P. Malecki, M. Piotrowska, K. Sochaczewska, and S. Piotrowski "Electronic Music Production in Ambisonics-Case Study"
J. Audio Eng. Soc., vol. 68, no. 1/2, pp. 87–94, (2020 January/February.).
DOI: https://doi.org/10.17743/jaes.2019.0048

Electronic Music Production in Ambisonics—Case Study

PAWEL MALECKI,[1] *AES Member*, **MAGDALENA PIOTROWSKA,**[2] *AES Member*,
(pawel.malecki@agh.edu.pl) (mplewa@multimed.org)

KATARZYNA SOCHACZEWSKA[1], **AND SZYMON PIOTROWSKI**[3]
(ka.sochaczewska@gmail.com) (psychosoundstudio@gmail.com)

[1]AGH University of Science and Technology, Faculty of Mechanical Engineering and Robotics, Department of Mechanics and Vibroacoustics, Al. Mickiewicza 30, 30-059 Kraków, Poland
[2]Faculty of ETI, Multimedia Systems Department Gdansk University of Technology, G. Narutowicza 11/12, 80-233 Gdańsk, Poland
[3]Psychosound Studio, Wrocławska 28, 30-011 Kraków, Poland

This paper introduces a case study of original electronic music production using ambisonics and describes the spatial remix of a stereo composition. The main goal was to explore the potential of all dimensions and extended space. Remarks on the mixing process, composition for 3D systems, the creative factor, and aesthetics are followed by listening tests. The subjective evaluation involved a panel of experts and included stereo and ambisonic playback, as well as binaural rendering. The authors describe the capabilities and challenges of electronic music production using ambisonics. These conclusions, along with the outcomes of the performed experiment, show the potential of 3D audio as a successful medium for electronic music production.

0 INTRODUCTION

Spatial sound requires extraordinary technical resources and efforts, from both the creator and the recipient. So far, spatial sound systems have not gained significant popularity among mass recipients. Ambisonics is one of the available 3D audio production and playback systems and researchers have recently investigated the technical and perceptual aspects of various solutions [1–4]. 3D audio systems, such as channel-based, scene-based, and object-based, enable a new creative dimension to be established and a unique artistic experience becomes possible. These are the reasons why spatial sound productions with immense, but also uncountable, potential are still developing. Numerous works discuss the technical methods of audio production using ambisonics, as well as case studies of music applications, perceptual aspects, and composition and mixing processes for 3D systems [5–11]. The main aim presented by [11] is to create music that actively engages our spatial perception. The authors discuss what should be taken into consideration during the composition process for 3D systems, and describe the scene-based approach, where space is treated as a musical parameter.

Channel-based approach due to the standardization and content availability is still the most popular technique for spatial sound productions. There is a variety of audio engineering tools within these formats. Standards such as 5.1 or 7.1 are fixed with a constant number of channels and very limited possibilities regarding geometrical changes. On the other hand, object-based techniques allow to achieve very high spatial clarity and can be decoded to any multichannel layout. They are the current standard in cinema productions and can be found in high fidelity home theaters. The most popular object-based format is an expensive and closed source, which is a disadvantage especially for scientific purposes. Alternatively, ambisonics is an open-source technology with a very strong science background [12]. The simple B-format (1st order ambisonics) is being employed in global VR (Virtual Reality) solutions such as Google VR or Facebook 360. HOA (Higher Order Ambisonics) gives great spatial opportunities while its limitations are a small number of dedicated tools and no support of multichannel tracks (over 16 channels) in the most popular DAWs. Currently it is a big challenge to incorporate existing standards (such as popular VST plug-ins) to any spatial sound formats workflow. Also, the important milestone would be the development of a multichannel streaming platform both for potential listeners and content delivery.

In the presented research, the authors introduce the case study of original electronic music production and describe an ambisonic remix of a stereo piece. The main emphasis was placed on the utilization of all dimensions and extended space. Remarks on the mixing process, composition for 3D systems, a creative factor, and aesthetics are followed by listening tests performed by a panel of experts. The subjective evaluation included stereo and ambisonic playback, as well as binaural rendering. In particular, binaural rendering is currently a promising topic, with more information regarding this matter, such as localization, front-back-up-down confusion, and sound quality, found, e.g., in [13–15].

The musical work described in this paper has been publicly presented numerous times in the form of an ambisonic remix to a wide, specialized audience (i.e., during the 144 AES Convention in Milan [16] and klingt gut! International Symposium on Sound in Hamburg [17] in 2018). During these events, the discussed music piece was played back through a multichannel ambisonic system and was followed by a discussion and comments from the audience. The remainder of the paper is organized as follows: Sec. 1 briefly introduces a composition overview and background of the project, together with a short outline of a stereo mix, while the spatial sound production, technical system, and methods are detailed in Sec. 2. Sec. 3 shows the listening evaluation process along with the obtained results. The concluding remarks are presented in Sec. 4.

1 MUSICAL CONTENT

The music pieces used in the course of the presented study are the “ANTON” original compositions created by one of the authors [18]. Four tracks were selected from a conceptual, electronic music album that relates to interstellar travel and man’s desolate journey through the universe. This concept and imaginary space, along with spatial and three-dimensional thinking while composing, inspired the authors to introduce these music ideas in 3D sound systems using ambisonics. The majority of sounds come from synthesizers, with approximately 80% from virtual instruments (i.e., Serum, Sylenth. ANA, Signal, Alchemy) and the others from hardware (i.e., DSI evolver, Alesis ION, Roland Juno 60) supplemented with fragments of sampled vocal recordings. The sounds are diverse and vary between tracks. The compositions are based mostly on melody and harmony while percussive elements play a secondary role. Stereo mixes were intended to be wide, deep, and spacious. Numerous reverbs, with a long decay time, overlap, and interfere thereby creating a changing and unreal space. The position of sound sources is fixed, with the impression of movement only coming from stereophonic timbres.

The production and mixing processes were performed using a hybrid system including DAW along with an analogue signal chain consisting of an equalizer and dynamic processors. During the final stage, the tracks were mastered using a digital and analogue chain. A bus was an important part of the process, because it allowed the desired timbre and character to be achieved. The sound was pulsing and saturated; bandwidth limited to 50 Hz – 12 kHz, with smooth sounds, without transients, and without a large dynamic range. It is important to note that the compositions, production, and mixing process were linked, as they were performed by one person through numerous iterations. The music used in the experiment is available for recipients in stereo [18] as well as binaural renders of ambisonic production [19].

Fig. 1. AGH UST Laboratory of Auralization

Table 1. Speakers’ coordinates in ambisonic set-up utilized in the course of research

No.	Azimuth φ[°]	Altitude θ[°]	No.	Azimuth φ[°]	Altitude θ[°]
1	0	0	9	68	–45
2	45	0	10	158	–45
3	90	0	11	–113	–45
4	135	0	12	–23	–45
5	180	0	13	23	45
6	–135	0	14	113	45
7	–90	0	15	–158	45
8	–45	0	16	–68	45

2 AMBISONIC MIX

The spatial remix of stereo compositions creates numerous opportunities as well as challenges. On the one hand, a spatial set-up obviously gives more space to be utilized and the tools, along with human creativity, can lead to more selective and impressive outcomes. On the other hand, the sounds previously used in stereo might not be sufficient to take advantage of the whole available space. This issue can be approached, for example, by creative use of effects and automation or, in some cases, might require changes in the arrangement. Moreover, due to the lack of dynamic bus processing in ambisonics, one might find it difficult to create a consistent sound scene and desired interaction between sounds.

2.1 Technical Description

The ambisonic mix presented in this study was performed in the AGH UST Laboratory of Auralization. The control room (dimensions 3.9 m x 6.7 m x 2.8 m) is equipped with a 16-channel, spherical loudspeaker set-up (Fig. 1). The room is acoustically treated and the average RT20 is 0.15 s, calculated from the RT20 at 500 Hz and 1,000 Hz. Sixteen loudspeakers (Genelec 6010) are set up on a 3.5 m sphere radius within evenly distributed polar coordinates (Tab. 1). The set is supplemented with a PSI Sub A225-M subwoofer.

Decoding to a given speaker layout was performed using an Allrad algorithm [20].

Spatial production was performed using Reaper software and the following audio plugins dedicated to 3D audio production: ambiX and mcfx by Matthias Kronlachner [21], the IEM Plug-in Suite created by staff and students of the Institute of Electronic Music and Acoustics in Graz [22], and Harpex [23]. It is important to note that other ambisonic tools are available for this kind of task. The advantage of the chosen solutions is their open-source license (most of them), scientifically based algorithms, and extensive documentation. All the employed plugins also have complex configuration options that are directly related to their functions while many commercial solutions tend to include ambiguous parameters. Therefore, the intuitive automation manipulation and intentional space shaping was possible in the course of the case study. Since other tools with similar functionality exists and new tools are constantly developed, the choice of the utilized software can be a matter of individual preference.

2.2 Spatial Remixing Strategy

The main assumption of the remix was to add another spatial dimension without interfering with the musical concept and the sound timbre. For this purpose, stereo stems included in the original composition were exported. It is worth noting that some of the effects (reverbs and delays) were also an important part of timbre, and therefore they were also included in stems. All the decisions related to spatialization were made by the composer or with his consent. Additionally, the musical role of the given elements of the mix was analyzed and, on this basis, the location and size in space were attributed. The validity of each performed change was verified by referencing the stereo mix in situ. Four parts of the composition described in Sec. 1 (approximately 20 minutes) were remixed into ambisonics. Due to the diversity in music material, this process was realized in separate sessions.

2.3 Implementation

2.3.1 Stems

Every element of the original stereo mix was rendered by the composer as a stereo stem and, in this form, imported into a 3D remix set-up. During the process, the authors did not intend to change either timbre or dynamics. Although it is important to note that some changes resulting from the modification of the location and size of the sources occurred. Based on the presented case study an observation that ambisonic tools might introduce some changes in the sound source timbre should be taken into consideration during production process. Each track was equipped with 5th-order ambisonic tools enabling positioning and other signal modification in 3D spaces, with the signal subsequently being sent directly to the decoder. Reverb (also in an ambisonic format) was added in the parallel bus and this signal was summed in the decoder track.

2.3.2 Perceived Phantom Source Size and Position

Comparing to the stereo mix, the atmospheres and backgrounds in the spatial set-up have been significantly increased in size and their signal has been directed to almost all directions, both as direct sound and reverb. The main elements of the mix were arranged mainly in front of the listener, but the distance between them was significantly increased in relation to the stereo mix. In addition, their sizes were increased and more complex movements were introduced. Elements that performed a secondary or temporary part in the mix were often located in non-obvious and atypical positions (e.g., behind or above the listener).

2.3.3 Spatial Mapping

Additionally, for selected instruments, spatial mapping was performed. This operation is based on assigning a particular pitch to a specific, fixed position in space. Spatial mapping was performed in order to make the listener perceive instruments of considerable size and depth. The sound source size term in multichannel audio is treated quite intuitively because there is no formal definition or measure of this psychoacoustic phenomena. In many cases desired size is achieved through amplitude or phase differences between physical loudspeakers. Also reverbs, delays or other tools are often used for this purpose as well. Applied spatial mapping is another idea to emphasize source size sensation. This operation was carried out using automation (Fig. 2). The results of this process were especially rewarding for the composer and integrate well with synthesized sounds.

2.3.4 Abstract Space

An important element of the remix was the creation of unique sound spaces with an intriguing and abstract character. This was accomplished by transformation of the ambisonic signal and position modulation by random signals. The atmosphere transformation was performed by synchronous and slow rotation of the ambisonic signal (mostly yaw and roll). This type of modification was subjected mainly to continuous sounds with a slow-changing envelope. Ambient sounds of an impulse character were positioned using fast and randomized panning. This effect was obtained using the functions of the Reaper program, which allows modulation of any values of any plug-in parameters. In this case, horizontal and vertical angles of the ambisonic panner of the given signal were modulated using a random signal with a uniform distribution.

2.3.5 Movement

The location of selected phantom sources was variable over time. The position and size of the apparent sound sources were automated. All movement was synchronized with the rhythm and dynamics changes to enhance the musical sense. Movements were introduced carefully to avoid distracting the listener and therefore do not take too much attention. Fig. 2 shows the automation curve of the position and size of an apparent sound source for selected phrases. Under the wave envelope of the one track there are four

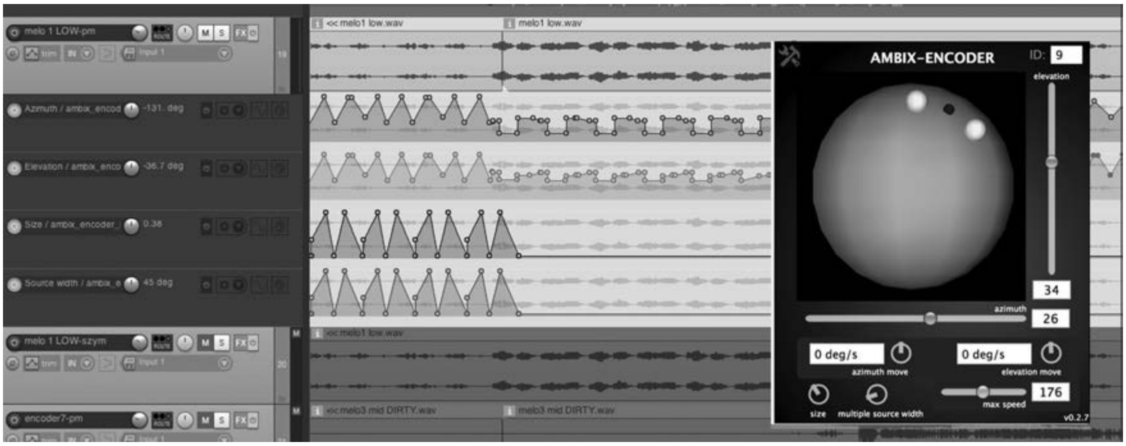

Fig. 2. Exemplary track automation of source position and size in Reaper program

dedicated automation curves that modify sound source Azimuth, Elevation, Size, and Source Width. These curves correspond to the sound source, represented by the dark spot on the sphere that is shown on the right side of the Fig. 2. The ambisonic encoder also shows two bright balls, which illustrate source width.

2.3.6 Ambisonic Mix Finalization

For all four remixed pieces, original animation consisting of old video footage from the 1950s and ’60s was prepared. Due to the visual to auditory impact relation [24], it was important to carefully adjust the musical rhythm with the video cuts. The video content was based on stereo renders. In order to synchronize ambisonic content with visuals, all the remixed parts were rendered as separate 32-channel 5th-order ambisonic files, which allowed sufficient transitions to be made.

2.4 Experience and Aesthetics

Most of all, mixing in ambisonics provides amazing capabilities in terms of the space that can be utilized. At the same time, sources that have been packed into a 60 stereo set-up now can cover almost the whole sphere. The aim of the ambisonic mix was to carry the same story as the stereo track. A 3D set-up provides new opportunities but also some limitations, which result in some compromises, especially regarding reverberation and dynamics. During the mixing in ambisonics, rules that apply to the traditional stereo mix were also employed. The authors followed the requirements of the well balanced mix. The most meaningful sound sources (i.e., melody and the main rhythm) were positioned in one place, mostly in the middle or “everywhere.” Gravity towards one direction was avoided (frequency balance among the whole space is also important). Creating large and spacious instruments is something novel when compared to stereo and, along with spatial mapping, can complement electronic productions involving synthesized sounds. One of the main challenges was not to disturb the listener. Source movements, meaningful sounds from different directions, and surprising sounds, especially at the back, were avoided, so the listener could follow the story carried by the music. The other demanding part was the lack of ambisonic bus processing. The workflow, where whole sonic material can be processed as an entity creates many opportunities to shape the sound scene in regard to the timbre, dynamics, and localization. While this approach is commonly used in the channel based audio, in the scene based applications, tools are still very limited. These observations confirm that this issue might be addressed in future technical solutions. Although ambisonic dynamic and reverb plugins provide great opportunities, their palette of sound is limited when compared to plugins available in the stereo format. Creativity in sound placement is definitely a great opportunity to emphasize musical content but, at the same time, provides a chance to take the attention away from the music.

3 LISTENING TESTS

The methodology of the listening experiment was based on previous research including a 3D audio evaluation performed by the authors [25–28]. As described in Sec. 2, remixing electronic music produced in stereo into an ambisonic system is associated with many opportunities, but in some cases also compromises; therefore subjective evaluation involving a group of 15 experts was performed. The test consisted of two parts: preference choice between the ambisonic (16.1 system as described in Sec. 2.1) and stereo mix (2.1 system of 2 x Genelec 8030 + PSI Sub A225-M) and preference choice between the ambisonic binaural render versus the stereo mix (Beyerdynamic DT770 closed headphones). A binaural render from the 5th-order of ambisonics was performed using an IEM Plug-in [22].

The listening set-up was calibrated at the listening position (shared sweet spot for both ambisonic and stereo set-ups) using a SVAN 957 level meter up to 80 dB SPL +/–1 dB with a C weighting curve and 30 s integration period for every excerpt and system.

The group of listeners consisted of highly trained listeners: audio engineers (6), sound engineering teachers (4), musicians (2), and acoustical engineering students (3). Most of them reported previous experience with surround and/or immersive set-ups. The listeners were asked to mark their preference (version A or B) of 3 diverse excerpts of the electronic composition described in Sec. 2. They could switch between the versions an indefinite number of times and the duration of each playback was not limited. They rated their preference in four categories:

- Spaciousness
- Selectivity
- Timbre and Dynamics
- General impression.

Participants were instructed regarding all of the evaluated criteria, on how to interpret the categories and how to submit their answers. In all of the cases, listeners were supposed to determine which of two presented variants were preferred, according to their experience and liking. In the case of "Spaciousness" evaluation was based on the perceived space—which variant is more suitable for given music material. "Selectivity" referred to clarity and distinction of particular sources, layers, and instruments. "Timbre and dynamics" alluded preference regarding creation and balance in the area of dynamics and timbre. All these criteria supplemented "General preference," which answered the question: which variant do you consider as better, more interesting, and preferred.

In the case of stereo vs. ambisonics comparison, differences between both versions was obvious (direction of sound only from front speakers or from all around), therefore order of the samples was not randomized. Listeners could change between samples a non-limited number of times, so the order was not causing bias. During headphone playback, double-blind randomization was implemented. In both cases samples were presented as pairs.

The tests were run with a custom script in Reaper DAW, providing a screen and mouse user interface. Since answers have been evaluated in a pairwise preference test, the analysis of the results was based on the test of equal proportions. Evaluation methods as well as the user interface were utilized by authors in their previous research [25, 27, 28] related to the subjective evaluation of spatial and binaural content. Discussion on the listening test environment with extended description can be found in [26]. Besides the preference choice, additional comments from the participants were collected.

The participants were also asked to provide their impressions and comments. In some cases they reported that it was difficult to pick one version as they were so different from each other and that they appreciated both mixes. Some listeners also expressed the reasons behind their preference, such as: characteristic timbre in one of the versions, dynamics, or some reported that, due to being used to stereo or surround set-ups, they found any sound from above very disturbing and distracting. As the music material was diverse among excerpts, the listeners had different comments related to each of them. The first fragment (louder and heavily processed with bus processing) was described as more consistent in the stereo version, while the third piece was complimented for its spaciousness and "breathing character." It is worth noting that the stereo mix was perceived as

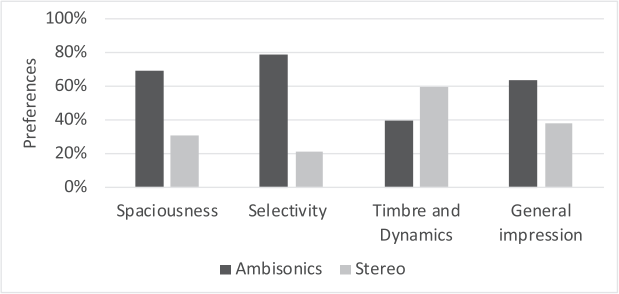

Fig. 3. Listeners' preference results of stereo mix to its ambisonic remix comparison

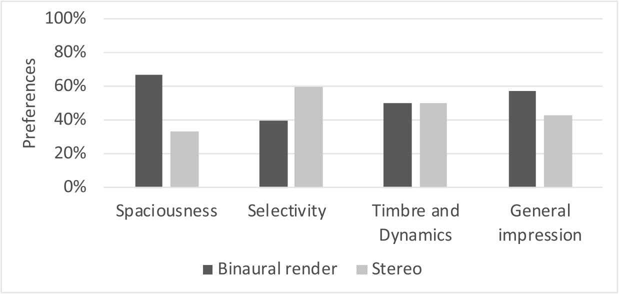

Fig. 4. Listeners' preference results of the comparison of the stereo mix to the binaural render of the ambisonic remix

louder, even though it was calibrated with the same SPL(C) as the ambisonic mix, as described in the previous paragraph. This observation is consistent with the dependency of perceived loudness on the direction of the sound [29].

The results obtained from the first part of the listening test (stereo versus ambisonic mix) are presented in Fig. 3. The outcomes show the significant preference of the ambisonic mix in terms of spaciousness and selectivity. On the other hand, the listeners preferred the timbre and dynamics of the stereo version, although this difference is not as substantial. The perception of timbre and dynamics might be caused by a lack of bus processing in the ambisonic mix. In stereo mix, master bus processing allowed to shape dynamics and timbre based on the chosen tools (both in analog and digital domain) and subtle settings. Unfortunately, at this moment in scene based audio, choice of this kind of tool is very limited, what affects the final experience of the sonic material. The conclusions regarding spaciousness and selectivity are intuitive and are coherent with expectations. Despite some timbre and dynamics concerns, most of the listeners preferred the ambisonic version over the stereo one.

The results shown in Fig. 4 represent a comparison of the stereo version and the automatic binaural downmix from ambisonics. As in the loudspeaker playback, spaciousness is preferred in the ambisonic version, but the results are the opposite in the case of selectivity. The assessment of timbre and dynamics is slightly higher in stereo, while the general preference is in favor of the binaural variant. It is worth recalling that there was no HRTF individualization or equalization.

3.1 Basic Statistics

The level for rejection of a null hypothesis for a normal binominal distribution for the results with given population N=45 (15 participants x 3 excerpts) and the confidence level p=0.05 is 62%. According to this requirement, all results with higher value are treated as statistically significant. For stereo vs. ambisonic, there is significant preference in case of 3 out of all for characteristics (except from “timbre and dynamics”). In stereo vs. binaural comparison, this requirement was fulfilled only in the case of “spaciousness.” It is important to note, that despite statistical significance, results cannot be treated as general outcomes due to the experiment bias.The presented work is based on the case study and outcomes represent preference of the selected group of experts.

4 CONCLUSIONS AND DISCUSSION

This paper includes a complex description of a musical composition in the stereo version and the process of an ambisonic remix. Both versions were produced with the collaboration of the composer, who is also a co-author of this publication. The stereo and ambisonic mixes, as well as the stereo and binaural renders, were included in a subjective evaluation performed by experts.

Based on the results, numerous conclusions can be formed. However, it is important to remember that listening materials, as well as panel experts, are limited and many variables could influence the results. The compared stereo and ambisonic mixes differ not only in terms of space but also with regard to the timbre and dynamics. Moreover, bus processing was applied to the stereo tracks, while there is still a place for specialized tools in ambisonics, as well as standards and good practices. The authors decided to use “real-world” music material instead of creating an artificial laboratory situation where this bias could be eliminated. Therefore, the results cannot be interpreted as general rules, but their interpretation can supplement qualitative remarks.

In general, the listeners preferred the spaciousness and selectivity of the ambisonic version over the stereo one. The obtained results are coherent with the outcomes of other studies focused on the differences between stereo and multichannel reproduction [30–32]. The most interesting conclusion might be formulated based on the comparison between the stereo and binaural renders of the ambisonic mix.

The results clearly show the preference of spaciousness of the binaural version, and the general preference also indicated binaural as preferred. Due to the popularity of headphone playback, the obtained results show the potential of ambisonic productions with targeted binaural playback.

Presented results and conclusions found in the case study focused on the technical workflow, aesthetic possibilities, production, and listening experience. As they are promising, future work should cover more extensive music sets and involve analysis based on the objective features analysis such as in [33]. The outcomes of the presented work, along with the results of the listening tests executed in the course of this study, are a good example of how ambisonics can enrich electronic music and bring it to “another dimension.”

5 ACKNOWLEDGMENT

This research has been partly supported by the AGH UST Project no. 16.16.130.942.

6 REFERENCES

[1] H. Lee, “Capturing 360 Audio Using an Equal Segment Microphone Array (ESMA),” *J. Audio Eng. Soc.*, vol. 67, pp. 13–26 (2019 Jan./Feb.), https://doi.org/10.17743/jaes.2018.0068.

[2] L. Riitano and J. E. Medina Victoria, “Comparison between Different Microphone Arrays for 3D-Audio,” presented at the *144th Convention of the Audio Engineering Society* (2018 May), convention paper 9980; http://www.aes.org/e-lib/browse.cfm?elib=19497.

[3] H. Wittek and G. Theile, “Development and Application of a Stereophonic Multichannel Recording Technique for 3D Audio and VR,” presented at the *143rd Convention of the Audio Engineering Society* (2017 Oct.) convention paper 9869, http://www.aes.org/e-lib/browse.cfm?elib=19266.

[4] K. Sato and Y. Haneda, “Filter Design of a Circular Loudspeaker Array Considering the Three-Dimensional Directivity Patterns Reproduced by Circular Harmonic Modes,” presented at the *142nd Convention of the Audio Engineering Society* (2017 May), convention paper 9767, http://www.aes.org/e-lib/browse.cfm?elib=18643.

[5] M. Neukom, “Ambisonic Panning,” presented at the *123rd Convention of the Audio Engineering Society* (2007 Oct.), convention paper 7297.

[6] E. Benjamin, A. Heller, and R. Lee, “Why Ambisonics Does Work,” presented at the *129th Convention of the Audio Engineering Society* (2010 Nov.) convention paper 8242.

[7] T. Carpentier, “Ambisonic Spatial Blur,” presented at the *142nd Convention of the Audio Engineering Society* (2017 May), convention paper 9729.

[8] M. Chapman, “New Dimensions for Ambisonics,” presented at the *124th Convention of the Audio Engineering Society* (2008 May), convention paper 7478.

[9] M. Frank, F. Zotter, and A. Sontacchi, “Producing 3D Audio in Ambisonics,” presented at the *AES 57th International Conference: The Future of Audio Entertainment Technology–Cinema, Television, and the Internet* (2015 Mar.), conference paper 6-3.

[10] T. Nishiyama, T. Nagata, S. Ogata, T. Hasue, and M. Kashiwagi, “Recording and Mixing Techniques for Ambisonic Sound Production,” presented at the *2018 AES International Conference on Spatial Reproduction—Aesthetics and Science* (2018 Jul.), conference e-Brief 1-4.

[11] P. Lennox, T. Myatt, and J. Vaughan, “From Surround to True 3-D,” presented at the *AES 16th International Conference: Spatial Sound Reproduction* (1999 Mar.), conference paper 16-011.

[12] F. Zotter and M. Frank, *Ambisonics* (Springer, 2019).

[13] G. Reardon, A. Roginska, P. Flanagan, J. S. Calle, A. Genovese, G. Zalles, M. Olko, and C. Jerez, "Evaluation of Binaural Renderers: A Methodology," presented at the *143rd Convention of the Audio Engineering Society* (2017 Oct.), convention e-Brief 359.

[14] M. Zaunschirm, C. Schörkhuber, and R. Höldrich, "Binaural Rendering of Ambisonic Signals by Head-Related Impulse Response Time Alignment and a Diffuseness Constraint," *J. Acous. Soc. Amer.*, vol. 143, no. 6, pp. 3616–3627 (2018), https://doi.org/10.1121/1.5040489.

[15] T. McKenzie, D. Murphy, and G. Kearney, "Diffuse-Field Equalisation of Binaural Ambisonic Rendering," *Applied Sciences*, vol. 8, no. 10, p. 1956 (2018), https://doi.org/10.3390/app8101956.

[16] P. Malecki and S. Piotrowski, ANTON—Universe in 3D: Ambisonics in Electronic Music Production (2018) (accessed February 3, 2019), http://www.aes.org/events/144/immersiveaudio/?ID=5949.

[17] P. Malecki and S. Piotrowski, Electronic Music in Ambisonics (2018) (accessed February 3, 2019), https://klangsymposium.sched.com/speaker/szymon_aleksander_pi otrowski_and_pawel_m.70f9q1x.

[18] *t.est* (2018) (accessed February 3, 2019), http://tdotest.bandcamp.com.

[19] *t.est - Project Anton (Binaural ambisonic render)* (2018) (accessed February 18, 2019), https://youtu.be/uInlGKmsIeg.

[20] F. Zotter and M. Frank, "All-Round Ambisonic Panning and Decoding," *J. Audio Eng. Soc.*, vol. 60, pp. 807–820 (2012 Oct.).

[21] M. Kronlachner and F. Zotter, "Warping and Directional Loudness Manipulation Tools for Ambisonics," EAA Joint Symposium on Auralization and Ambisonics, TU Berlin, Germany, April 3–5, 2014, *accessed January 2019*, vol. 2 (2016).

[22] *IEM Plug-in Suite* (2018) (accessed February 3, 2019), https://plugins.iem.at.

[23] S. Berge and N. Barrett, "High Angular Resolution Planewave Expansion," *Proceedings of the 2nd International Symposium on Ambisonics and Spherical Acoustics May*, pp. 6–7 (2010).

[24] B. Kunka and B. Kostek, "New Aspects of Virtual Sound Source Localization Research—Impact of Visual Angle and 3-D Video Content on Sound Perception," *J. Audio Eng. Soc.*, vol. 61, pp. 280–289 (2013 May).

[25] P. Kleczkowski, A. Krol, and P. Malecki, "Multichannel Sound Reproduction Quality Improves with Angular Separation of Direct and Reflected Sounds," *J. Audio Eng. Soc.*, vol. 63, pp. 427–442 (2015 Jun.), https://doi.org/10.17743/jaes.2015.0054.

[26] K. Sochaczewska, P. Malecki, D. Czopek, and J. Wiciak, "Soundscape Recording: Review of Approaches," presented at the *142nd Convention of the Audio Engineering Society* (2017 May), convention e-Brief 311, http://www.aes.org/e-lib/browse.cfm?elib=18687.

[27] K. Sochaczewska, "The Study of Recording of Klezmer Music Using Ambisonics and Binaural Microphone Techniques" Bechelor thesis, AGH University of Technology, Faculty of Mechanical Engineering and Robotics, Krakow, Poland (2017).

[28] K. Sochaczewska, *Consumer grade surround sound system—calibration, mix and mastering*, Master's thesis, AGH University of Technology, Faculty of Mechanical Engineering and Robotics, Krakow, Poland (2017).

[29] V. P. Sivonen and W. Ellermeier, "Directional Loudness in an Anechoic Sound Field, Head-Related Transfer Functions, and Binaural Summation," *J. Acoust. Soc. Amer.*, vol. 119, no. 5, pp. 2965–2980 (2006), https://doi.org/10.1121/1.2184268.

[30] H. Wierstorf, C. Hold, and A. Raake, "Listener Preference for Wave Field Synthesis, Stereophony, and Different Mixes in Popular Music," *J. Audio Eng. Soc.*, vol. 66, pp. 385–396 (2018 May).

[31] J. Francombe, T. Brookes, and R. Mason, "Evaluation of Spatial Audio Reproduction Methods (Part 1): Elicitation of Perceptual Differences," *J. Audio Eng. Soc.*, vol. 65, pp. 198–211 (2017 Mar.), https://doi.org/10.17743/jaes.2016.0070.

[32] J. Francombe, T. Brookes, R. Mason, and J. Woodcock, "Evaluation of Spatial Audio Reproduction Methods (Part 2): Analysis of Listener Preference," *J. Audio Eng. Soc.*, vol. 65, pp. 212–225 (2017 Mar.), https://doi.org/10.17743/jaes.2016.0071.

[33] A. Rosner and B. Kostek, "Automatic Music Genre Classification Based on Musical Instrument Track Separation," *J. Intel. Infor. Sys.*, vol. 50, no. 2, pp. 363–384 (2018), https://doi.org/10.1007/s10844-017-0464-5.

THE AUTHORS

Pawel Malecki

Magdalena Piotrowska

Katarzyna Sochaczewska

Szymon Piotrowski

Pawel Malecki was born in Gorlice, Poland, in 1985. He received his M.Sc. (Tech.) in mechanics and Ph.D. degree in vibroacoustics from the AGH University in Cracow, Poland, 2013. His thesis, "Evaluation of Objective and Subjective Factors of Highly Reverberant Acoustic Field," focused on the psychoacoustic aspects of choral music perception in the diffused field. His main research tools were auralization techniques combined with ambisonics. He works at the AGH University and also as a freelance sound engineer and acoustic consultant. Dr. Malecki has designed the acoustics of the recording studio in the new AGH University Media Center. He is a former officer of the Polish Acoustical Society.

•

Magdalena Piotrowska received a doctoral degree with distinction at the Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology (GUT), Poland. Her thesis, "Automatic Mood Indexing of Music Excerpts based on Correlation Between Subjective Evaluation and Feature Vector," was dedicated to Music Information Retrieval. She is a member of the Audio Engineering Society where she served as Governor and recently as Chair of the Education Committee. Her main scientific interests are music information retrieval, psychoacoustics, signal processing, and connections between science and practical application in the field of audio technology.

•

Katarzyna Sochaczewska is currently a Ph.D student in Departament of Mechanics and Vibroacoustics in Kraków AGH UST, Poland, where she also received her M.Sc. degree in acoustical engineering in 2017. During bachelor and master studies her main focus was on binaural and surround sound. She now proceeds in the multichannel audio while being involved in various projects on ambisonics, including soundscapes, SIR measurements, psychoacoustical research on spaciousness, and sound localization. She has a strong interest in 3D music production, proving it with award winning compositions. She is also a sound engineer with experience in live sound, studio recording, and broadcasting.

•

Szymon Aleksander Piotrowski is a producer, sound engineer, musician, and composer. He pursues his passion to sound all around the world, recently in Stavanger, Norway, and at the Banff Centre, Canada. He composes original music and sound design for games and videos (i.e., Jeremy Collins "Wild Inside," Burden of Command, Joanna Polak). His electronic music within the t.est project is inspired by cities visited, people, and nature. In t.est, Szymon brings together musicians from different genres, from classical, jazz, and electronic (Simen Kiil Halvorsen, Jexus, Magnus Murphy Joelson, Julie Hasfjord, Michał Wierba, and others). His last project "ANTON" relates to interstellar travel and man's desolate journey through the universe. Szymon has recently performed as a drummer with MLDV, Michał Wierba Trio, and Segue. He has recorded and co-produced bands from different music genres, such as: BenReddik (NOR), Clock Machine (PL), Phenix (FR) MagBalay (PL), Annalie Wilson (UK), and many others. He is CEO of Psychosound Studio and Learn How to Sound foundation. His scientific interests are related to applications in audio engineering and music production.