T. Schmele and A. Garriga,
"Spatial Matrix Synthesis,"
J. Audio Eng. Soc., vol. 72, no. 4, pp. 193–210 (2024 Apr.).
https://doi.org/10.17743/jaes.2022.0130.

Spatial Matrix Synthesis

TIMOTHY SCHMELE,[1, 2, *] *AES Associate Member,* **AND ADAN GARRIGA**[1]
(tim.schmele@eurecat.org) (adan.garriga@eurecat.org)

[1] *Eurecat, Technology Center of Catalonia, Multimedia Technologies, Barcelona, Spain*
[2] *Institute for Music Informatics and Musicology, University of Music, Karlsruhe, Germany*

Spatial Matrix synthesis is presented in this paper. This modulation synthesis technique creates acoustic velocity fields from acoustic pressure signals by using spatial transformation matrices, thus generating complete sound fields for spatial audio. The analysis presented here focuses on orthogonal rotation matrices in both two and three dimensions and compares the results in each scenario with other sound modulation synthesis methods, including amplitude and frequency modulation. As an alternative method for spatial sound synthesis that exclusively modifies the acoustic velocity vector through effects comparable to those created by both amplitude and frequency modulations, Spatial Matrix synthesis is argued to generate inherently spatial sounds, giving this method the potential to become a new musical instrument for spatial music.

0 INTRODUCTION

The general motivation to improve the acoustic spatial image in audio reproduction has mostly been the desire to feel as if "being there" when listening to a recording [1]. *Panning*, in particular, a technique in which prerecorded sounds are artificially placed in between two or more loudspeakers in post, often also referred to as *spatialization*, is a problem to which many different solutions have since been proposed. These include the ubiquitous two-channel stereophony, or more elaborate techniques often employed for full spherical audio reproduction, such as VBAP [2] or *Ambisonics* [3, 4]. Engineers and artists have since used these panning techniques to go beyond naturalism and create hyperreal productions [5]. Composers of electroacoustic music embraced virtual source panning techniques in a multitude of ways and incorporated them into their compositional practice [6–10].

Despite this, comparatively little investigation has gone into the effects caused by the velocity of panned sounds. Studies analyzing spatial audio reproduction methods with regard to moving sources mostly focus on the Doppler shift [11–13]. It has also been remarked that slow moving sources produce negligible Doppler shift effects [14]. Naturally, few objects in the real world move fast enough to cause noticeable spectral distortions, and it has also been shown that people's ability to follow a sound rotating around them breaks down at about three rotations per second [15]. Thus, why should anyone be concerned with angular velocities higher than the above? Recently, the authors have shown that a sound source rotating with a constant angular frequency causes a frequency split to appear in the velocity field at the center of rotation [16]. Due to the spectral distortions that occur in this case, these techniques can lead to approaches that tightly fuse spatialization and sound synthesis into one process, arguably creating *inherently spatial* sounds.

Most common synthesis methods, such as AM [17, 18] or FM [19–21], usually act on the audio pressure signal. Existing approaches that intend to combine spatialization with sound synthesis usually treat spatialization as an added effect on top of the synthesis method, and not as an integral part of it. For example, time-based approaches, such as granular synthesis, spatialize each grain of audio individually [22–25], while spectral approaches give each extracted frequency band an individual position in space [23, 26–28]. It has also been proposed to reuse the trajectories generated by wave terrain synthesis as a way to more tightly combine the sound synthesis technique with its spatialization [26, 29, 30]. Methods that use spatial parameters to modulate the signal include varying the source distance, which produces timbral changes that are similar to the effects of FM synthesis, due to rapid changes in the amount of Doppler shift applied to the source [31–33], as well as the expansion of the pairwise panning technique to 3D, to be able to pan a sound source at audio rate using *wavetables* [34–36].

One important factor of a sound's spatial appearance is its spatial extent. The perception of the spatial extent, or Apparent Source Width, is often linked to the Interaural Cross Correlation coefficient [37–39]. Spatially extended sources are produced by vibrating finite-sized plates or

*To whom correspondence should be addressed, email: tim.schmele@eurecat.org, Last updated: March 11, 2024.

spheres [40] but also occur naturally in psychoacoustic appearances, such as the architectural control of reflections in a concert hall [41] or waves crashing down on beach fronts [38]. The Interaural Cross Correlation can be controlled by the amount of correlation present in a coherent source. Taking the ubiquitous stereo panning as an example, two correlated loudspeakers fuse together to appear as one source perceptually. This correlation can be gradually broken up by utilizing frequency-dependent delays and distributing spectral bands in space [38, 39, 41].

As is the case with spatial granular or spectral synthesis approaches mentioned above, while a sound is either distributed temporally or spectrally in space, it is still coherent and perceived as a single sound event, but its spatial extent is altered proportionally to the amount of decorrelation [26, 39, 41, 42]. Thus, after the ability to follow a rotating sound breaks down beyond about three rotations per second [15], a sound rotating at higher angular velocities is effectively redistributed in space around the listener. While it remains coherent, it is decorrelated, and thus its Apparent Source Width is effectively increased.

This paper presents *Spatial Matrix* (SM) synthesis, a method for generating sounds by synthesizing velocity fields from acoustic pressure signals. Inspired by the physical model of a rotating source [16], SM synthesis includes applying rotation matrices to a source in both 2D and 3D. Unlike most other methods that combine spatialization with sound synthesis, which also affects the pressure field, SM synthesis neither alters the source nor pressure field at the center; all effects achieved by SM synthesis are exclusively located in the velocity field. Also, unlike methods that make use of the "Doppler FM" effect [31–33], the approach presented here maintains the source distance constant and therefore achieves all effects described below by changing the source's angle of incidence to the listener. The authors therefore argue that spatialization is at the core of this method and not simply an afterthought.

This paper is divided into two main parts, each looking at the 2D and 3D cases for SM synthesis separately. SEC. 1 will first define 2D SM synthesis in general, followed by an investigation into different types of rotation matrices. The section will also go into some more complex scenarios, such as using multiple sources and rotations. In turn SEC. 2 defines 3D SM synthesis and expands the analysis on the most interesting cases of SEC. 1, with particular focus on those phenomena that were previously not possible in the 2D case. It shall be seen that in the 3D case, using only rotations, the resulting spectrum can be shaped parametrically. After that, an implementation example for SM synthesis shall be discussed in SEC. 3 using Ambisonics. The paper will then conclude in SEC. 4, discussing the findings and bring forth proposals for future work.

1 SM SYNTHESIS IN 2D

For the 2D case, a harmonic sound source $p_{\text{rad}}(t) = A\cos(\omega_s t + \phi)$ in a free field is considered, with amplitude A, angular frequency ω_s, and phase ϕ, located on the perimeter of a disc of radius $R > 0$ at a position defined by the angle θ (see Fig. 1). 2D SM synthesis will first be briefly defined in general in SEC. 1.1. However, for the remainder of this section, these investigations shall be focused on rotation matrices, as motivated by the findings from analyzing the physical model of a rotation source shown in [16]. In [16] it was shown that the acoustic pressure radiated from a rotating source modulates the amplitude of the velocity components, which causes a spectrum split in the velocity field at the center of rotation. Due to the source's constant distance from the center, however, the periodicity of the pressure is maintained, meaning that the pressure field remains unaltered when compared to a stationary source.

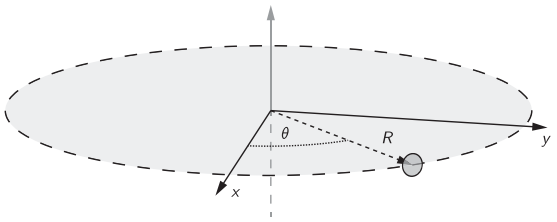


Fig. 1. Sound source in a ring of radius R, θ being the angular coordinate.

1.1 General Definition in 2D

In two dimensions, the acoustic field is completely determined by the acoustic pressure p and the acoustic velocity vector $\mathbf{v} = (v_x, v_y)$. In the most general terms, the 2D SM synthesis is defined as a transformation of the velocity field produced by an input sound pressure signal $p_{\text{in}}(t)$. In accordance with the findings in [16], the output pressure field measured at the center of the rotation shall be considered the same as the input, $p_{\text{out}}(t) = p_{\text{in}}(t)$, and define the input vector $\mathbf{v}_{\text{in}}(t) = (C_x p_{\text{in}}(t), C_y p_{\text{in}}(t))$, where $C_x = C\cos(\theta_0)$ and $C_y = C\sin(\theta_0)$ are constants, with $C = \frac{1}{\rho c}$, in which ρ is the density of the fluid and c the speed of sound inside this fluid, and θ_0 being the initial source position. SM synthesis thus reads

$$\mathbf{v}_{\text{out}}(t) = \mathbf{M}(t) \cdot \mathbf{v}_{\text{in}}(t). \tag{1}$$

In general, $\mathbf{M}(t)$ is a time-dependent transformation matrix. Given an acoustic pressure input $p_{\text{in}}(t)$, Eq. (1) synthesizes an output velocity field, $\mathbf{v}_{\text{out}}(t) = (v_{x_{\text{out}}}(t), v_{y_{\text{out}}}(t))$.

1.2 Rotational SM Synthesis in 2D

More specifically, orthogonal transformations, in particular on rotation matrices, will be focused on. Therefore, in the case of a 2D rotation in the velocity field, the transformation matrix $\mathbf{M}(t) = \mathbf{R}(\theta(t))$ reads

$$\mathbf{R}(\theta(t)) = \begin{bmatrix} \cos\theta(t) & -\sin\theta(t) \\ \sin\theta(t) & \cos\theta(t) \end{bmatrix}. \tag{2}$$

For the analysis of the rest of this section, consider the case of a harmonic sound source of angular frequency ω_s:

$$p_{\text{in}}(t) = p_{\text{out}}(t) = A\cos(\omega_s t). \tag{3}$$

With this in mind, Eq. (1) now reads

$$\mathbf{v}_{\text{out}}(t) = \begin{pmatrix} CA\cos(\theta(t)+\theta_0)\cos(\omega_s t) \\ CA\sin(\theta(t)+\theta_0)\cos(\omega_s t) \end{pmatrix}. \tag{4}$$

From now on, $\theta_0 = 0$, meaning the starting position of the source is in the front, shall be considered. In the following subsections, different scenarios for $\theta(t)$ shall be analyzed.

1.2.1 $\theta(t) = \alpha$

Keeping the angle $\theta(t) = \alpha$ constant in Eq. (4) results in the following complete sound field:

$$p_{\text{out}}(t) = A\cos(\omega_s t), \tag{5}$$

$$v_{x_{\text{out}}}(t) = CA\cos(\alpha)\cos(\omega_s t), \tag{6}$$

$$v_{y_{\text{out}}}(t) = CA\sin(\alpha)\cos(\omega_s t), \tag{7}$$

which is equivalent to the source at a different angle shifted by α.

1.2.2 $\theta(t) = \omega_r t$

Likewise, if $\theta(t)$ is replaced for $\omega_r t$ in Eq. (4), where ω_r is the constant rotational frequency, the following is obtained:

$$p_{\text{out}}(t) = A\cos(\omega_s t), \tag{8}$$

$$v_{x_{\text{out}}}(t) = \frac{CA}{2}\Big[\cos((\omega_s+\omega_r)t) + \cos((\omega_s-\omega_r)t)\Big], \tag{9}$$

$$v_{y_{\text{out}}}(t) = \frac{CA}{2}\Big[\sin((\omega_s+\omega_r)t) - \sin((\omega_s-\omega_r)t)\Big]. \tag{10}$$

Note that one may define $A = A'$ to comply with the physical model analyzed in [16]. However, if $CA = 1$, $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$ in Eqs. (9) and (10) are similar to those for ring modulation (RM) synthesis [17, 18]:

$$RM = \cos(\omega_c t)\cdot\cos(\omega_m t), \tag{11}$$

where ω_c is the frequency of the carrier, which can be compared to ω_s, while ω_m is the modulation frequency, which can be compared to the rotational frequency ω_r. The result for $v_{y_{\text{out}}}(t)$ in Eq. (10) deviates from RM synthesis by demonstrating a phase shift of $-\frac{\pi}{2}$ for the sum and $\frac{\pi}{2}$ for the difference.

RM synthesis causes the carrier to disappear while only the two side-bands remain. However, Eq. (8) also needs to be taken into account, which adds ω_s to the overall sound field. AM synthesis, in turn, uses a unipolar modulator and, thus, maintains the carrier in its result [17, 18]:

$$AM = \cos(\omega_c t)\cdot\frac{1}{2}\Big[\cos(\omega_m t) + 1\Big]. \tag{12}$$

Therefore, if the entire sound field is considered, including the pressure component $p_{\text{out}}(t)$ and velocity components $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$, the spectral makeup looks more akin to AM synthesis.

The perpendicular phase relationship between $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$ hints at the novel spatial property that this synthesis approach features. Not only are $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$ decorrelated, but if one neglects the velocity field and remove the spatial aspect of this synthesis approach, one would be left with only the input frequency ω_s in $p_{\text{out}}(t)$. All side-bands exist exclusively in the spatial velocity components of the signal. As such, the authors tend to call new sounds synthesized with SM synthesis (or similar approaches) as *inherently spatial*.

1.2.3 $\theta(t) = \frac{\alpha}{2}t^2$

Consider now an accelerated trajectory, $\theta(t) = \frac{\alpha}{2}t^2$, where α is the angular acceleration. Substituting in Eq. (4), the complete sound field now becomes

$$p_{\text{out}}(t) = A\cos(\omega_s t), \tag{13}$$

$$v_{x_{\text{out}}}(t) = \frac{CA}{2}\Big[\cos\Big((\omega_s+\frac{\alpha}{2}t)t\Big) + \cos\Big((\omega_s-\frac{\alpha}{2}t)t\Big)\Big], \tag{14}$$

$$v_{y_{\text{out}}}(t) = \frac{CA}{2}\Big[\sin\Big((\omega_s+\frac{\alpha}{2}t)t\Big) - \sin\Big((\omega_s-\frac{\alpha}{2}t)t\Big)\Big]. \tag{15}$$

In this case the two side-bands generated in the velocity field depart from ω_s linearly with time.

1.2.4 $\theta(t) = D\sin(\omega_o t)$

In this scenario for $\theta(t)$ the focus is turned from a constant acceleration to a periodic change in rotation velocity. Returning to Eq. (4), setting $\theta(t) = D\sin(\omega_o t)$, the following would instead be received:

$$p_{\text{out}}(t) = A\cos(\omega_s t), \tag{16}$$

$$v_{x_{\text{out}}}(t) = CA\cos(D\sin(\omega_o t))\cos(\omega_s t), \tag{17}$$

$$v_{y_{\text{out}}}(t) = CA\sin(D\sin(\omega_o t))\cos(\omega_s t), \tag{18}$$

with D being a constant that controls the *depth*, or extent, of the rotation variation and ω_o its angular frequency. To exemplify the spectral composition, $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$ in Eqs. (16)–(18) can be written as a harmonic series using the Bessel functions of the first kind. Using Eqs. (6) and (7) from APPENDIX A.1, they can be written as infinite sums:

$$v_{x_{\text{out}}}(t) = CA\sum_{n=-\infty}^{\infty} J_{2n}(D)\cos\Big((\omega_s+2n\omega_o)t\Big), \tag{19}$$

$$v_{y_{\text{out}}}(t) = CA\sum_{n=-\infty}^{\infty} J_{2n+1}(D)\sin\Big((\omega_s+(2n+1)\omega_o)t\Big). \tag{20}$$

Eqs. (19) and (20) demonstrate that harmonically oscillating around the rotation velocity creates a series of multiples of the modulation frequency ω_o around ω_s, each determined in amplitude by the corresponding Bessel function $J_n(D)$ in dependence of the depth parameter D. In the case of $v_{x_{\text{out}}}(t)$ these are strictly even multiples, while $v_{y_{\text{out}}}(t)$ contains all odd multiples.

Viewed together, the velocity vector $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$ form a spectrum that is similar to FM synthesis [19, 20]:

$$FM = A \sin\Big(\omega_c t + I \sin(\omega_m t)\Big), \tag{21}$$

where the carrier frequency ω_c is modulated by the modulating frequency ω_m by $\pm I$, a constant usually referred to as *modulation index*, expressed as the ratio $I = D/\omega_m$ between the modulation depth D and ω_m. Eq. (21) can similarly be written as a harmonic series using Bessel functions of the first kind [20]:

$$FM = A \sum_{n=-\infty}^{\infty} J_n(I) \sin\Big((\omega_c + n\omega_m)t\Big). \tag{22}$$

While the frequency values are identical in both cases, the main difference lies in the $\pm\frac{\pi}{2}$ phase shift between even and odd multiples of the modulation frequency ω_m in SM synthesis. Furthermore, the even and odd multiples are distributed in $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$ respectively, creating an additional spatial decorrelation effect not present in FM synthesis.

Just as in the case of FM synthesis, the amplitudes of the source in SM synthesis (which is equivalent to the carrier in FM) and side-band components are determined by the Bessel functions J_n. As such, SM synthesis can make use of the same techniques FM synthesis uses to make these undulations simulate the behavior of known acoustic instruments [18, 20]. The constant C can be seen as a weighing control between the pressure component $p_{\text{out}}(t)$ and the velocity components $\mathbf{v}_{\text{out}}(t)$. Thus, SM synthesis enables control between the fundamental frequency ω_s and the overtones.

Conceptually it should be noted that rotational SM synthesis does not modulate the source frequency directly, as FM synthesis does with the carrier. The source being rotated still emits the same constant frequency ω_s, and $p_{\text{out}}(t) = p_{\text{in}}(t)$ still holds true. Instead, the speed of *rotation* is being modulated at a frequency ω_o around a center rotation frequency of $\omega_r = 0$ and all side-bands solely exist in the spatial components $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$.

Inspired by Eq. (21), $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$ in Eqs. (16)–(18) can be altered to see what would happen if a base rotation $\omega_r > 0$ was introduced. This means that $\theta(t) = \omega_r t + D\sin(\omega_o t)$ and the velocity components read

$$v_{x_{\text{out}}}(t) = CA \cos\Big(\omega_r t + D \sin(\omega_o t)\Big) \cos(\omega_s t), \tag{23}$$

$$v_{y_{\text{out}}}(t) = CA \sin\Big(\omega_r t + D \sin(\omega_o t)\Big) \cos(\omega_s t). \tag{24}$$

Similarly to how Eq. (22) was obtained, a Bessel function expansion can be applied to get the harmonic description of $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$:

$$v_{x_{\text{out}}}(t) = \frac{CA}{2} \sum_{n=-\infty}^{\infty} J_n(D)\big[\cos((\omega_s + \omega_r + n\omega_o)t) + \cos((\omega_s - \omega_r - n\omega_o)t)\big], \tag{25}$$

$$v_{y_{\text{out}}}(t) = \frac{CA}{2} \sum_{n=-\infty}^{\infty} J_n(D)\big[\sin((\omega_s + \omega_r + n\omega_o)t) - \sin((\omega_s - \omega_r - n\omega_o)t)\big], \tag{26}$$

meaning that, for $\omega_r > 0$, when compared to Eqs. (19) and (20), an additional spectrum split akin to RM synthesis is introduced. More interestingly though, is that $\omega_r > 0$ introduces the additional odd and even multiples in both $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$ respectively, increasing the similarities between the velocity components.

1.3 Multiple-Source Rotational SM Synthesis in 2D

SM synthesis can be enriched by adding more sources, which are all rotated by the same rotation matrix $\mathbf{R}$. Consider the following for N sources:

$$p_{\text{in}}(t) = p_{\text{out}}(t) = \sum_{n=1}^{N} A_n \cos(\omega_{s_n} t), \tag{27}$$

$$\mathbf{v}_{\text{out}}(t) = \mathbf{R}(\theta(t)) \cdot \sum_{n=1}^{N} \begin{pmatrix} CA_n \cos(\theta_{0_n}) \cos(\omega_{s_n} t) \\ CA_n \sin(\theta_{0_n}) \cos(\omega_{s_n} t) \end{pmatrix}, \tag{28}$$

If $\theta_{0_n} = 0$, such that $C_{x_n} = C\cos(\theta_{0_n}) = C$ and $C_{y_n} = C\sin(\theta_{0_n}) = 0$ for all n, as discussed in SEC. 1.1, then this is equivalent to a complex source made up of all frequencies ω_{s_n} being modified by $\mathbf{R}(\theta(t))$.

However, with SM synthesis, the spatial distribution can additionally be modified by controlling the relative angular distances between each source, with $\theta_{0_n} \neq \theta_{0_m}$ for $n \neq m$ and $n, m \leq N$, to further decorrelate the signal. Thus, unlike what was assumed in SEC. 1.2, $C_{y_n} = 0$ cannot strictly be assumed for all n, and the solutions for rotational SM synthesis become slightly more complicated:

$$v_{x_{\text{out}}}(t) = \sum_{n=0}^{N} \big[C_{x_n} A_n \cos\theta(t) \cos(\omega_{s_n} t) - C_{y_n} A_n \sin\theta(t) \sin(\omega_{s_n} t)\big], \tag{29}$$

$$v_{y_{\text{out}}}(t) = \sum_{n=0}^{N} \big[C_{x_n} A_n \sin\theta(t) \cos(\omega_{s_n} t) + C_{y_n} A_n \cos\theta(t) \sin(\omega_{s_n} t)\big]. \tag{30}$$

Eqs. (29) and (30) can be compared with multiple-carrier FM (MCFM) synthesis for $N > 1$ carriers with angular frequency ω_{c_n}:

$$MCFM = \sum_{n=0}^{N} A_n \sin\Big(\omega_{c_n} t + I_n \sin(\omega_m t)\Big), \tag{31}$$

with individual modulation indices I_n, where ω_{c_0} is the fundamental carrier frequency and all subsequent carrier frequencies $\omega_{c_n} = k_n \omega_{c_0}$ for $n > 0$ with $k_n > 1, k_n \in \mathbb{N}$. All comparisons to MCFM synthesis can equally apply if $\theta_{0_n} = 0$ for all n, as described above. However, similar to how FM synthesis parameterizes the amplitudes A_n and modulation indices I_n [20], all angles θ_{0_k} can also be

regarded as parameters to transition between spatial configurations (e.g., from a point source, $\theta_{0_n} = 0$, to a distributed one).

1.4 Multiple-Rotation SM Synthesis in 2D

Another comparison to traditional synthesis techniques would be to use several modulators. RM and AM are often applied in several iterations to enrich the spectrum of the target sound. However, 2D rotational SM synthesis using constant angular velocities cannot enrich the target spectrum any further by applying successive rotations. This is due to the fact that this operation is equivalent to an addition of the angular parameters, i.e., $\mathbf{R}(\theta_1(t)) \cdot \mathbf{R}(\theta_2(t)) = \mathbf{R}(\theta_1(t) + \theta_2(t))$. By extension, rotating any complex acoustic field that already contains rotating sources only alters their angle of rotation and shifts the resulting side-bands by the sum of both the existing and the added rotation.

Conversely, applying successive harmonically oscillating rotations do produce more side bands. For a single, front-facing source rotated by two successive frequency-modulated rotation matrices, defining $\mathbf{v}_{\text{in}}(t) = (CA\cos(\omega_s t), 0)$ and $\tau_i(t) = \omega_{r_i} + D_i \sin(\omega_{o_i} t)$ for each rotation i,

$$\mathbf{v}_{\text{out}}(t) = \mathbf{R}(\tau_2(t))\mathbf{R}(\tau_1(t))\mathbf{v}_{\text{in}}(t), \tag{32}$$

the formulae found in APPENDIX A.1 can be applied to get

$$\begin{aligned} v_{x_{\text{out}}}(t) = \frac{CA}{2} \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} & J_{n_1}(D_1) J_{n_2}(D_2) \\ & \big[\cos((\omega_s + \omega_{r_1} + \omega_{r_2} + n_1\omega_{o_1} + n_2\omega_{o_2})t) \\ & + \cos((\omega_s - \omega_{r_1} - \omega_{r_2} - n_1\omega_{o_1} - n_2\omega_{o_2})t)\big], \end{aligned} \tag{33}$$

$$\begin{aligned} v_{y_{\text{out}}}(t) = \frac{CA}{2} \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} & J_{n_1}(D_1) J_{n_2}(D_2) \\ & \big[\sin((\omega_s + \omega_{r_1} + \omega_{r_2} + n_1\omega_{o_1} + n_2\omega_{o_2})t) \\ & - \sin((\omega_s - \omega_{r_1} - \omega_{r_2} - n_1\omega_{o_1} - n_2\omega_{o_2})t)\big], \end{aligned} \tag{34}$$

where it can be seen that the base rotation values ω_{r_1} and ω_{r_2} merely add together, forming a two-sided offset to the series of overtones, comparable to Eqs. (25) and (26) in SEC. 1.2.4. However, a second series of side-bands $\pm n_2\omega_{o_2}$ appears around each side-band of the series $\pm n_1\omega_{o_1}$.[1]

This technique is comparable to a similar approach in FM synthesis called *parallel* multi-modulator FM (PMMFM) [18]:

$$PMMFM = A \sin\Big(\omega_c + I_1 \sin(\omega_{m_1} t) + I_2 \sin(\omega_{m_2} t)\Big), \tag{35}$$

the spectral composition of which is given by [21]:

$$\begin{aligned} PMMFM = & \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \\ & J_{n_1}(I_1) J_{n_2}(I_2) \sin\big((\omega_c + n_1\omega_{m_1} + n_2\omega_{m_2})t\big). \end{aligned} \tag{36}$$

[1] For a generalization of Eqs. (32)–(34), see APPENDIX A.2.

For a better comparison, $\omega_{r_1} = 0$ and $\omega_{r_2} = 0$. In this case, similar to Eqs. (19) and (20) in SEC. 1.2.4, certain terms cancel each other out, and the frequencies "distribute themselves" across $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$:

$$\begin{aligned} v_{x_{\text{out}}}(t) = CA \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} & \big[J_{2n_1}(D_1) J_{2n_2}(D_2) \\ & \cos\big((\omega_s + 2n_1\omega_{o_1} + 2n_2\omega_{o_2})t\big) \\ & + J_{2n_1+1}(D_1) J_{2n_2+1}(D_2) \\ & \cos\big((\omega_s + (2n_1+1)\omega_{o_1} + (2n_2+1)\omega_{o_2})t\big)\big], \end{aligned} \tag{37}$$

$$\begin{aligned} v_{y_{\text{out}}}(t) = CA \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} & \big[J_{2n_1}(D_1) J_{2n_1+1}(D_2) \\ & \sin\big((\omega_s + 2n_1\omega_{o_1} + (2n_2+1)\omega_{o_2})t\big) \\ & + J_{2n_1+1}(D_1) J_{2n_2}(D_2) \\ & \sin\big((\omega_s + (2n_1+1)\omega_{o_1} + 2n_2\omega_{o_2})t\big)\big]. \end{aligned} \tag{38}$$

When viewed together, Eqs. (37) and (38) are very similar to Eq. (36), apart from the phase differences; all possible combinations of even and odd multiples of ω_{o_1} and ω_{o_2} are present across $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$.

Another approach, different to PMMFM synthesis is that of *series* multi-modulator FM (SMMFM), where a primary modulator ω_{m_1} is modulated by a secondary modulator ω_{m_2} [18]:

$$SMMFM = A \sin\big(\omega_c + I_1 \sin(\omega_{m_1} t + I_2 \sin(\omega_{m_2} t))\big). \tag{39}$$

This approach can be directly applied to SM synthesis, by defining $\theta(t) = D\sin(\omega_o t + I\sin(\omega_m t))$.

2 SM SYNTHESIS IN 3D

In the 3D case, the same harmonic sound source as in SEC. 1 shall be considered, but this time located on the surface of a sphere of radius $R > 0$ (see Fig. 2). Using spherical coordinates, its position can be defined using the azimuth angle ϕ and inclination angle θ. Similarly to SEC. 1.1, 3D SM synthesis in general will first be defined, and then the focus of these investigations turns toward rotation matrices.

As will be shown, the case of a rotation in a great circle of the sphere is a rotated version of the 2D case already analyzed in SEC. 1.1 and, thus, results in two side-bands in the velocity field. However, the authors shall show how these results are related to arbitrary rotations anywhere on the sphere by including intermediary static rotations as analysis tools. Exploiting the fact that successive rotations can be oriented perpendicularly in 3D, it shall be shown that 3D SM synthesis is capable of producing rather unique spectra when compared to more traditional synthesis methods. Finally, Sec. 2.3.3 will look at how combining multi-rotation 3D SM synthesis with intermediary parametric rotations yields the capability to shape the resulting spectra using only spatial rotation transformations on a single harmonic sound source.

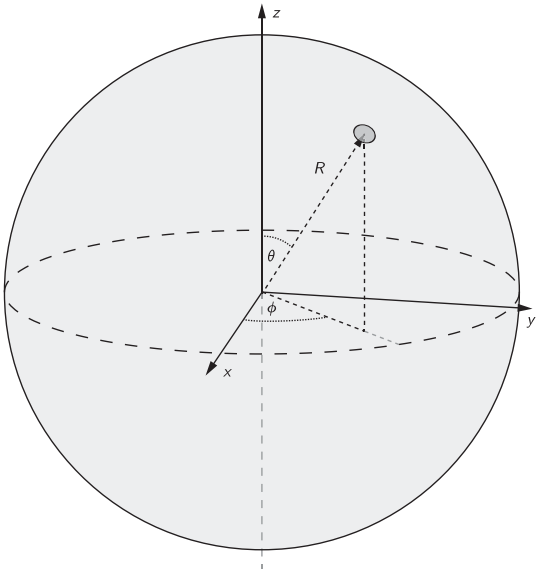


Fig. 2. Sound source on the surface of a sphere of radius R. θ is the inclination angle, and ϕ is the azimuth angle.

2.1 General Definition in 3D

As was done in 2D, the acoustic pressure p and the acoustic velocity vector $\mathbf{v} = (v_x, v_y, v_z)$ shall be considered. Again, it shall also be considered that the output pressure field is the same as the input, $p_{\text{out}}(t) = p_{\text{in}}(t)$. For the remainder of this section, Eq. (1) shall be considered in three dimensions, redefining $\mathbf{v}_{\text{in}}(t) = (C_x p_{\text{in}}(t), C_y p_{\text{in}}(t), C_z p_{\text{in}}(t))$, where C_x, C_y, and C_z again are constants. The constants are usually defined as

$$C_x = C \sin\theta_0 \cos\phi_0, \tag{40}$$

$$C_y = C \sin\theta_0 \sin\phi_0, \tag{41}$$

$$C_z = C \cos\theta_0, \tag{42}$$

where, again, $C = \frac{1}{\rho c}$, as defined in Eq. (1), and θ_0 and ϕ_0 are the initial spherical angles of the sound source (see Fig. 2).

2.2 Rotational SM Synthesis in 3D

As in SEC. 1.2, these investigations will be oriented along the physical results from [16], meaning focus will be on orthogonal rotation matrices. Based on Euler's rotation theorem, it is known that any rotation in 3D space can be seen as a rotation around a unitary vector $\hat{\mathbf{e}} = (e_x, e_y, e_z)$. Due to spherical symmetry of the problem, any orientation for $\hat{\mathbf{e}}$ is equivalent. Choosing thus the xy-plane as the rotation plane, the rotation matrix around the z-axis is received:

$$\mathbf{R}_z(\gamma) = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{43}$$

Just like in SEC. 1.2, a harmonic sound source of angular frequency ω_s shall be considered:

$$p_{\text{in}}(t) = p_{\text{out}}(t) = A\cos(\omega_s t). \tag{44}$$

Expanding Eq. (1) and using Eq. (43) with $\gamma = \phi(t)$, the following is obtained:

$$\mathbf{v}_{\text{out}}(t) = \begin{pmatrix} CA\sin\theta_0\cos(\phi(t)+\phi_0)\cos(\omega_s t) \\ CA\sin\theta_0\sin(\phi(t)+\phi_0)\cos(\omega_s t) \\ CA\cos\theta_0\cos(\omega_s t) \end{pmatrix}. \tag{45}$$

The choices for $\phi(t)$ have equivalent results to what was studied in SEC. 1.2 if the initial source position is horizontally in the front, i.e., $\phi_0 = 0$ and $\theta_0 = \frac{\pi}{2}$. Eq. (45) shows that the choice for $0 \le \theta_0 \le \frac{\pi}{2}$ acts as a non-linear volume control for the overtones produced, with $\theta_0 = \frac{\pi}{2}$ yielding the maximum effect, as it represents the rotation around the great circle. If $\mathbf{R}_z$ was to be replaced with any axis of rotation $\hat{\mathbf{e}}$, then a general requirement to achieve the same maximum effect could be formulated as

$$\mathbf{v}_0 \cdot \hat{\mathbf{e}} = 0, \tag{46}$$

where $\mathbf{v}_0 = (\sin\theta_0\cos\phi_0, \sin\theta_0\sin\phi_0, \cos\theta_0)$.

2.3 Multiple-Rotation SM Synthesis in 3D

Using the example and results of SEC. 2.2, the harmonic sound source from Eq. (44) shall be considered at a starting position $\theta_0 = \frac{\pi}{2}$ and $\phi_0 = 0$, which results in $C_x = C$ and $C_y = C_z = 0$. If a constant angular velocity rotation $\phi(t) = \omega_r t$ in Eq. (45) is considered, then it is known from SEC. 1.4 that adding an additional, on-axis constant angular velocity rotation will not result in any additional side-bands. However, in 3D space, an added rotation can be oriented with respect to a different axis. The result of two successive, off-axis rotations shall be analyzed as a function of angular distance by adding an intermediate constant angle rotation. To achieve the greatest angular distance, this intermediate rotation should ideally be oriented perpendicularly. Let this be a rotation around the x-axis:

$$\mathbf{R}_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}. \tag{47}$$

To generalize, this intermediate rotation full be further rotated using an additional rotation $\mathbf{R}_z(\gamma)$. Redefining $\mathbf{v}_{\text{in}}(t) = (CA\cos(\omega_s t), 0, 0)$, the following series of rotations is received:

$$\mathbf{v}_{\text{out}}(t) = \mathbf{R}_z(\omega_{r_2} t)\mathbf{R}_z(\gamma)\mathbf{R}_x(\alpha)\mathbf{R}_z(\omega_{r_1} t)\mathbf{v}_{\text{in}}(t), \tag{48}$$

which leads to the following set of equations for the velocity field:

$$\begin{aligned} v_{x_{\text{out}}}(t) = CA[& \\ & \frac{1+\cos\alpha}{4}(\cos((\omega_s + \omega_{r_1} + \omega_{r_2})t + \gamma) \\ & \quad + \cos((\omega_s - \omega_{r_1} - \omega_{r_2})t - \gamma)) \\ & + \frac{1-\cos\alpha}{4}(\cos((\omega_s + \omega_{r_1} - \omega_{r_2})t - \gamma) \\ & \quad + \cos((\omega_s - \omega_{r_1} + \omega_{r_2})t + \gamma))], \end{aligned} \tag{49}$$

$$v_{y_{\text{out}}}(t) = CA[\frac{1+\cos\alpha}{4}(\sin((\omega_s+\omega_{r_1}+\omega_{r_2})t+\gamma) - \sin((\omega_s-\omega_{r_1}-\omega_{r_2})t-\gamma)) - \frac{1-\cos\alpha}{4}(\sin((\omega_s+\omega_{r_1}-\omega_{r_2})t-\gamma) - \sin((\omega_s-\omega_{r_1}+\omega_{r_2})t+\gamma))], \quad (50)$$

$$v_{z_{\text{out}}}(t) = \frac{CA\sin\alpha}{2}[\sin((\omega_s+\omega_{r_1})t) - \sin((\omega_s-\omega_{r_1})t)]. \quad (51)$$

As expected, the rotation $\mathbf{R}_z(\gamma)$ acts as an offset for $\mathbf{R}_z(\omega_{r_2}t)$, which would affect the phase in $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$, and $\gamma = 0$ can therefore be considered. For $\alpha = 0$, $v_{z_{\text{out}}}(t) = 0$ and the amplitude coefficient for the terms containing the difference $\omega_{r_1} - \omega_{r_2}$ is zero; as in the case of 2D, i.e., if both rotations are done on the same rotation axis, the number of side-bands remains the same, and the resulting rotation frequency is the sum $\omega_{r_1} + \omega_{r_2}$. Similarly, $\alpha = \pi$ flips the rotation by 180° and inverts the rotation direction. Thus, the combined rotation frequency is $\omega_{r_1} - \omega_{r_2}$ instead.

If α tends toward $\alpha = \frac{\pi}{2}$, however, the amplitude coefficients of the terms containing the sum and difference each reach $\frac{1}{4}$. At this point, the input signal is effectively modulated by two compound frequencies $\omega_{r_1} + \omega_{r_2}$ and $\omega_{r_1} - \omega_{r_2}$, as well as ω_{r_1}, resulting in two additional modulations around ω_s. Considering $\theta_0 = \frac{\pi}{2}$ and $\phi_0 = 0$, it can be easily shown that for $\alpha = \frac{\pi}{2}$ and $\gamma = 0$ in Eq. (48), the series of rotations $\mathbf{R}_z(\omega_{r_2}t)\mathbf{R}_z(0)\mathbf{R}_x(\frac{\pi}{2})\mathbf{R}_z(\omega_{r_1}t)$ can be replaced with an equivalent, simpler sequence of two perpendicular rotations $\mathbf{R}_z(\omega_{r_2}t)\mathbf{R}_y(-\omega_{r_1}t)$ with

$$\mathbf{R}_y(\beta) = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}. \quad (52)$$

Thus, again due to spherical symmetry as evidenced by $\mathbf{R}_z(\gamma)$, it can be generalized that any two perpendicular rotation matrices can be used in order to achieve an additional spectral split akin to multiple-modulator RM synthesis. However, $p_{\text{in}}(t) = p_{\text{out}}(t)$ remains unaltered, meaning that the entire acoustic field will always contain the input signal $\cos(\omega_s t)$. Thus, after two perpendicular rotations in 3D, the spectral output of SM synthesis is directly comparable to neither RM nor AM synthesis. Without loss of generality, the investigations in the following sections can henceforth be limited to using only the elemental rotations $\mathbf{R}_x$, $\mathbf{R}_y$ and $\mathbf{R}_z$, as they correspond to the perpendicular components of the velocity vector.

2.3.1 Chaining Pairwise Perpendicular Constant Angular Velocity Rotations

To further enrich the acoustic field using constant angular velocity rotations, any number of rotations can be applied to the same source, as long as they do not share the same rotation axis with their predecessor. Considering the conclusions above, a chain of N elemental rotations shall be considered:

$$\mathbf{v}_{\text{out}}(t) = \mathbf{R}_{a_N}(\omega_{r_N}t)\ldots\mathbf{R}_{a_1}(\omega_{r_1}t)\mathbf{v}_{\text{in}}(t), \quad (53)$$

where each $\mathbf{R}_{a_n}$, $a_n \in \{x, y, z\}$, $n > 1$, defines the axis of rotation $\hat{\mathbf{e}}_n$, with $a_1 \in \{y, z\}$ and $\hat{\mathbf{e}}_{n+1} \cdot \hat{\mathbf{e}}_n = 0$ for each neighboring pair of rotations.

The spectral composition of Eq. (53) can become quite complex, as it depends on the exact sequence of rotations $\mathbf{R}_x$, $\mathbf{R}_y$, or $\mathbf{R}_z$. However, there are two generalizing observations that help make the analysis somewhat easier:

- Because it is required that rotation $\mathbf{R}_{a_{n+1}}$ must be different than $\mathbf{R}_{a_n}$, a_{n+1} can only be one of two choices: arranged in a cycle, either the *cyclic* element, i.e., in the order shown in Fig. 3, or the *anticyclic* one can be chosen.
- Each elemental rotation has one component of the velocity vector $\mathbf{v}$ that coincides with their axis of rotation. As such, the components of the velocity vector relative to $\mathbf{R}_{a_n}$ can be referred to using the same cyclic convention: one on-axis component $v_{\circ_n}(t)$, one *cyclic* off-axis component $v_{\circlearrowright_n}(t)$, and one *anticyclic* off-axis component $v_{\circlearrowleft_n}(t)$.

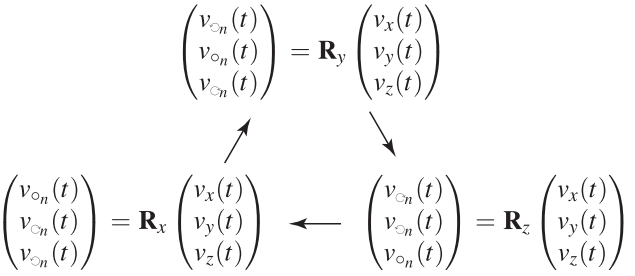


Fig. 3. Order of rotations visualized as a cycle and their corresponding relative on and off-axis components.

For example, for $\mathbf{R}_z$ as the nth rotation, the relative on-axis component $v_{\circ_n}(t)$ is equivalent to the absolute $v_z(t)$, $v_{\circlearrowright_n}(t)$ to $v_x(t)$ and $v_{\circlearrowleft_n}(t)$ to $v_y(t)$.

Furthermore, the sequence of N rotations can be viewed as a sequence of pairwise relations, where $(a_{n-1} \circlearrowright a_n)$ if a_n is the *cyclic* element to a_{n-1} and $(a_{n-1} \circlearrowleft a_n)$ conversely. The result after n rotations can then be given as

$$v_{\circ_n}(t) = \begin{cases} v_{\circlearrowright_{n-1}}(t) \text{if } (a_{n-1} \circlearrowright a_n), \\ v_{\circlearrowleft_{n-1}}(t) \text{if } (a_{n-1} \circlearrowleft a_n), \end{cases} \quad (54)$$

$$v_{\circlearrowright_n}(t) = \begin{cases} v_{\circlearrowright_{n-1}}(t)\cos(\omega_{r_n}t) - v_{\circ_{n-1}}(t)\sin(\omega_{r_n}t) \\ \qquad \text{if } (a_{n-1} \circlearrowright a_n), \\ v_{\circ_{n-1}}(t)\cos(\omega_{r_n}t) - v_{\circlearrowright_{n-1}}(t)\sin(\omega_{r_n}t) \\ \qquad \text{if } (a_{n-1} \circlearrowleft a_n), \end{cases} \quad (55)$$

$$v_{\circlearrowleft_n}(t) = \begin{cases} v_{\circlearrowleft_{n-1}}(t)\sin(\omega_{r_n}t) + v_{\circ_{n-1}}(t)\cos(\omega_{r_n}t) \\ \qquad \text{if } (a_{n-1} \circlearrowright a_n), \\ v_{\circ_{n-1}}(t)\sin(\omega_{r_n}t) + v_{\circlearrowleft_{n-1}}(t)\cos(\omega_{r_n}t) \\ \qquad \text{if } (a_{n-1} \circlearrowleft a_n), \end{cases} \quad (56)$$

for $n \geq 0$, with the initial source given by $v_{\circ_0}(t) = CA\cos(\omega_s t)$ and $v_{\circlearrowright_0}(t) = v_{\circlearrowleft_0}(t) = 0$. Considering that the source in Eq. (53) is equivalent to $a_0 = x$, this can inform

Table 1. Comparison of the potential *maximum* number of frequencies possible with several iterations of multi-rotation 3D SM versus multiple-modulator RM and AM synthesis on a single source/carrier, respectively.

n	0	1	2	3	4	5	6	7	8	9	10	11	12	...
m_{SM_n}	1	3	7	17	45	121	329	897	2,449	6,689	18,273	49,921	136,385	...
m_{RM_n}	1	2	4	8	16	32	64	128	256	512	1,024	1,048	4,096	...
m_{AM_n}	1	3	9	27	81	243	729	2,187	6,561	19,683	59,049	17,7147	531,441	...

the initial choice in Eqs. (54)–(56), as either $(a_0 \circlearrowright a_1)$ or $(a_0 \circlearrowleft a_1)$. The final spectral makeup after N rotations therefore only depends on the sequence of *cyclic* or *anticyclic* relations between neighboring elements a_n.

Using Eqs. (54)–(56), and choosing each ω_n accordingly, the *maximum* number of overtones that could possibly be received can be determined: the ring modulated results of $v_{\circlearrowright_{n+1}}(t)$ and $v_{\circlearrowleft_{n+1}}(t)$ are similar, apart from the phase, meaning that $v_{\circlearrowleft_{n+1}}(t)$ does not add any new frequency values when compared to $v_{\circlearrowright_{n+1}}(t)$ and vice versa. However, $v_{\circ_{n+1}}(t)$ is not ring-modulated by $\omega_{r_{n+1}}$ and ideally does not have any frequencies in common with $v_{\circlearrowright_{n+1}}(t)$ (or $v_{\circlearrowleft_{n+1}}(t)$). Therefore, the result in $v_{\circlearrowright_{n+1}}(t)$ (or $v_{\circlearrowleft_{n+1}}(t)$) ring modulates *all* possible frequency values from rotation n, meaning that the number of frequency values is doubled, if it is considered that ideally no two overtones should overlap. Similarly, $v_{\circ_{n+1}}(t)$ includes the result of a doubling of frequency values from the prior rotation $n-1$. Thus, the *maximum* possible number of side-bands m'_{SM} after n rotations is given by

$$m'_{SM_n} = 2(m'_{SM_{n-1}} + m'_{SM_{n-2}}), \tag{57}$$

for $n \geq 1$, and with $m'_{SM_{-1}} = 0$ and $m'_{SM_0} = 1$, which is a binomial transform of the generalized k-Fibonacci sequence for $k = 2$ and distance $r = 2$ [43, 44]. The *maximum* number of frequencies in the acoustic field includes the pressure component $p_{\text{in}}(t) = p_{\text{out}}(t) = A\cos(\omega_s t)$ and is given by

$$m_{SM_n} = m'_{SM_n} + 1, \tag{58}$$

for $n \geq 1$, and with $m_{SM_0} = 1$, since ω_s is ideally not found in the velocity vector.

In comparison, the frequency combination of multiple-modulator RM synthesis can be determined iteratively using Eq. (11):

$$RM_n = RM_{n-1}\cos(\omega_{m_n}), \tag{59}$$

with $RM_0 = \cos(\omega_c t)$, and the *maximum* number of frequencies m_{RM} after n modulators therefore grows by

$$m_{RM_n} = 2m_{RM_{n-1}}, \tag{60}$$

for $n \geq 1$, and with $m_{RM_0} = 1$. Similarly, multiple-modulator AM synthesis based on Eq. (12) is given by

$$AM_n = AM_{n-1} \cdot \frac{1}{2}(\cos(\omega_{m_n} t) + 1), \tag{61}$$

with $AM_0 = \cos(\omega_c t)$, and the *maximum* number of frequencies m_{AM} after n modulators grows by

$$m_{AM_n} = 3m_{AM_{n-1}}, \tag{62}$$

for $n \geq 1$, and with $m_{AM_0} = 1$. The three approaches are compared in Table 1 for the first values of n, demonstrating that the potential spectral complexity of SM synthesis lies somewhere in between that of multiple-modulator RM and AM synthesis. Therefore, while 2D SM synthesis seems comparable to simple RM and AM synthesis in terms of spectral buildup, SM synthesis sets itself apart in 3D.

2.3.2 Harmonically Oscillating Rotations in 3D

The authors also want to expand their investigations into harmonically oscillating rotations in 3D. In line with the previous methodology, the relationship between a harmonically oscillating rotation and the initial source position will be investigated first. Using $\mathbf{R}_z(\omega_r + D\sin(\omega_o t))$ in the 3D version of Eq. (1) results in

$$\mathbf{v}_{\text{out}}(t) = \begin{pmatrix} CA\sin\theta_0\cos(\omega_r + \phi_0 + D\sin(\omega_o t))\cos(\omega_s t) \\ CA\sin\theta_0\sin(\omega_r + \phi_0 + D\sin(\omega_o t))\cos(\omega_s t) \\ CA\cos\theta_0\cos(\omega_s t) \end{pmatrix}, \tag{63}$$

where, again, ϕ_0 is a simple phase offset to ω_r and θ_0 acts as a nonlinear volume control for the overtones produced, similar to Eq. (45). Thus, henceforth, $\phi_0 = 0$ and $\theta_0 = \frac{\pi}{2}$.

Next, the relationship between two harmonically oscillating rotations shall be examined. Following the same strategy and assumptions that were used in Eq. (48), defining $\tau_i(t) = \omega_{r_i} t + D_i \sin(\omega_{o_i} t)$, and rotating two harmonically oscillating rotations away from each other:

$$\mathbf{v}_{\text{out}}(t) = \mathbf{R}_z(\tau_2(t))\mathbf{R}_z(\gamma)\mathbf{R}_x(\alpha)\mathbf{R}_z(\tau_1(t))\mathbf{v}_{\text{in}}(t), \tag{64}$$

which resolves into the following harmonic description:

$$\begin{aligned} v_{x_{\text{out}}}(t) = \frac{CA}{4} \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} J_{n_1}(D_1) J_{n_2}(D_2) \big(& (1+\cos\alpha) \\ & \big[\cos((\omega_s + \omega_{r_1} + \omega_{r_2} + n_1\omega_{o_1} + n_2\omega_{o_2})t + \gamma) \\ & + \cos((\omega_s - \omega_{r_1} - \omega_{r_2} - n_1\omega_{o_1} - n_2\omega_{o_2})t - \gamma)\big] \\ & + (1-\cos\alpha)\big[\cos((\omega_s - \omega_{r_1} + \omega_{r_2} \\ & - n_1\omega_{o_1} + n_2\omega_{o_2})t + \gamma) + \cos((\omega_s + \omega_{r_1} - \omega_{r_2} \\ & + n_1\omega_{o_1} - n_2\omega_{o_2})t - \gamma)\big]\big), \end{aligned} \tag{65}$$

$$v_{y_{\text{out}}}(t) = \frac{CA}{4} \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} J_{n_1}(D_1) J_{n_2}(D_2) \big((1+\cos\alpha) \big[\sin((\omega_s + \omega_{r_1} + \omega_{r_2} + n_1\omega_{o_1} + n_2\omega_{o_2})t + \gamma) - \sin((\omega_s - \omega_{r_1} - \omega_{r_2} - n_1\omega_{o_1} - n_2\omega_{o_2})t - \gamma) \big] + (1-\cos\alpha) \big[\sin((\omega_s - \omega_{r_1} + \omega_{r_2} - n_1\omega_{o_1} + n_2\omega_{o_2})t + \gamma) - \sin((\omega_s + \omega_{r_1} - \omega_{r_2} + n_1\omega_{o_1} - n_2\omega_{o_2})t - \gamma) \big] \big), \quad (66)$$

$$v_{z_{\text{out}}}(t) = \frac{CA}{2} \sin\alpha \sum_{n=-\infty}^{\infty} J_n(D_1) \big(\sin((\omega_s + \omega_{r_1} + n\omega_{o_1})t) - \sin((\omega_s - \omega_{r_1} - n\omega_{o_1})t) \big). \quad (67)$$

The result for $\alpha = 0$ is consistent with the 2D case [cf. Eqs. (33) and (34)]. However, once the harmonically oscillating rotations are turned away from each other, i.e., α tends toward $\frac{\pi}{2}$, ω_{r_1} and ω_{r_2} do not simply add together but form an additional RM, resulting in a new set of harmonics. Also, while $\mathbf{R}_z(\gamma)$ acts as a simple offset for $\mathbf{R}_z(\omega_{r_2}t + D_2\sin(\omega_{o_2}t))$, its influence on the harmonic distribution is much more pronounced, as evidenced by the result for $\omega_{r_1} = 0$ and $\omega_{r_2} = 0$:

$$v_{x_{\text{out}}}(t) = CA \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \big(\cos(\alpha) J_{2n_1+1}(D_1) \big[\cos(\gamma) J_{2n_2+1}(D_2) \cos((\omega_s + (2n_1+1)\omega_{o_1} + (2n_2+1)\omega_{o_2})t) - \sin(\gamma) J_{2n_2}(D_2) \sin((\omega_s + (2n_1+1)\omega_{o_1} + 2n_2\omega_{o_2})t) \big] + J_{2n_1}(D_1) \big[\cos(\gamma) J_{2n_2}(D_2) \cos((\omega_s + 2n_1\omega_{o_1} + 2n_2\omega_{o_2})t) - \sin(\gamma) J_{2n_2+1}(D_2) \sin((\omega_s + 2n_1\omega_{o_1} + (2n_2+1)\omega_{o_2})t) \big] \big), \quad (68)$$

$$v_{y_{\text{out}}}(t) = CA \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} \big(\cos(\alpha) J_{2n_1+1}(D_1) \big[\cos(\gamma) J_{2n_2}(D_2) \sin((\omega_s + (2n_1+1)\omega_{o_1} + 2n_2\omega_{o_2})t) + \sin(\gamma) J_{2n_2+1}(D_2) \cos((\omega_s + (2n_1+1)\omega_{o_1} + (2n_2+1)\omega_{o_2})t) \big] + J_{2n_1}(D_1) \big[\cos(\gamma) J_{2n_2+1}(D_2) \sin((\omega_s + 2n_1\omega_{o_1} + (2n_2+1)\omega_{o_2})t)) + \sin(\gamma) J_{2n_2}(D_2) \cos((\omega_s + 2n_1\omega_{o_1} + 2n_2\omega_{o_2})t) \big] \big), \quad (69)$$

$$v_{z_{\text{out}}}(t) = \frac{CA}{2} \sum_{n=-\infty}^{\infty} \sin(\alpha) J_n(D_1) \big[\sin((\omega_s + n\omega_{o_1})t) - \sin((\omega_s - n\omega_{o_1})t) \big]. \quad (70)$$

Once again, for $\alpha = 0$ and $\gamma = 0$, the results are equal to the 2D case [cf. Eqs. (37) and (38)]. However, the choice for γ now controls how the different combinations between even and uneven multiples of ω_{o_2} are distributed across $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$. With $0 < \alpha \leq \frac{\pi}{2}$ the harmonics comprised of odd multiples of ω_{o_1} from $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$ are gradually removed and replaced with harmonics in $v_{z_{\text{out}}}(t)$ instead. Because $\mathbf{R}_z(D_1\sin(\omega_{o_1}t))$ is effectively being rotated away from the common rotation axis, the emerging harmonics in $v_{z_{\text{out}}}(t)$ are composed of multiples of ω_{o_1} only.

2.3.3 Spectral Shaping Using Parametric Time-Independent Rotations

In the previous subsections, intermediate constant rotations were used to investigate the effect of the relative orientation between successive rotation matrices, as well as the source itself. It was shown that certain overtones fade out, while others emerge or even flip their phase. For complex multi-rotation SM synthesis scenarios, this effect can be utilized as a parameter for spectral shaping; the fading of overtones affects subsequent modulating rotations. This can already be applied to the initial source position.

To illustrate this, consider the following example using $\mathbf{v}_{\text{in}}(t) = (CA\cos(\omega_s t), 0, 0)$:

$$\mathbf{v}_{\text{out}}(t) = \mathbf{R}_y(\omega_{r_2}t)\mathbf{R}_x(\omega_{r_1}t)\mathbf{R}_z(\gamma)\mathbf{v}_{\text{in}}(t), \quad (71)$$

with $f_s = 2$ kHz, $f_1 = 12$ kHz, and $f_2 = 6$ kHz, where $\omega_{r_n} = 2\pi f_n$, which results in overtones at 4, 8, 10, 14, 16, and 20 kHz in the velocity field:

$$v_{x_{\text{out}}}(t) = \Big(\frac{\cos\gamma}{2} + \frac{\sin\gamma}{4}\Big) \big(\cos(4000\cdot 2\pi t) + \cos(8000\cdot 2\pi t)\big) - \frac{\sin\gamma}{4}\big(\cos(16000\cdot 2\pi t) + \cos(20000\cdot 2\pi t)\big), \quad (72)$$

$$v_{y_{\text{out}}}(t) = \frac{\sin\gamma}{2}\big(\cos(10000\cdot 2\pi t) + \cos(14000\cdot 2\pi t)\big), \quad (73)$$

$$v_{z_{\text{out}}}(t) = \Big(\frac{\sin\gamma}{4} - \frac{\cos\gamma}{2}\Big) \big(\sin(4000\cdot 2\pi t) + \sin(8000\cdot 2\pi t)\big) + \frac{\sin\gamma}{4}\big(\sin(16000\cdot 2\pi t) + \sin(20000\cdot 2\pi t)\big). \quad (74)$$

As can be seen in Eqs. (72)–(74), which overtone is audible at which magnitude depends on the parameter γ. In Fig. 4 the spectra of the components of the velocity field, $\overline{V}_{x_{\text{out}}}(f)$, $\overline{V}_{y_{\text{out}}}(f)$ and $\overline{V}_{z_{\text{out}}}(f)$ are shown as a function of γ in order to better visualize the variations of magnitude. The modulations of the overtones at 4 and 8 kHz have their roots at $\gamma = 2.034$ and $\gamma = 5.176$ in $\overline{V}_{x_{\text{out}}}(f)$, as well as $\gamma = 1.107$ and $\gamma = 4.249$ in $\overline{V}_{z_{\text{out}}}(f)$ respectively, while the modulations of all other overtones have their roots at $\gamma = 0$ and $\gamma = \pi$ instead. $\mathbf{R}_z(\gamma)$ in Eq. (71) is equivalent to a source being rotated horizontally in the *xy*-plane. Therefore, the example shows that the initial source position alters the spectral balance between the overtones and affects the spectral composition of the output velocity field $\mathbf{v}_{\text{out}}(t)$.

To present a more complex example, a harmonic series shall be constructed for 1 kHz that spans the entire human audible hearing range (see Fig. 5) using only constant angular velocity rotation matrices. This can be achieved using four perpendicular rotation matrices, with $f_s = 1$ kHz, $f_1 =$

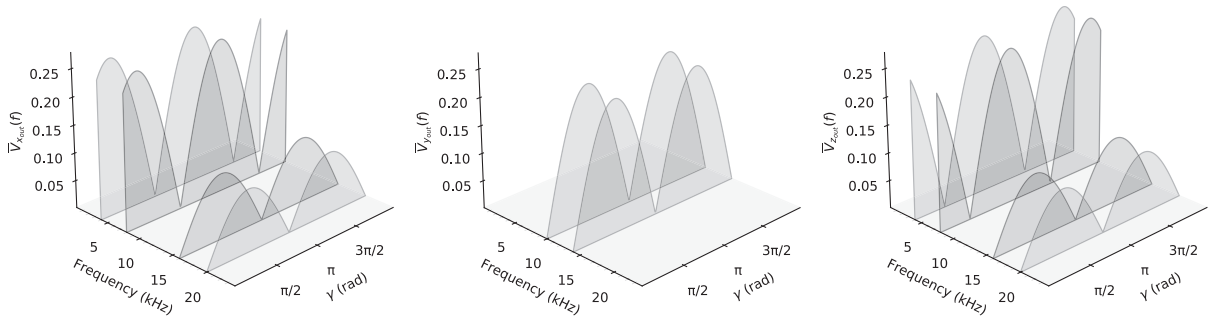


Fig. 4. The spectrum in each component of the velocity vector **v** as a function of γ, as given in Eq. (71), with $f_s = 2$ kHz, $f_1 = 12$ kHz, and $f_2 = 6$ kHz.

12 kHz, $f_2 = 6$ kHz, $f_3 = 2$ kHz, and $f_4 = 1$ kHz. Additionally a parametric constant rotation shall be inserted, and the initial source position on the z-axis shall be defined, meaning that $\mathbf{v}_{\text{in}}(t) = (0, 0, CA\cos(\omega_s t))$:

$$\mathbf{v}_{\text{out}}(t) = \mathbf{R}_y(\omega_{r_4} t)\mathbf{R}_x(\alpha)\mathbf{R}_z(\omega_{r_3} t)\mathbf{R}_y(\omega_{r_2} t)\mathbf{R}_x(\omega_{r_1} t)\mathbf{v}_{\text{in}}(t). \tag{75}$$

For Eq. (75), the even and odd harmonics are distributed between $v_{x_{out}}(t)$ (or $v_{z_{out}}(t)$) and $v_{y_{out}}(t)$ respectively, as can be seen in Fig. 5. The variations in magnitude differ quite extensively between many overtones. Thus altering α in Eq. (75) results in numerous combinations of different weightings between overtones, resulting in a multitude of possible timbres using only a combination of rotations on a single harmonic source.

Finally, the authors want to show that this spectral shaping technique can also be used with harmonically oscillating rotations in 3D. For this example, they chose a simple scenario, similar to what was used in SEC. 2.3.2, using two rotations $\mathbf{R}_z$ with an intermediate, parametric constant rotation. Defining $\tau_i(t) = D_i \sin(\omega_{o_i} t)$ in this case, as well as $\mathbf{v}_{\text{in}}(t) = (CA\cos(\omega_s t), 0, 0)$, the following is obtained:

$$\mathbf{v}_{\text{out}}(t) = \mathbf{R}_z(\tau_2(t))\mathbf{R}_x(\alpha)\mathbf{R}_z(\tau_1(t))\mathbf{v}_{\text{in}}(t). \tag{76}$$

The harmonic series depicted in Fig. 6 results from defining $f_s = 1$ kHz, $f_1 = 7$ kHz, and $f_2 = 2$ kHz as well as $D_1 = 2.5$ and $D_2 = 2.5$ in Eq. (76).

In this case, most harmonics are contained in $v_{x_{\text{out}}}(t)$ and $v_{y_{\text{out}}}(t)$, with only some even harmonics represented in $v_{z_{\text{out}}}(t)$ as well. It can be seen that the influence of α only extends to the even harmonics in this example, having roots at $\alpha = \pm\frac{\pi}{2}$, while the odd harmonics remain unaffected. While this turns α into a spectral parameter controlling the balance of even harmonics against odd harmonics, some select even harmonics do reemerge in $v_{z_{\text{out}}}(t)$, since the roots for the overtones in $v_{z_{\text{out}}}(t)$ are located at $\alpha = 0$ and $\alpha = \pi$ instead. However, unlike altering the depth parameters D_i (as is a common technique with FM synthesis), where most overtones would follow complex undulations that would also affect the magnitudes of odd harmonics, α synchronously reduces or increases the magnitudes of *only* even harmonics, giving it a slightly different application and justifying using both approaches for different goals in spectral shaping. In APPENDIX A.3 the interested reader may find additional information on some sound examples that are similar to those shown in this subsection and that were prepared to provide a more tangible illustration.

3 IMPLEMENTING SM SYNTHESIS

Rotating physical sources around a listener at the velocities described here could be considered impractical. Therefore, the authors suggest implementing a simulation of a rotating source using a multichannel loudspeaker array. An example of a simulation of a source rotating at high velocities using a circular and spherical secondary source array by the means of Wavefield Synthesis [45, 46] is demonstrated in [16]. However, simulating the sound field of a rotating source at high fidelity using Wavefield Synthesis may also not be considered particularly practical in most cases.

Instead, the authors have implemented SM synthesis using a virtualization in the spherical harmonic domain and

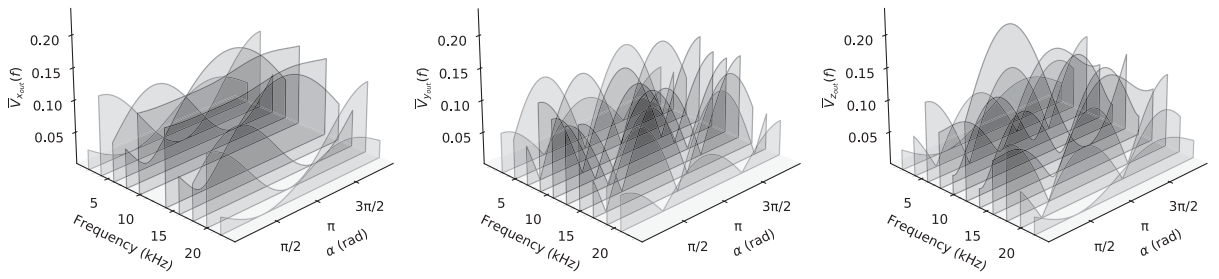


Fig. 5. The spectrum in each component of the velocity vector **v** as a function of α, as given in Eq. (75), with $f_s = 1$ kHz, $f_1 = 12$ kHz, $f_2 = 6$ kHz, $f_3 = 2$ kHz, and $f_4 = 1$ kHz.

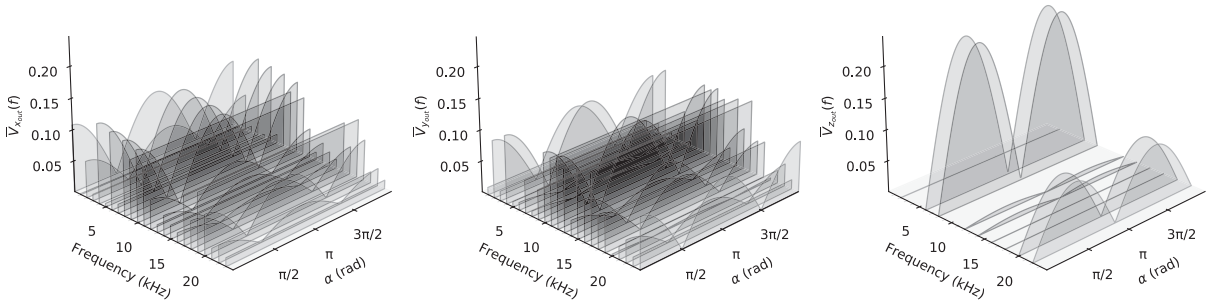


Fig. 6. The spectrum in each component of the velocity vector **v** as a function of α, as given in Eq. (76), with $f_s = 1$ kHz, $f_{m_1} = 7$ kHz, $f_{m_2} = 2$ kHz, and $D_1 = D_2 = 2.5$.

decoding the simulated velocity field with the help of Ambisonics [3, 4, 47]. Because Ambisonics decomposes the acoustic field as a series of spherical harmonics from a central reference point, it is relatively straight forward to implement SM synthesis so that it can be reconstructed with most speaker layouts or even in binaural over headphones.

It has been shown in [47] that the acoustic pressure at a position **r**, generated by sources outside of the listening area, can be approximated by the following truncated series:

$$p(\mathbf{r}) = \sum_{m=0}^{M} i^m J_m(kr) \sum_{\substack{0\le n\le m \\ \sigma=\pm 1}} B_{mn}^{\sigma} Y_{mn}^{\sigma}(\phi, \theta), \tag{77}$$

for the radius r and wave number $k = 2\pi f/c$, where Y_{mn}^{σ} represents the spherical harmonic function for order m and degree n and B_{mn}^{σ} are weighting coefficients and are associated with the spherical harmonics expansion. For a plane wave, this is given as [47, 48]:

$$B_{mn}^{\sigma} = p_0 Y_{mn}^{\sigma}(\phi, \theta), \tag{78}$$

with p_0 being the pressure at the source. Eq. (78) shows that a source may be encoded simply using gain factors on a source signal $s(t)$.

For first-order Ambisonics, only the terms corresponding to the zeroth and first order are considered. An encoding technique of a virtual sound source in first-order Ambisonics can thus be derived from Eq. (78) as follows [49]:

$$W(t) = a_0 s(t), \tag{79}$$

$$X(t) = a_1 s(t) \cos\phi \sin\theta, \tag{80}$$

$$Y(t) = a_1 s(t) \sin\phi \sin\theta, \tag{81}$$

$$Z(t) = a_1 s(t) \cos\theta, \tag{82}$$

where a_0 and a_1 are normalization factors for the zeroth and first order, respectively, and depend on the standard adhered to. Thus, if the input signal is set to $s(t) = cos(\omega_s t)$ in Eqs. (79)–(82) and a constant rotation is chosen for $\phi = \omega_{r_1} t$ and $\theta = \omega_{r_2} t$ along both the azimuth and inclination, the same result as demonstrated in [16] is achieved, and the analysis done in SECS. 1 and 2 applies.

The four components of first-order Ambisonics can be mapped to the pressure and velocity components of the acoustic field at the central listening position, respectively [3, 47, 49, 50]. Moreover, though, the spherical harmonic representation can be regarded as an ideal free field in which SM synthesis may be implemented to produce the desired acoustic field. An Ambisonic signal can therefore be rotated by applying one of the rotation matrices given in SECS. 2.2 and 2.3 to the first-order components [51]. Also, it should be noted that several sources encoded in Eqs. (79)–(82) may simply be summed together into a single Ambisonic signal to achieve more complex scenes. Referring back to SEC. 1.3, this implies that not only can a simple source be rotated, but an entire periphonic scene of arbitrary complexity.

Therefore, for a practical implementation of rotational SM synthesis, one can directly take advantage of this representation and implement time-varying rotation matrices using input control signals $s_i(t)$ that represent the angles of rotation. For example, a unipolar sawtooth wave can be easily mapped to the angle of rotation, representing a full revolution at the sawtooth wave's frequency. Likewise, a sinusoidal input signal with varying amplitude, using the same mapping as above, would represent a harmonic oscillation around a fixed angle $\omega_r = 0$, where the input control signal's amplitude represents the angular distance D of the oscillation—or *depth* of the modulation. Furthermore, the separate normalization factors for each order can be used to overcome the level difference encountered in [16] by using $a_1' = ca_1$ with a correction factor c, similarly to how $CA = 1$ was set in SEC. 1.2.2 to achieve a better comparison with AM and FM synthesis.

To better illustrate this, an example for a z-axis rotation is given in Fig. 7, using a similar style as was used in [20]. A sine tone generator is used as an input sound in this example. The input frequency is given as f_s at full scale, to be able to scale it later separately by the factors a_0 and a_1 for each spherical harmonic order. A sawtooth generator generates a unipolar control signal at a frequency f_r at full scale. The cos and sin modules accept the control signal and map the full scale range to the range 2π. The resulting coefficients are used to scale the input signal and convert it into the spherical harmonic domain. The final signal, comprising the components W, X, and Y, can then be combined with any other signal already encoded in Ambisonics.

Finally, the spherical harmonic domain will need to be decoded. There are several techniques to decode an Ambisonic signal to a loudspeaker array or headphones, depending on the selected optimization criterion among other factors [4, 52]. Nevertheless, a general decoder to obtain the

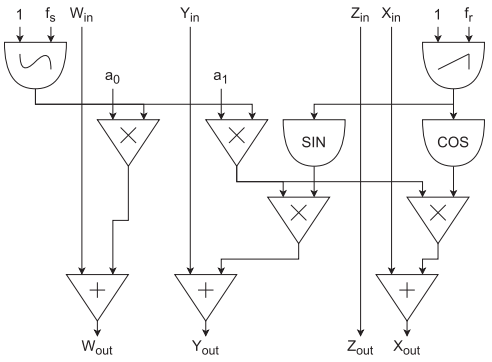


Fig. 7. An implementation diagram for a sound generated by SM synthesis using a z-axis rotation in the first-order spherical harmonic domain. Note that the channel order is given in the common Ambisonics Channel Number format [64].

loudspeaker signals $\mathbf{x}$ for a finite number of loudspeakers L is given as [52]

$$\mathbf{x} = \mathbf{D}\,\mathrm{diag}\{\mathbf{a}_N\}\chi_N, \tag{83}$$

where $\mathbf{D}$ is a sampling decoder, dependent on the spherical harmonics $Y_{nm}(\phi_l, \theta_l)$, with ϕ_l and θ_l describing the position for a loudspeaker l; $\mathbf{a}_N$ is the side-lobe suppressing weights; and χ_N is the encoded Ambisonic signal for N harmonics. Thus, if the loudspeaker positions ϕ_l and θ_l remain constant, the operation is linear and the frequency content found in the spherical harmonic representation is not altered.

The spatial reconstruction of the acoustic field using a finite set of points over a sphere, however, may suffer from spatial artifacts and can limit the sweet spot area. A common approach to mitigate these is to use higher-order Ambisonics, which the authors plan to investigate more thoroughly in the context of SM synthesis in future work. The decoding techniques presented in [4, 52] should serve as a starting point, as this is an issue in spatial audio and Ambisonics in general.

4 CONCLUSION

In this paper, the authors presented a method for audio synthesis in the acoustic velocity field using spatial transformation matrices. They looked at the various secondary effects caused by different rotation behaviors of a sound source in both 2D and 3D and compared them to more traditional synthesis approaches.

It has been shown that the spectral result in the velocity field using a single constant angular rotation in 2D produces a spectrum split akin to RM synthesis. Yet, because the pressure component maintains the input signal, it can be said that the spectral result in the entire acoustic field is more comparable to AM synthesis. Increasing the spectral complexity by using successive constant angular rotations is not possible in 2D, due to the fact that applying successive rotations is the same as merely adding their angles of rotation.

An additional interesting case that was shown are harmonically oscillating rotations in 2D, where it was found that the spectral result in the velocity field is comparable to FM synthesis instead. If no base rotation is present and the rotation oscillates around a center direction, then the even and odd multiples are distributed in the x and y components of the velocity field, respectively, increasing the spatial decorrelation. Moreover, chaining successive harmonically oscillating rotations is comparable to PMMFM, meaning that the spectrum can similarly be enriched by applying more harmonically oscillating rotations to the same source.

It was found that certain limitations in 2D could be overcome in 3D. Particularly, it is possible to split the spectrum found in the velocity field after n constant angular rotations by applying an additional rotation with a rotation axis (ideally) perpendicular to the previous one. The spectral result is unique to 3D SM synthesis, compared to traditional synthesis methods. An iterative method to determine the spectral potential after n pairwise perpendicular rotations has shown that the potential spectral complexity of 3D SM synthesis lies in between the potential of multi-modulator RM and AM synthesis.

Moreover, it was shown that rotating the rotation axes of multi-rotation SM synthesis against each other has the potential to be a parameter for spectral shaping. Since the spectral result of a 3D multi-rotation system depends on the relative orientations of successive rotation axes, both constant angular velocity rotations as well as harmonically oscillating rotations can be rotated against each other to produce varying weightings of the overtones produced.

Finally, a practical implementation of SM synthesis within the spherical harmonic domain using Ambisonics was outlined in SEC. 3. Linear ramps, like a sawtooth wave, can be used as input control signals to control the speed at which the source rotates around the circle. Moreover, a sinusodial control signal then represents a harmonically oscillating movement, where the amplitude of the control signal defines the depth. Using Ambisonics, the result of SM synthesis can reproduced over a loudspeaker array without any loss in spectral content. However, spatial artifacts due to the truncation of the Ambisonic order and the finite number of loudspeakers in the array may cause spatial artifacts in the reproduced sound.

4.1 Beyond Rotational SM Synthesis

This paper has focused on orthogonal rotation matrices as one solution to SM synthesis. However, the general definition has been specifically chosen to encompass any spatial transformation matrix. Considering that SM synthesis will most likely always be implemented in virtual environments, this can also include non-orthogonal matrices.

Thus, future work may also include the analysis of alternate solutions for SM synthesis. One inspiration could be Ambisonics itself, where many spatial transformations have already been well established. For example, mirroring [51] uses a diagonal matrix $\mathbf{T} = \mathrm{diag}\{\mathbf{c}\}$ with a corresponding sign change sequence $\mathbf{c}$ and could be adapted to use a sequence of continuous, periodic sign changes $\tilde{\mathbf{c}} = (\cos(\omega_x t + \phi_x), \cos(\omega_y t + \phi_y), \cos(\omega_z t + \phi_z))$, which can

be considered a non-orthogonal generalization to rotational SM synthesis.

4.2 Inherently Spatial Sounds

Although SM synthesis moves a source along a circle around a listener, the speed of the movement is too high to be perceived [15], and the synthesis result is instead perceived as immersive and surrounding. Therefore, frequency distribution and decorrelation are an inevitable part of rotational SM synthesis. Unlike usual approaches to 3D sound reconstruction and synthesis, the implementation in SEC. 3 demonstrates that SM synthesis does not reconstruct point sources at certain positions but instead creates *inherently spatial* sounds.

Because all sound synthesis happens in the velocity field by the means of spatial transformations, any effect done by SM synthesis theoretically disappears if analyzed at the center of rotation via the pressure field exclusively. As such, *space* plays an integral part in the use and application of SM synthesis, and the authors therefore coined the term *inherently spatial* to denote that sounds created by this or similar methods exist not only in but specifically because of the use of *space*. Accordingly, SM synthesis differs from most previous spatial sound synthesis methods [22–30] in that they first synthesize a sound and apply spatialization on the synthesized result afterward, as opposed to achieving synthesis through spatialization. Future work will encompass the analysis of *inherently spatial* sounds from two perspectives: first, from a musicological perspective, i.e., what are the compositional implications of *inherently spatial* sounds, and, second, perceptually, in how far *inherently spatial* sounds affect the perception of *space*.

Composition happens both on a macroscopic as well as a microscopic scale [53], and it may be argued that sculpting a sound using SM synthesis also sculpts a space. Thus, the authors identify SM synthesis as a potential tool to aid and reflect on musicological research in spatial music, that is, as a tool for both the technical analysis of existing works as well as the contemplation on the potential of spatial music. For his composition *Sirius*, Stockhausen used a modified version of his *Rotationstisch*, with which he could quickly pan sounds by mechanically rotating a loudspeaker at the center of a ring of eight microphones reproduced through an octophonic loudspeaker setup thereafter [54]. In an interview conducted in 1977, he described perceiving sounds that were rotated "fast" by this technique as both "standing still" while also "vibrating" [55].

Previous research has already shown that the auditory spatial sensation of immersion can be further subdivided and refined, which could help analyze the output of SM synthesis in this aspect. For example, in [56] it has been argued in favor of a distinction between the sensation of 2D *envelopment* and 3D *engulfment*, the latter being unique to auditory height perception, and [57] has expanded this research to compare different methods of decorrelation using qualitative perceptual gradients of the above. Composers of spatial music have also attempted to classify spatial impressions as a matter of perspectives [8, 58], while others have described spatial sounds and metaphors as a matter of *spatial texture* [59, 60]. The fact that the perception of frequency and spatial location are not independent [61–63] hints at the interplay of frequency and space and the complications regarding this matter. Thus, concerning its perception, the analysis SM synthesis can be approached from another angle entirely in future work: in how far does SM synthesis, as well as other spatial synthesis methods, not only create sounds spectrally but also create sensation of *a space*, as perceived through audition.

4.3 Future Work and Commercial Viability

Additionally to the points already mentioned in this section, future work will also have to focus on practical implementations of SM synthesis. The approach discussed in SEC. 3 describes only a basic implementation of the raw, bare-bones algorithm. However, a commercially viable synthesizer requires work on additional details that go beyond the basic synthesis process, as described here. This may include effects and other modifications as well as studies on the user experience of a potential product, in both software and hardware. In particular, these effects and processes most likely need to be adapted to the 3D spatial domain. Furthermore, as addressed, e.g., in [27], user experience studies would extend to the interface of such a product, as it may require careful consideration of which parameter to expose, while others may be grouped into meta-parameters, since the user is presented with a different approach to sound synthesis that strictly creates results in the 3D spatial domain.

5 ACKNOWLEDGMENT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 957185. The authors want to thank Umut Sayın for a careful reading of the manuscript. Also, the authors are grateful for the reviewer's comments, which have contributed toward improving the quality of the paper.

6 REFERENCES

[1] F. Rumsey, *Spatial Audio* (Routledge, New York, NY, 2001), 1st ed. https://doi.org/10.4324/9780080498195.

[2] V. Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning," *J. Audio Eng. Soc.*, vol. 45, no. 6, pp. 456–466 (1997 Jun.).

[3] M. A. Gerzon, "Periphony: With-Height Sound Reproduction," *J. Audio Eng. Soc.*, vol. 21, no. 1, pp. 2–10 (1973 Feb.).

[4] J. Daniel, *Représentation de Champs Acoustiques, Application à la Transmission et à la Reproduction de Scènes Sonores Complexes dans un Contexte Multimédia*, Ph.D. thesis, University of Paris VI, Paris, France (2001 Jul.).

[5] J. R. Gonzalez, "Sound Hyperreality in Popular Music: On the Influence of Audio Production in Our Sound Expectations," in E. Encabo (Ed.), *Sound in Motion: Cinema, Videogames, Technology and Audiences* (Cambridge Scholars Publishing, Newcastle upon Tyne, UK, 2018).

[6] K. Stockhausen, "Musik im Raum," in *Texte zur Elektronischen und Instrumentalen Musik, Band 1: Aufsätze 1952-1962 zur Theorie des Komponierens* (M. DuMont Schauberg, Cologne, Germany, 1963).

[7] J. M. Chowning, "The Simulation of Moving Sound Sources," *J. Audio Eng. Soc.*, vol. 19, no. 1, pp. 2–6 (1971 Jan.).

[8] D. Smalley, "Spectromorphology: Explaining Sound-Shapes," *Organ, Sound*, vol. 2, no. 2, pp. 107–126 (1997 Jul.). https://doi.org/10.1017/S1355771897009059.

[9] N. Barrett, "Spatio-Musical Composition Strategies," *Organ. Sound*, vol. 7, no. 3, pp. 313–323 (2002 Jun.). https://doi.org/10.1017/S1355771802003114.

[10] M. A. J. Baalman, "Spatial Composition Techniques and Sound Spatialisation Technologies," *Organ. Sound*, vol. 15, no. 3, pp. 209–218 (2010 Oct.). https://doi.org/10.1017/S1355771810000245.

[11] A. Franck, A. Gräfe, T. Korn, and M. Strauß, "Reproduction of Moving Sound Sources by Wave Field Synthesis: An Analysis of Artifacts," in *Proceedings of the 32nd AES International Conference: DSP For Loudspeakers* (2007 Sep.), paper 10.

[12] J. Ahrens and S. Spors, "Reproduction of Moving Virtual Sound Sources With Special Attention to the Doppler Effect," presented at the *124th Convention of the Audio Engineering Society* (2008 May), paper 7363.

[13] M. J. Morrell and J. D. Reiss, "Inherent Doppler Properties of Spatial Audio," presented at the *129th Convention of the Audio Engineering Society* (2010 Nov.), paper 8278.

[14] B. B. Boren and M. Ericson, "Motion Simulation in the Environment for Auditory Research," *Proc. Mtgs. Acoust.*, vol. 14, no. 1, paper 015006 (2011 Oct.). https://doi.org/10.1121/1.4704670.

[15] F.-X. Féron, I. Frissen, J. Boissinot, and C. Guastavino, "Upper Limits of Auditory Rotational Motion Perception," *J. Acoust. Soc. Am.*, vol. 128, no. 6, pp. 3703–3714 (2010 Dec.). https://doi.org/10.1121/1.3502456.

[16] T. Schmele and A. Garriga, "Sound Spectrum Modulation Generated by Circularly Moving Sound Sources," presented at the *154th Convention of the Audio Engineering Society* (2023 May), paper 10657.

[17] H. S. Black, *Modulation Theory* (Van Nostrand, Princeton, NJ, 1953).

[18] C. Roads, *Computer Music Tutorial* (MIT Press, Cambridge, MA, 1996).

[19] J. R. Carson, "Notes on the Theory of Modulation," *Proc. Inst. Electr. Eng.*, vol. 10, no. 1, pp. 57–64 (1922 Feb.). https://doi.org/10.1109/JRPROC.1922.219793.

[20] J. M. Chowning, "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation," *J. Audio Eng. Soc.*, vol. 21, no. 7, pp. 526–534 (1973 Sep.).

[21] M. Le Brun, "A Derivation of the Spectrum of FM With a Complex Modulating Wave," *Comput. Music J.*, vol. 1, no. 4, pp. 51–52 (1977 Nov.).

[22] D. Kim-Boyle, "Sound Spatialization With Particle Systems," in *Proceedings of the 8th International Conference on Digital Audio Effects (DAFx)*, pp. 65–68 (Madrid, Spain) (2005 Sep.).

[23] D. Kim-Boyle, "Spectral and Granular Spatialization With Boids," in *Proceedings of the 32nd International Computer Music Conference (ICMC)*, pp. 139–142 (New Orleans, LA) (2006 Nov.).

[24] S. Wilson, "Spatial Swarm Granulation," in *Proceedings of the 34th International Computer Music Conference (ICMC)*, paper 127 (Belfast, Northern Ireland) (2008 Aug.).

[25] E. Bates, *The Composition and Performance of Spatial Music*, Ph.D. thesis, Trinity College Dublin, Dublin, Ireland (2009 Aug.).

[26] S. James and C. Hope, "2D and 3D Timbral Spatialisation: Spatial Motion, Immersiveness, and Notions of Space," in *Proceedings of the 39th International Computer Music Conference (ICMC)*, pp. 77–84 (Perth, Australia) (2013 Aug.).

[27] D. Kim-Boyle, "Spectral Spatialization – An Overview," in *Proceedings of the 34th International Computer Music Conference (ICMC)*, paper 86 (Belfast, Northern Ireland) (2008 Aug.).

[28] R. H. Torchia and C. Lippe, "Techniques for Multi-Channel Real-Time Spatial Distribution Using Frequency-Domain Processing," in *Proceedings of the International Conference on New Interfaces for Musical Expression (NIME)*, pp. 116–119 (Hamamatsu, Japan) (2004 Jun.).

[29] A. Mills and R. C. de Souza, "Gestural Sounds by Means of Wave Terrain Synthesis," presented at the *VI Brazilian Symposium on Computer Music* (Rio de Janeiro, Brazil) (1999 Jul.).

[30] S. James, "Spectromorphology and Spatiomorphology of Sound Shapes: Audio-Rate AEP and DBAP Panning of Spectra," in *Proceedings of the 41st International Computer Music Conference (ICMC)*, pp. 279–285 (Denton, TX) (2015 Sep.).

[31] R. L. Jenison, M. F. Neelon, R. A. Reale, and J. F. Brugge, "Synthesis of Virtual Motion in 3D Auditory Space," in *Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1096–1100 (Hong Kong, China) (1998 Nov.). https://doi.org/10.1109/IEMBS.1998.747062.

[32] R. McGee, "Spatial Modulation Synthesis," in *Proceedings of the 41st International Computer Music Conference (ICMC)*, pp. 246–249 (Denton, TX) (2015 Sep.).

[33] R. M. McGee, *Scanning Spaces: Paradigms for Spatial Sonification and Synthesis*, Ph.D. thesis, University of California, Santa Barbara, CA (2015 Dec.).

[34] T. Schmele and I. Gómez, "Exploring 3D Audio for Brain Sonification," in *Proceedings of the 18th International Conference on Auditory Display*, pp. 22–28 (Atlanta, GA) (2012 Jun.). https://doi.org/10.13140/2.1.2312.0009.

[35] T. Schmele and J. J. Lopez, "Comparisons Between VBAP and WFS Using Spatial Sound Synthesis," presented

at the *153th Convention of the Audio Engineering Society* (2022 Oct.), paper 10622.

[36] T. Schmele and N. Reppel, "Emulating Vector Base Amplitude Panning Using Panningtable Synthesis," presented at the *155th Convention of the Audio Engineering Society* (2023 Oct.), paper 119.

[37] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization* (MIT Press, Cambridge, MA, 1997).

[38] G. Potard and I. Burnett, "Decorrelation Techniques for the Rendering of Apparent Sound Source Width in 3D Audio Displays," in *Proceedings of the 4th International Conference on Digital Audio Effects (DAFx)*, pp. 280–284 (Naples, Italy) (2004 Oct.).

[39] M. Weger, G. Marentakis, and R. Höldrich, "Auditory Perception of Spatial Extent in the Horizontal and Vertical Plane," in *Proceedings of the 19th International Conference on Digital Audio Effects (DAFx)*, pp. 301–308 (Brno, Czech Republic) (2016 Sep.).

[40] J. Ahrens and S. Spors, "Two Physical Models for Spatially Extended Virtual Sound Sources," presented at the *131st Convention of the Audio Engineering Society* (2011 Oct.), paper 8483.

[41] G. S. Kendall, "The Decorrelation of Audio Signals and Its Impact on Spatial Imagery," *Comput. Music J.*, vol. 19, no. 4, pp. 71–87 (1995 Dec.).

[42] S. James, "A Classification of Multi-Point Spectral Sound Shapes," in *Proceedings of the Sonic Environments Conference*, pp. 57–64 (Brisbane, Australia) (2016 Jul.).

[43] S. Falcón, "Binomial Transform of the Generalized k-Fibonacci Numbers," *Commun. Math. Appl.*, vol. 10, no. 3, pp. 643–651 (2019 Sep.). https://doi.org/10.26713/cma.v10i3.1221.

[44] OEIS, "A002605," http://oeis.org/A002605 (accessed June 5, 2021).

[45] R. Rabenstein and S. Spors, "Spatial Aliasing Artifacts Produced by Linear and Circular Loudspeaker Arrays Used for Wave Field Synthesis," presented at the *120th Convention of the Audio Engineering Society* (2006 May), paper 6711.

[46] J. Ahrens, R. Rabenstein, and S. Spors, "The Theory of Wave Field Synthesis Revisited," presented at the *124th Convention of the Audio Engineering Society* (2008 May), paper 7358.

[47] J. Daniel, S. Moreau, and R. Nicol, "Further Investigations of High-Order Ambisonics and Wavefield Synthesis for Holophonic Sound Imaging," presented at the *114th Convention of the Audio Engineering Society* (2003 Mar.), paper 5788.

[48] P. M. Morse and K. U. Ingard, *Theoretical Acoustics* (Princeton University Press, Princeton, NJ, 1968).

[49] N. Rozenn, "Sound Field," in A. Roginska and P. Geluso (Eds.), *Immersive Sound: The Art and Science of Binaural and Multi-Channel Audio*, pp. 276–310 (Routledge, New York, NY, 2018).

[50] F. Zotter and M. Frank, "XY, MS, and First-Order Ambisonics," in *Ambisonics: A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*, Springer Topics in Signal Processing, vol. 19, pp. 1–22 (Springer, Cham, Switzerland, 2019). https://doi.org/10.1007/978-3-030-17207-7_1.

[51] F. Zotter and M. Frank, "Signal Flow and Effects in Ambisonic Productions," in *Ambisonics: A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*, Springer Topics in Signal Processing, vol. 19, pp. 99–129 (Springer, Cham, Switzerland, 2019). https://doi.org/10.1007/978-3-030-17207-7_5.

[52] F. Zotter and M. Frank, "Ambisonic Amplitude Panning and Decoding in Higher Orders," in *Ambisonics: A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality*, Springer Topics in Signal Processing, vol. 19, pp. 53–98 (Springer, Cham, Switzerland, 2019). https://doi.org/10.1007/978-3-030-17207-7_4.

[53] A. Di Scipio, "Compositional Models in Xenakis's Electroacoustic Music," *Perspect. New Music*, vol. 36, no. 2, pp. 201–243 (1998 Summer). https://doi.org/10.2307/833529.

[54] M. Clarke and P. Manning, "The Influence of Technology on the Composition of Stockhausen's *Octophonie*, With Particular Reference to the Issues of Spatialisation in a Three-Dimensional Listening Environment," *Organ. Sound*, vol. 13, no. 3, pp. 177–187 (2008 Nov.). https://doi.org/10.1017/S1355771808000277.

[55] D. Felder and K. Stockhausen, "An Interview With Karlheinz Stockhausen," *Perspect. New Music*, vol. 16, no. 1, pp. 85–101 (1977 Autumn-Winter). https://doi.org/10.2307/832850.

[56] R. Sazdov, G. Paine, and K. Stevens, "Perceptual Investigation Into Envelopment, Spatial Clarity, and Engulfment in Reproduced Multi-Channel Audio," presented at the *Proceedings of the 31st AES International Conference: New Directions in High Resolution Audio* (2007 Jun.), paper 25.

[57] H. Lynch and R. Sazdov, "A Perceptual Investigation Into Spatialization Techniques Used in Multichannel Electroacoustic Music for Envelopment and Engulfment," *Comput. Music J.*, vol. 41, no. 1, pp. 13–33 (2017 Mar.). https://doi.org/10.1162/COMJ_a_00401.

[58] D. Smalley, "Space-Form and the Acousmatic Image," *Organ. Sound*, vol. 12, no. 1, pp. 35–58 (2007 Apr.). https://doi.org/10.1017/S1355771807001665.

[59] K. L. Hagan, "Textural Composition: Aesthetics, Techniques, and Spatialization for High-Density Loudspeaker Arrays," *Comput. Music J.*, vol. 41, no. 1, pp. 34–45 (2017 Mar.). https://doi.org/10.1162/COMJ_a_00395.

[60] E. Nystrom, "Morphology of the Amorphous: Spatial Texture, Motion and Words," *Organ. Sound*, vol. 22, no. 3, pp. 336–344 (2017 Nov.). https://doi.org/10.1017/S1355771817000498.

[61] E. Rusconi, B. Kwan, B. L. Giordano, C. Umilta, and B. Butterworth, "Spatial Representation of Pitch Height: The SMARC Effect," *Cognition*, vol. 99, no. 2, pp. 113–129 (2006 Mar.). https://doi.org/10.1016/j.cognition.2005.01.004.

[62] Z. Eitan and R. Y. Granot, "How Music Moves: Musical Parameters and Listeners Images of Motion," *Mu-*

sic Percept., vol. 23, no. 3, pp. 221–248 (2006 Feb.). https://doi.org/10.1525/mp.2006.23.3.221.

[63] Z. Eitan and R. Timmers, "Beethoven's Last Piano Sonata and Those Who Follow Crocodiles: Cross-Domain Mappings of Auditory Pitch in a Musical Context," *Cognition*, vol. 114, no. 3, pp. 405–422 (2010 Mar.). https://doi.org/10.1016/j.cognition.2009.10.013.

[64] M. Chapman, W. Ritsch, T. Musil, et al., "A Standard for Interchange of Ambisonic Signal Sets. Including a file Standard With Metadata," in *Proceedings of the Ambisonics Symposium*, pp. 1–6 (Graz, Austria) (2009 Jun.).

[65] C. Armstrong, L. Thresh, D. Murphy, and G. Kearney, "A Perceptual Evaluation of Individual and Non-Individual HRTFs: A Case Study of the SADIE II Database," *Appl. Sci.*, vol. 8, no. 11, paper 2029 (2018 Oct.).

A.1 BESSEL FORMULAE

Throughout the paper, the following expansions using Bessel functions of the first kind have been used:

$$\begin{aligned}\cos(D\sin(\theta)) &= J_0(D) + 2\sum_{n=1}^{\infty} J_{2n}(D)\cos(2n\theta)\\ &= \sum_{n=-\infty}^{\infty} J_{2n}(D)\cos(2n\theta),\end{aligned} \tag{84}$$

$$\begin{aligned}\sin(D\sin(\theta)) &= 2\sum_{n=1}^{\infty} J_{2n-1}(D)\sin((2n-1)\theta)\\ &= \sum_{n=-\infty}^{\infty} J_{2n-1}(D)\sin((2n-1)\theta),\end{aligned} \tag{85}$$

where the following has been used:

$$\sum_{n=-\infty}^{\infty} J_{2n}(D) f(2n) = \sum_{n=-\infty}^{\infty} J_{2n}(D) f(-2n), \tag{86}$$

$$\begin{aligned}&\sum_{n=-\infty}^{\infty} J_{2n-1}(D) f(2n-1)\\ &\quad = -\sum_{n=-\infty}^{\infty} J_{2n-1}(D) f(-2n-1).\end{aligned} \tag{87}$$

Eqs. (86) and (87) can be easily derived using the well-known relation:

$$J_n(D) = (-1)^n J_{-n}(D). \tag{88}$$

As a consequence, the following is obtained:

$$\cos\phi\cos(D\sin\theta) = \sum_{n=-\infty}^{\infty} J_{2n}(D)\cos(\phi + 2n\theta), \tag{89}$$

$$\cos\phi\sin(D\sin\theta) = \sum_{n=-\infty}^{\infty} J_{2n+1}(D)\sin(\phi + (2n+1)\theta), \tag{90}$$

$$\sin\phi\cos(D\sin\theta) = \sum_{n=-\infty}^{\infty} J_{2n}(D)\sin(\phi + 2n\theta), \tag{91}$$

$$\sin\phi\sin(D\sin\theta) = -\sum_{n=-\infty}^{\infty} J_{2n+1}(D)\cos(\phi + (2n+1)\theta), \tag{92}$$

as well as:

$$\cos(\phi + D\sin(\theta)) = \sum_{n=-\infty}^{\infty} J_n(D)\cos(\phi + n\theta), \tag{93}$$

$$\sin(\phi + D\sin(\theta)) = \sum_{n=-\infty}^{\infty} J_n(D)\sin(\phi + n\theta)\,. \tag{94}$$

A.2 CONCATENATING HARMONICALLY OSCILLATING ROTATIONS IN 2D

Consider a frequency-modulated rotation matrix of the following type:

$$\begin{aligned}&\mathbf{R}(\omega_r + D\sin(\omega_m t))\\ &\quad = \begin{bmatrix}\cos(\omega_r + D\sin(\omega_m t)) & -\sin(\omega_r + D\sin(\omega_m t))\\ \sin(\omega_r + D\sin(\omega_m t)) & \cos(\omega_r + D\sin(\omega_m t))\end{bmatrix}.\end{aligned} \tag{95}$$

To show the effect of two or more such rotations applied to the same source, the strategy used in [21], which replaces one of the time-dependent terms with a generic variable C_1. Because $\mathbf{R}(\theta_1)\cdot\mathbf{R}(\theta_2) = \mathbf{R}(\theta_1 + \theta_2)$, the compound matrix $\mathbf{R}(C_1)\cdot\mathbf{R}(D_1\sin(\omega_{m_1}t)) = \mathbf{R}_c(C_1 + D_1\sin(\omega_{m_1}t))$, where $C_1 = C_2 + D_2\sin(\omega_{m_2}t) + \omega_{r_1}t$ shall be considered. For the purpose of only concatenating two frequency-modulated rotation matrices, $C_2 = \omega_{r_2}t$ can be considered, which will prove useful later. Also, because $m_{c_{11}} = m_{c_{22}}$ and $m_{c_{21}} = -m_{c_{12}}$, only the first column of $\mathbf{R}_c$ shall be considered:

$$\begin{aligned}m_{c_{11}} &= \cos(C_1)\cos(D_1\sin(\omega_{m_1}t))\\ &\quad - \sin(C_1)\sin(D_1\sin(\omega_{m_1}t)),\end{aligned} \tag{96}$$

$$\begin{aligned}m_{c_{21}} &= \sin(C_1)\cos(D_1\sin(\omega_{m_1}t))\\ &\quad + \cos(C_1)\sin(D_1\sin(\omega_{m_1}t)).\end{aligned} \tag{97}$$

Using Eqs. (93) and (94), the following is received:

$$m_{c_{11}} = \sum_{n_1=-\infty}^{\infty} J_{n_1}(D_1)\cos(C_1 + n_1\omega_{m_1}t), \tag{98}$$

$$m_{c_{21}} = \sum_{n_1=-\infty}^{\infty} J_{n_1}(D_1)\sin(C_1 + n_1\omega_{m_1}t). \tag{99}$$

$C_1 = C_2 + D_2\sin(\omega_2 t) + \omega_{r_1}t$ is now unpacked. In order to simplify the next steps, $C_2^* = C_2 + \omega_{r_1}t + n_1\omega_{m_1}t$ may also be defined:

$$m_{c_{11}} = \sum_{n_1=-\infty}^{\infty} J_{n_1}(D_1)\cos(C_2^* + D_2\sin(\omega_{m_2}t)), \tag{100}$$

$$m_{c_{21}} = \sum_{n_1=-\infty}^{\infty} J_{n_1}(D_1)\sin(C_2^* + D_2\sin(\omega_{m_2}t)). \tag{101}$$

The resulting factors are identical to Eqs. (96) and (97) after applying the trigonometric identity in each case, meaning

the previous steps can be applied once more, using the new terms:

$$m_{c_{11}} = \sum_{n_1=-\infty}^{\infty} J_{n_1}(D_1) \sum_{n_2=-\infty}^{\infty} J_{n_2}(D_2) \cos(C_2^* + n_2\omega_{m_2}t), \tag{102}$$

$$m_{c_{21}} = \sum_{n_1=-\infty}^{\infty} J_{n_1}(D_1) \sum_{n_2=-\infty}^{\infty} J_{n_2}(D_2) \sin(C_2^* + n_2\omega_{m_2}t). \tag{103}$$

What remains is to unpack C_2^*, and the nearly final solution is received:

$$m_{c_{11}} = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} J_{n_1}(D_1)J_{n_2}(D_2) \cos(C_2 + n_2\omega_{m_2}t + \omega_{r_1}t + n_1\omega_{m_1}t), \tag{104}$$

$$m_{c_{21}} = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} J_{n_1}(D_1)J_{n_2}(D_2) \sin(C_2 + n_2\omega_{m_2}t + \omega_{r_1}t + n_1\omega_{m_1}t). \tag{105}$$

If only two frequency-modulated rotation matrices are considered, $C_2 = \omega_{r_2}t$ can be set. However, the steps in Eqs. (100)–(105) can be repeated, setting $C_i = C_{i+1} + D_{i+1}\sin(\omega_{m_{i+1}}t) + \omega_{r_i}t$ and $C_{i+1}^* = C_{i+1} + \sum_{k=0}^{i}(\omega_{r_k}t + n_k\omega_{m_k}t)$. Thus, for a series of harmonically oscillating rotations applied to a front-facing source,

$$\mathbf{v}_{\text{out}}(t) = \mathbf{R}(\omega_{r_N} + D_N \sin(\omega_{m_N}t)) \ldots \mathbf{R}(\omega_{r_1} + D_1 \sin(\omega_{m_1}t)) \begin{pmatrix} CA\cos(\omega_s t) \\ 0 \end{pmatrix}, \tag{106}$$

the following harmonic output is therefore received:

$$v_{x_{\text{out}}}(t) = \frac{CA}{2} \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} J_{k_1}(D_1) \ldots J_{k_N}(D_N) \Big[\cos((\omega_s + \omega_{r_1} + k_1\omega_{m_1} + \cdots + \omega_{r_N} + k_N\omega_{m_N})t) + \cos((\omega_s - \omega_{r_1} - k_1\omega_{m_1} - \cdots - \omega_{r_N} - k_N\omega_{m_N})t)\Big], \tag{107}$$

$$v_{y_{\text{out}}}(t) = \frac{CA}{2} \sum_{k_1=-\infty}^{\infty} \cdots \sum_{k_N=-\infty}^{\infty} J_{k_1}(D_1) \ldots J_{k_N}(D_N) \Big[\sin((\omega_s + \omega_{r_1} + k_1\omega_{m_1} + \cdots + \omega_{r_N} + k_N\omega_{m_N})t) - \sin((\omega_s - \omega_{r_1} - k_1\omega_{m_1} - \ldots - \omega_{r_N} - k_N\omega_{m_N})t)\Big]. \tag{108}$$

A.3 SM SYNTHESIS SOUND EXAMPLES

The following appendix will briefly give some additional information on the sound examples provided with this publication.[2] In general the examples closely follow the approach taken in each example of SEC. 2.3.3. In each example, a sound was constructed using Eqs. (71), (75), and (76), respectively, and the spectral shaping effect is demonstrated by slowly changing the parameter of each respective parametric time constant rotation matrix. However, because SEC. 2.3.3 features examples with fundamental frequencies in the kilohertz range for better visual display, lower values were chosen for the audible examples because the high frequencies may be harsh to the ear.

Each example was rendered to binaural stereo, using subject D1's binaural room impulse response from the SADIE database [65]. As mentioned throughout the text above, because the sounds created using multiple-rotation SM synthesis cannot be replicated using RM, AM, or FM without filtering, an approximation using additive synthesis [18] over mono was done in each case. For a better comparison of the spatial qualities created using SM synthesis, the additive synthesis' mono signal is also rendered using a single loudspeaker virtually placed in front of the listener using binaural stereo. In order to imitate the change in tone color as the parametric time-independent rotation changes, the envelope of each harmonic in all three velocity components of the respective SM synthesis sound was extracted and then applied to each frequency produced via additive synthesis. Note, however, that, due to SM synthesis' spatial distribution of harmonics across the velocity vectors, as well as phase differences between frequencies shared across some velocity vectors, the summation in the 1D mono signal using additive synthesis results in a tone color that does not perfectly match that of SM synthesis in all cases.

The SM synthesis values used in each example are the following:

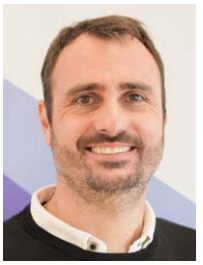
- Example 1 uses Eq. (71) with $f_s = 200$ kHz, $f_1 = 1.2$ kHz, and $f_2 = 600$ kHz.
- Example 2 uses Eq. (75) with $f_s = 500$ Hz, $f_1 = 6$ kHz, $f_2 = 3$ kHz, $f_3 = 2$ kHz, and $f_4 = 500$ Hz.
- Example 3 uses Eq. (76) with $f_s = 250$ kHz, $f_1 = 3.5$ kHz, and $f_2 = 1$ kHz, as well as $D_1 = 0.3$ and $D_2 = 0.6$.

The order of each example is the same in each case. Each example always starts with 5 s of the additive synthesis approximation, weighting the harmonics according to the starting values of each respective envelope. Then, the envelope is traced over the course of 5 s. After that, the final spectrum is heard for another 5 s. Each example then switches immediately to the SM synthesis case. Again, each sound is first presented for 5 s with the parameter for the respective parametric time constant rotation at 0. Then, each respective parametric time constant rotation is then slowly changed to rotate from 0 to 2π over the course of 5 s. Finally, the sound with the parametric time constant rotation at 2π is heard for a last 5 s before each example concludes.

[2] The examples can be heard using the following link: https://github.com/multimedia-eurecat/SMS.

THE AUTHORS

Timothy Schmele



Dr. Adan Garriga

Timothy Schmele is a doctoral candidate at the Institute of Music Informatics and Musicology of the University of Music, Karlsruhe, Germany. He is also a researcher at the Multimedia Technologies research group in Eurecat. He holds Master's degrees in sound and music computing from the University Pompeu Fabra and in music technology from the University of Limerick. He has been a lecturer at the University of Barcelona; the University of Music, Karlsruhe; and the SAE Institute. Recognition for his compositions include the Luigi Russolo first and grand prizes as well as the Karl Szcuka support grant.

•

Dr. Adan Garriga is the Director of the Multimedia Technologies research group in Eurecat, Barcelona, Spain. He holds a Ph.D. in Theoretical Physics from the University of Barcelona and M.B.A. from the EAE Business School. He also holds the Piano Degree from the Conservatori del Liceu. He has been lecturer at the University of Barcelona, University of the Balearic Islands, and Pompeu Fabra University. At the latter, he acted as deputy director of the Computer Science Studies. He has authored scientific publications in physics, mathematics, and acoustics in indexed journals. He has been the Principal Investigator of several European projects on immersive audio.