tau), (t-\tau_1) - \tau) q(\tau) \, \mathrm{d}\tau \right] \check{h}_{l,r}(\vec{x}_r(t), \vec{\alpha}_r(t), \vec{y}_s(o(t-\tau_1)), \tau_1) \, \mathrm{d}\tau_1$$

$$= \int \psi_m(t-\tau_1) \check{h}_{l,r}(\vec{x}_r(t), \vec{\alpha}_r(t), \vec{y}_s(o(t-\tau_1)), \tau_1) \, \mathrm{d}\tau_1 . \qquad (25)$$

This describes a subsequent binaural rendering of a mono distance-rendered signal with a retarded source position. Thus, in a first step we render the mono sound and simultaneously save the emission time for each output sample and do a simple binaural rendering of the mono sound using a source position that belongs to this saved retarded time in a second step. The binaural rendering part in the discrete case is then given by

$$\psi_{l,r}[k] = \sum_{\ell=0}^{N-1} \check{h}_{l,r}\left[\vec{x}_r[k], \vec{\alpha}_r[k], \vec{y}_s[o[k-\ell]], \ell\right] \psi_m[k-\ell] . \qquad (26)$$

For the radial (distance) rendered part we can set $\check{h}_{l,r}\left[\vec{x}_r[k], \vec{\alpha}_r[k], \vec{y}_s[\kappa], m\right] = \delta[m]$ in (24) and get

$$\psi_m[k] = \sum_{\kappa} \frac{r_0}{|\vec{x}_r[k] - \vec{y}_s[\kappa]|} \mathrm{I}\left(\left((k-\kappa) - \frac{|\vec{x}_r[k] - \vec{y}_s[\kappa]|}{Tc}\right)\right) q[\kappa] \qquad (27)$$

where we use a finite interpolator $\mathrm{I}(\cdot)$ with one-sided width w. Thus, the interpolator has to be evaluated only for κ in the range

$$-w \leq \kappa - k - \frac{|\vec{x}_r[k] - \vec{y}_s[\kappa]|}{Tc} \leq w . \qquad (28)$$

As an additional, external step, we need a discrete integration to get $q[k]$ from $s[k]$ and a discrete differentiation to get $p[k]$ from $\psi[k]$. In a first test this is done by first order integration and differentiation as

$$q[k+1] = q[k] + s[k]/T \quad \text{and} \quad p[k+1] = (\psi[k+1] - \psi[k])/T . \qquad (29)$$

For continuously operating systems an additional anti-windup in the integration can be useful.

(a) 1st order integration and differentiation

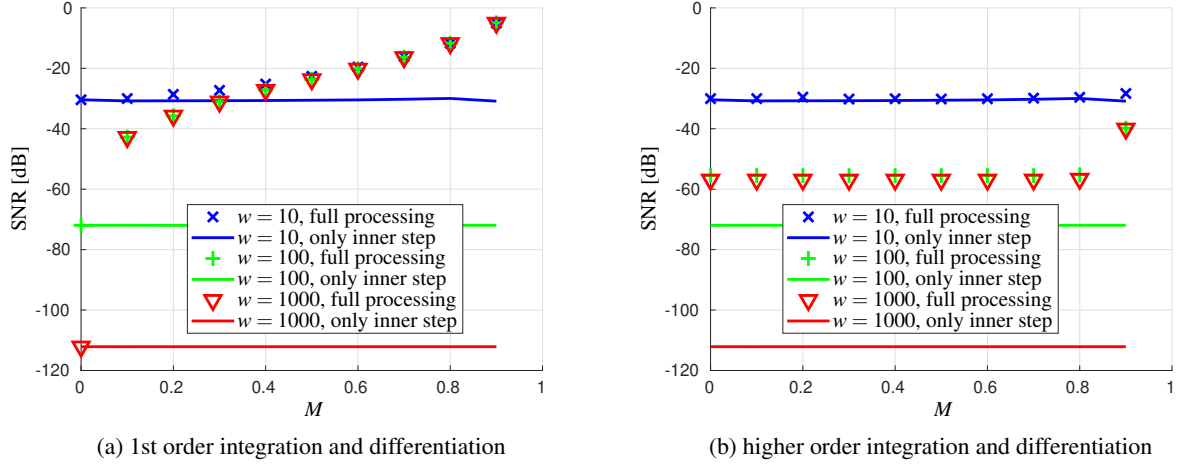(b) higher order integration and differentiation

Figure 4. SNR between simulated and predicted signal for various source velocities and window-width w, performance of full processing and inner time-retardation step compared.

5 EXPERIMENTAL RESULTS

The evaluation of the whole system consisting of radial and binaural part is quite complicated due to the absence of a ground-truth reference signal. We implemented the whole processing including the binaural receiver using a continuous-azimuth representation of the HRIR as described in [13] for motions in a 2D plane. Some examples were rendered, e.g., a passing ambulance that sounded plausible.

A way to evaluate the proposed method with objective measures is the comparison of physically expected features with the simulation results. Although it tackles only the part in (11) without the integration and differentiation steps, for the angular part we have done such comparisons in [12]. Thus, below we concentrate on the evaluation of the radial part. A verification of the continuous-time calculation was already done in the sections before. Therefore the numerical implementation is left for evaluation in this section. For the radial part we know the results for frequency and amplitude shift in theory. We consider a scene as shown in Fig. 2 with a 1 kHz sine as source signal. For this scene we know the frequency shift is given by $\omega_{\mathrm{Doppler}} = \omega_0 \frac{c-v_r}{c-v_s}$ and the amplitude shift is given by $A_{\mathrm{dyn}} = A_0 \frac{c(c-v_r)}{(c-v_s)^2}$. The results for various $M = v/c$ for independent source or receiver motion is shown in Fig. 3. We see that the numerical simulation meets the physical expectation very well.

For the radial part we can furthermore predict the received signal for the given scene by (18). We use a moving source in this case with various $M = v_s/c$ and compare the numerically simulated signal with the predicted one. Therefore we calculate an waveform distance defined by SNR $= \frac{\sum (u_{\mathrm{simulated}} - u_{\mathrm{predicted}})^2}{\sum u_{\mathrm{predicted}}^2}$ and show the results in Fig. 4a. The simulation was done for the window-widths $w = 10; 100; 1000$. We see that the full processing suffers from higher velocities. Therefore we separated the time retarding part (11) from the integration and differentiation and plotted the results of inner step (27) additionally in Fig. 4a. This part delivers constant performance over all velocities and therefore the deteriorating effect for larger velocities is caused by the implementation of integration and differentiation. When we use a more sophisticated integration and differentiation Fig. 4b shows an improvement in performance. Here we used $q[k] = q[k-1] + (s[k] + s[k-1])/2$ and an 8th-order central finite difference for differentiation. It should be noted that the SNR is a very rigorous measure as also small phase-shifts would dramatically affect it, even if they are non-audible.

6 CONCLUSIONS

In this paper we introduced a new approach to handle moving sound sources and receivers within virtual acoustic environments. The presented approach is positioned between phenomenological approaches which try to mimic physical features without calculating the whole true physical process and, on the other side, comprehensive physical simulations of the whole acoustic field which suffer from very high computational effort. In this paper we derived a description for moving sound sources and receivers from a deep physical starting point. Therefore, our approach handles many physical phenomena, such as Doppler shift, amplitude shift, or further small nonlinearities inherently. We showed the reproduction of these effects in theory and in numerical simulations. This paper is meant to introduce this kind of description for advanced binaural rendering in a general way. Further improvements in the numerical implementation are possible, e.g., by more sophisticated interpolation algorithms [6] or better integration and differentiation techniques.

ACKNOWLEDGEMENTS

REFERENCES

[1] D. R. Begault. *3D Sound for Virtual Reality and Multimedia*. Academic Press, Boston, Oct. 1994.

[2] J. Blauert, editor. *The Technology of Binaural Listening*. Springer, 2013.

[3] J. M. Chowning. The Simulation of Moving Sound Sources. *JAES*, 19(1):2–6, Jan. 1971.

[4] Y. Iwaya and Y. Suzuki. Rendering moving sound with the Doppler effect in sound space. *Applied Acoustics*, 68(8):916–922, Aug. 2007.

[5] R. Kronland-Martinet and T. Voinier. Real-Time Perceptual Simulation of Moving Sources: Application to the Leslie Cabinet and 3D Sound Immersion. *J. Audio, Speech, Music Process.*, 2008(1):849696, June 2008.

[6] T. I. Laakso, V. Valimaki, M. Karjalainen, and U. K. Laine. Splitting the unit delay [FIR/all pass filters design]. *IEEE Signal Processing Magazine*, 13(1):30–60, Jan. 1996.

[7] P. L. Lee and J. H. Wang. The simulation of binaural hearing caused by a moving sound source. *Computers & Structures*, 87(17):1102–1110, Sept. 2009.

[8] F. Meng, G. Behler, and M. Vorländer. A Synthesis Model for a Moving Sound Source Based on Beamforming. *Acta acustica united with acustica*, 104(2):351–362, 2018.

[9] P. M. Morse and K. U. Ingard. *Theoretical Acoustics*. Princeton University Press, 1968.

[10] J. O. Smith and D. P. Berners. Doppler Simulation and the Leslie. In *COST-G6 Conf. Digital Audio Effects (DAFx-02)*, pages 13–20, Sept. 2002.

[11] H. Strauss. Implementing Doppler Shifts for Virtual Auditory Environments. In *Audio Eng. Soc. Conv. 104*, May 1998.

[12] C. Urbanietz and G. Enzner. Binaural Rendering of Dynamic Head and Sound Source Orientation Using High-Resolution HRTF and Retarded Time. In *IEEE Int. Conf. Acoustics, Speech, Signal Process. (ICASSP)*, pages 566–570, Apr. 2018.

[13] C. Urbanietz and G. Enzner. Spatial-Fourier Retrieval of Head-related Impulse Responses from Fast Continuous-Azimuth Recordings in the Time-Domain. In *IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, pages 950–954, May 2019.