THE SPHEAR PROJECT UPDATE: REFINING THE OCTASPHEAR, A 2ND ORDER AMBISONICS MICROPHONE

Fernando Lopez-Lezcano
CCRMA, Stanford University
nando@ccrma.stanford.edu

ABSTRACT

This paper presents an update of the *SpHEAR (Spherical Harmonics Ear) project, created with the goal of using low cost 3D printers to fabricate Ambisonics microphones. The project includes all mechanical 3d models and electrical designs, as well as all the procedures and software needed to calibrate the microphones. Everything is shared through GPL/CC licenses and is available in a public GIT repository. [1] We will focus on the status of the eight-capsule OctaSpHEAR 2nd order microphone, with details of the evolution of its mechanical design and calibration.

1. INTRODUCTION

The soundfield microphone was designed in the 1970s by Michael Gerzon and Peter Craven [1] to capture the spherical harmonics of a soundfield up to first-order. It uses four capsules in a tetrahedral configuration, which are matrixed and equalized to derive the Ambisonics B-format signals that represent the soundfield. In 2012 Eric Benjamin published the design and preliminary evaluation of an eight capsule microphone [2], which can capture second order Ambisonics components and shows better performance in first order than the traditional tetrahedral microphone. Its capsules are located in the vertices of a square antiprism, and it can encode 8 of the 9 components of an Ambisonics 2nd order soundfield (figure 1). The R component cannot be recovered as it aliases to W. This design is the basis of our OctaSpHEAR (aka: Octathingy) microphone.

The SpHEAR project started at the end of 2015 with the design and construction of conventional tetrahedral prototypes [3]. These initial designs were followed by eight capsule prototypes [4], with a calibration procedure derived from the work on the tetrahedral microphones. This paper focuses primarily on the eight-capsule design. It presents two different acoustical and mechanical designs of the capsule array, and compares their raw and calibrated performance. It also explore optimizations of the encoding process in the high frequency range.

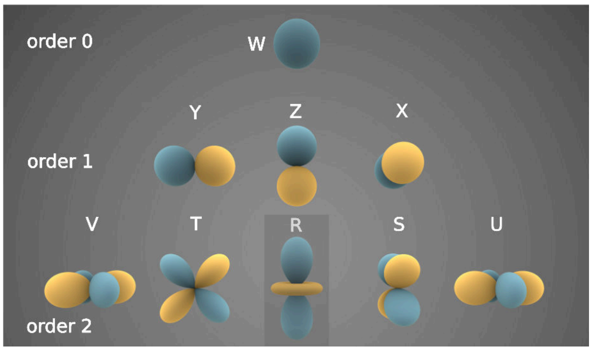

Figure 1. 2nd order spherical harmonics [2]

2. MECHANICAL DESIGN, VERSION 1 (V1)

The mechanical design of the OctaSpHEAR's first two prototypes was a direct derivation of the tetrahedral design. The capsule array is created out of individual capsule holders that assemble together like a 3D puzzle.

Figure 2. OctaSpHEAR v1 capsule array and individual capsule holder

The array was designed with a radius of 18mm, which is close to the minimum that can be obtained with 14mm diameter capsules.

The first two prototypes built have been extensively used for field recordings, and concert and event documentation at CCRMA, and their performance has been considered very adequate when compared to much more expensive microphones. Nevertheless, a plot of the raw frequency response of individual capsules in the array show

[1] https://cm-gitlab.stanford.edu/ambisonics/SpHEAR/
[2] https://commons.wikimedia.org/wiki/File:Spherical_Harmonics.png

problems that suggest a better design could improve the performance of the microphone.

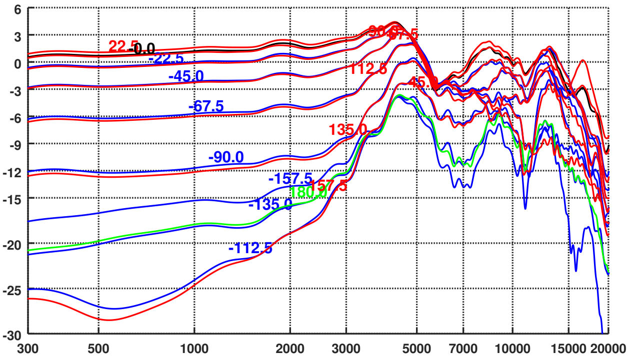

Figure 3. OctaSpHEAR v1 capsule #1 frequency response as a function of incident angle, as indicated on the corresponding trace (in degrees)

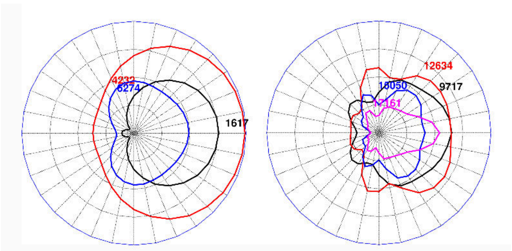

Figure 4. OctaSpHEAR v1 capsule polar patterns at different frequencies

The plots show a strong resonant peak at around 4.4KHz (and its harmonics) caused by the space enclosed by the eight capsules which creates a Helmholtz resonator with multiple necks. The resonances degrade the polar pattern of the capsule at frequencies above about 3KHz. The front to back ratio is reduced, and the capsules become more omnidirectional. This will introduce distortions in the shape of the recovered Ambisonics components.

3. MECHANICAL DESIGN, VERSION 2 (V2)

The resonances suggested a different approach (common to many existing commercial microphones) to the mechanical design of the array. The simple design was replaced by individual conical capsule holders that attach to a spherical core.

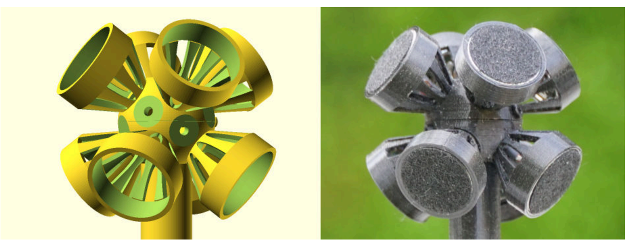

Figure 5. OctaSpHEAR v2 capsule array design

Mechanical design constraints forced us to use a bigger array radius than in version 1 (20.5mm instead of 18mm).

If we only attach one capsule holder to the version 2 design we can measure an almost ideal free field capsule response that still includes the effect of the capsule holder and the body of the microphone. This set of measurements helps us define a baseline performance for this capsule (Primo EMM200), and will help us understand how the rest of the microphone affects its performance.

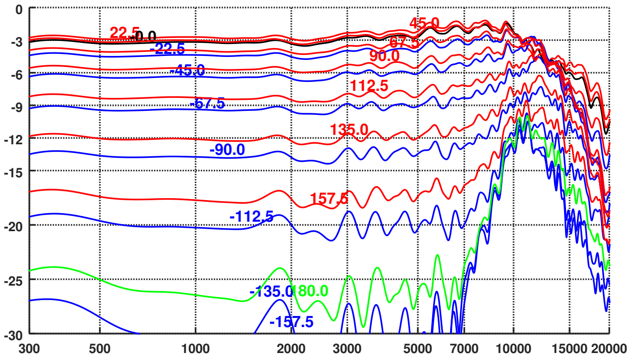

Figure 6. OctaSpHEAR v2, one capsule holder and capsule, frequency response as a function of incident angle

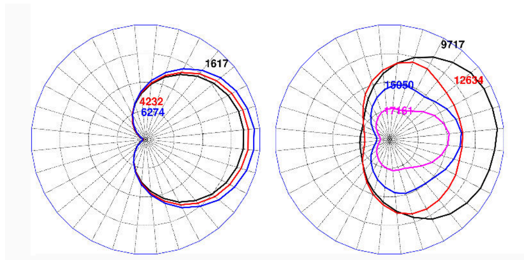

Figure 7. OctaSpHEAR v2, one capsule holder and capsule, polar pattern at different frequencies

Up to about 7Khz the capsule behaves almost like a perfect cardioid, above that we see a degradation of the polar pattern (figure 7) and it becomes more omnidirectional (an expected behavior in cardioid capsules).

Adding the other seven capsules changes the response as shown in figures 8 and 9.

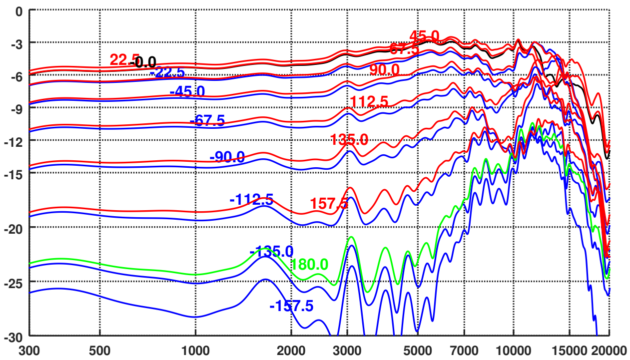

Figure 8. OctaSpHEAR v2 capsule #1 frequency response as a function of incident angle

The occlusion created by all the other capsules degrades the front to back ratio at low and mid frequencies, compared to the measurements of a single capsule. Even then,

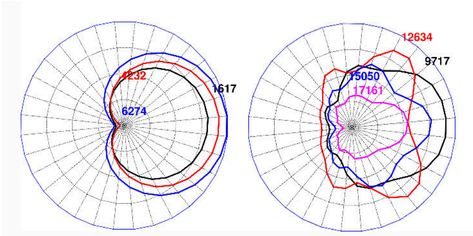

Figure 9. OctaSpHEAR v2 capsule #1 polar responses at different frequencies

the ratio is better than in the version 1 design, except at very low frequencies. The measurements confirm that the resonances at 4.4Khz are gone, as expected, and show that the polar patterns are more consistent over frequency.

In both designs, the polar patterns at very high frequencies in figure 4 and 9 show the shadowing effect of the other capsules and exhibit multiple lobes in their response.

4. MEASUREMENT AND CALIBRATION

As detailed in our previous paper [4], our microphones are measured using quasi-anechoic techniques, with an automated system based on a low cost modified robotic arm. The plots in this paper are derived from 32 equally spaced impulse response measurements in the horizontal plane and 150 measurements of an equally spaced 240 point spherical t-design [6] in the full sphere. Not all points in the t-design are reachable by the arm, which is currently limited due to its length to points lying between -24 and +54 degrees of elevation with respect to the horizontal plane. We obtain about 4.5mSecs of clean equalized impulse response data from each measurement.

The measured impulse responses are used to calibrate the microphone, that is, to create an encoder black box that converts the 8 capsule signals (A format) to 8 Ambisonics components (B format). In an ideal world the B format signals frequency response would be flat, they would be in phase over the full frequency range, and their polar patterns would match the theoretical ones and would not change over frequency.

A simple static 8×8 matrix cannot not satisfy these criteria as the spacing between capsules will create phase related boosts and cancellations in the B format signals above a transition frequency determined by the radius of the array. For our microphones this effect starts to show up at roughly 2KHz, and it can be mitigated by the design of suitable B format correction filters.

As shown in figure 6 the capsule itself does not behave like a cardioid at all frequencies. The polar plots in figure 7 shows it becomes a subcardioid and then a hypercardioid as frequency increases. The other capsules in the array and the structure that supports them also distorts its polar pattern, and the capsule show multiple lobes at very high frequencies (figures 4 and 9), as well as being more omnidirectional at low and mid frequencies in version 2.

These changes of directivity versus frequency will create frequency dependent distortions of the shape of the lobes of the recovered Ambisonics signals.

Finally, cardioid capsules have a peak in their on-axis response at high frequencies (for our capsules there is a 5dB boost at 12KHz, approximately, see figure 10). This peak gradually disappears at increasing angles of incidence from the axis of the capsule, and in our capsules the effect is almost gone at 35 degrees off-axis.

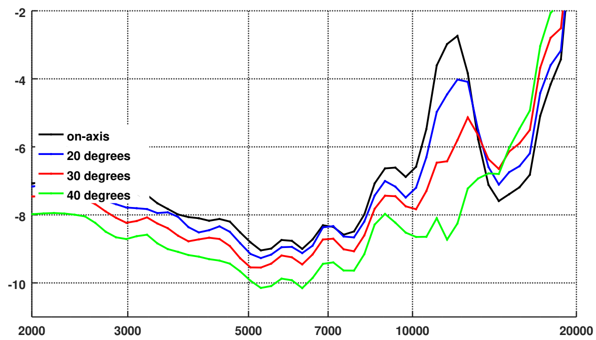

Figure 10. OctaSpHEAR v2, high frequency on-axis capsule resonance

Changes over frequency of the capsule polar patterns, and changes of the polar pattern that depend on the angle of incidence will create distortions in the recovered Ambisonics signals that we cannot really correct. We cannot fix capsule polar patterns, and all our processing and filtering is angle-invariant.

4.1 Encoding from A format to B format

In our current very simple encoder design strategy [4] we start by equalizing all capsules and then using singular value decomposition [7] to create an 8×8 static A to B conversion matrix in a range of frequencies where the capsules can be considered to be co-located. Using the capsule equalization filters and the A to B matrix, we create a first approximation of the B format signals, which deviates from theory above the transition frequency. These signals are then used to create B format equalization filters (figure 11) that mitigate those deviations.

For the horizontal components (WXYUV), the performance of the resulting encoder is different if we calculate it based on just horizontal plane measurements or all measurements. Horizontal only measurements yield better results (flatter frequency response, less variations at high frequencies). We choose to optimize the performance of the microphone for signals coming from the horizontal plane or from low positive or negative elevations, because this is the more likely source of sounds in “normal” recording conditions. So, the A to B matrix and B format correction filters for WXYUV are designed using horizontal plane measurements only. The encoder for the rest of the components (ZST) is designed using the full 3D set of measurements (we do not have a choice here as these are the measurements that have information about height). The final A to B matrix and B format correction filters are a merge of both.

Additionally, the second order components of the microphone are difference microphones, so their output drops at 6dB/octave throughout the full frequency range. The B format filters for those components include the needed boost to equalize them, and we add regularization high pass filters that limit the capsule self-noise amplification.

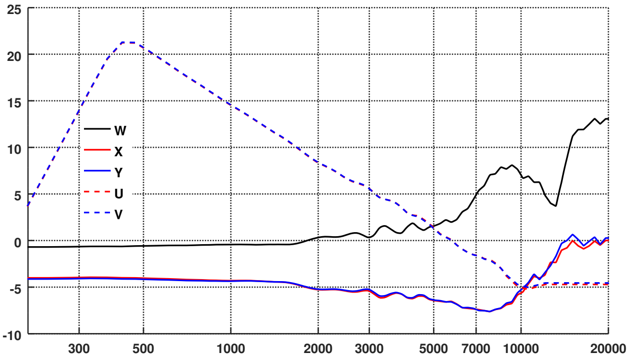

Figure 11. OctaSpHEAR v2 B format equalization filters

Even with these filters, the second order components are noticeably noisier than the first order components, so a set of defeatable expanders is included in the encoder to minimizes the noise for low level signals or silence.

4.2 Version 1 and 2, Ambisonics performance

Figures 12 through 19 show plots of the performance of the calibrated microphones. They include the frequency response of one 1st order (Y) and one 2nd order (V) Ambisonics signals in the horizontal plane, with the azimuth of the excitation signal as a parameter, and polar plots at different frequencies.

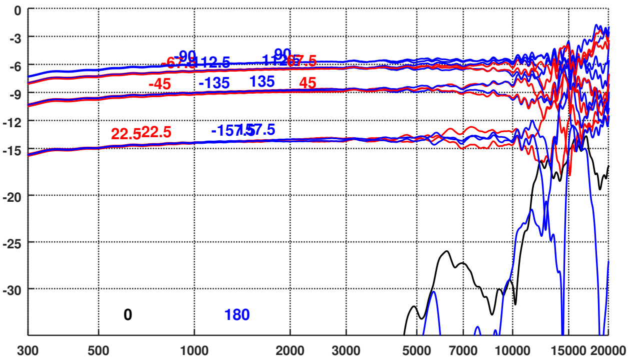

Figure 12. OctaSpHEAR v1 B format Y frequency response with the azimuth angle as a parameter

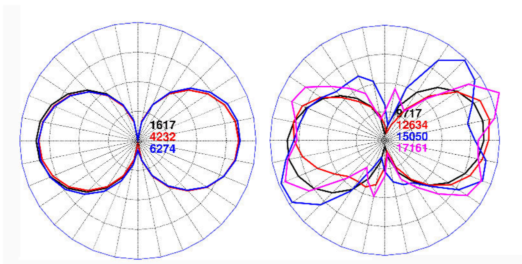

Figure 13. OctaSpHEAR v1 B format Y polar pattern

Both designs show very solid performance up to about 10-11Khz, with frequency response deviations from the ideal flat performance of only a few dB. Above that there is more spread of the response due to the changes in the polar patterns of the capsules.

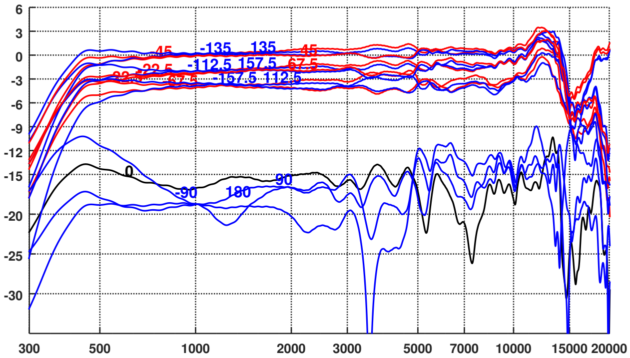

Figure 14. OctaSpHEAR v1 B format V frequency response with the azimuth angle as a parameter

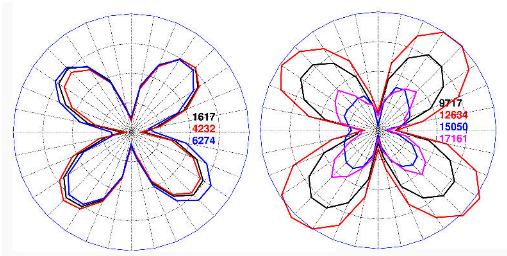

Figure 15. OctaSpHEAR v1 B format V polar pattern

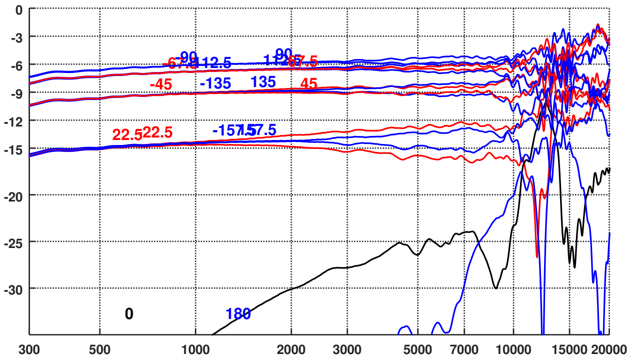

Figure 16. OctaSpHEAR v2 B format Y frequency response with the azimuth angle as a parameter

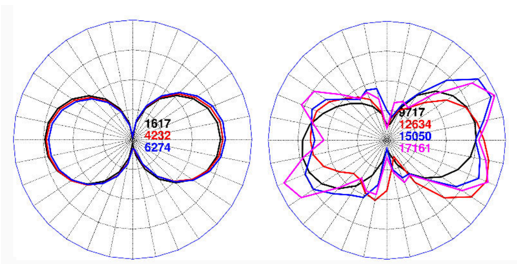

Figure 17. OctaSpHEAR v2 B format Y polar pattern

Version 2 shows slightly worse 1st order (Y) performance than version 1. The null of the lobes is shallower because the capsules have more omnidirectional behavior, and there is more spread in amplitude towards the back of the recovered lobes. On the other hand, version 2 shows better 2nd order (V) performance. The nulls are about 5dB deeper than in version 1, and the amplitude of the measured points is segregated into two groups only (ie: the shape of the lobes is more symmetrical than in version 1). The behavior near the 400Hz lower frequency limit is also marginally better.

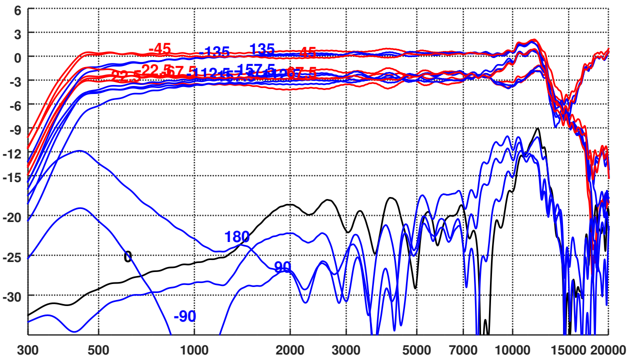

Figure 18. OctaSpHEAR v2 B format V frequency response with azimuth angle as a parameter

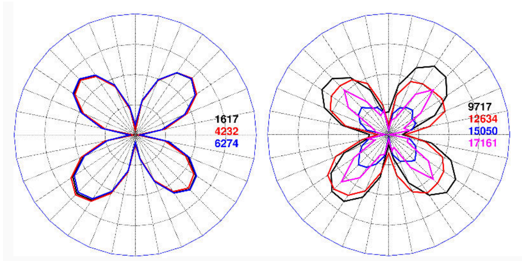

Figure 19. OctaSpHEAR v2 B format V polar pattern

It should be noted that the second order components only provide correct spatial information up to 11Khz, at which point we start seeing the effects of spatial aliasing.

The differences in the performance in the Ambisonics domain are smaller than expected, given the big differences in the behavior of the raw capsule signals, but the change in mechanical design in version 2 shows noticeable improvements, specially in the performance of the 2nd order signals.

4.3 Version 2, 3D performance plots

Figure 20 shows the 3D shape of the X 1st order component at low frequencies (from 800 to 1600Hz). The blue and green dots and their tessellation show the measured points, the red dots show the theoretical points of X for the full 240 point spherical t-design, normalized to the maximum of the measured points. The missing measurements (red dots without blue counterparts) are due to the limited reach of the robotic arm. Figure 21 and 22 show the measured and theoretical V and T second order components respectively. In all three plots there is a very good match between the theoretical and measured shapes.

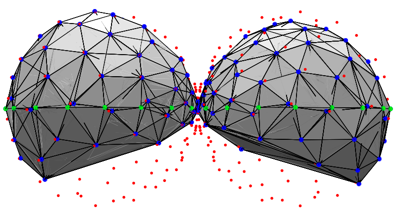

Figure 20. OctaSpHEAR v2 X component, measured (blue dots), measured in horizontal plane (green dots) and t-design (red dots)

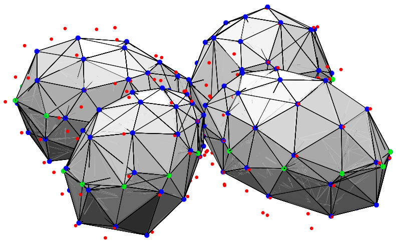

Figure 21. OctaSpHEAR v2 V component, measured (blue and green dots) and t-design (red dots)

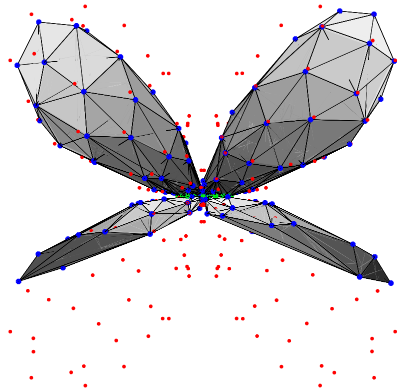

Figure 22. OctaSpHEAR v2 T component, measured (blue and green dots) and t-design (red dots)

4.4 Refining the encoder at very high frequencies

Close examination of the very high frequency behavior of Y shown in figures 12 and 16 indicates unexpected variations in the polar pattern.

Our software defines the shape of the B format equalization filters by measuring the power in logarithmically spaced bands for measurements near the peak of each recovered lobe. The inverse of this power profile is used to create the FIR filters. The following plots of Y versus frequency (from 5KHz to 20KHz) in all measured angles for elevations between -10 and 10 degrees (red traces), and between 20 and 40 degrees (blue traces), show how this approach fails at frequencies above 10KHz.

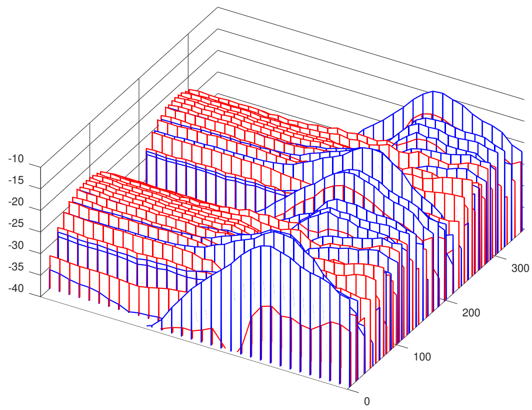

Figure 23. OctaSpHEAR v2 Y waterfall; red traces: -10 to 10 degrees elevation, blue traces: 20 to 40 degrees, 5KHz to 20KHz

At the top of the spectrum we see substantial energy peaks outside of the two Y lobes, while at or near the peak of the lobes we sometimes have a drop in level. Our averaging algorithm does not take into account the energy outside of the lobes, and as a result that frequency range is boosted by the B format equalization filters.

The effect is much worse for elevations above (and below) the horizontal plane (blue traces). This is most likely the result of the on-axis resonant peak of the capsules at 12KHz.

The amount of unintended boost at high frequencies depends on the elevation angle, and our B format filters are angle-invariant, so it is impossible to compensate for this effect. However, we can design additional filters (and merge them into the corresponding B format filters) that take into account the peaks in that frequency range. Figure 24 shows the filter shapes we arrive at if we average all measurements for Y, V and W.

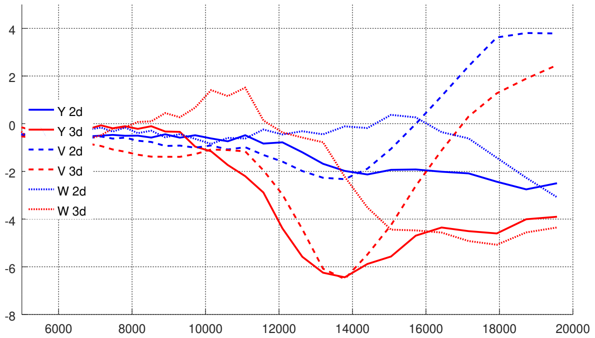

Figure 24. OctaSpHEAR v2 B format high frequency correction filters for Y, V and W

Red traces correspond to the full 3D measurement set, blue traces to the horizontal plane measurements. For both Y and V (and other components not shown) the correction filters for both sets of measurements agree on the frequency of the boost, but not for W (which should not be corrected). The amount of correction to be applied is a tradeoff between behavior in the horizontal plane and above and below it. Figure 25 shows the corrected spectrum if we choose a filter based on the full 3D set.

This is a calibration trade-off similar to the one that happens for the horizontal plane calibration of first order microphones. It is impossible to equalize equally well in all directions, and the choice of which signals to use to design the filters in that region will affect the behavior of the microphone.

5. CONCLUSIONS

Of the two designs we tested for this paper, version 2 (individual conical capsule holders with a central spherical core) has the best overall performance, specially at mid and high frequencies. We should stress that both versions (designs 1 and 2) perform very well, and the differences, advantages and disadvantages are subtle when listening to the results of a fully calibrated microphone. In particular, the first design (version 1) is surprisingly good in subjective terms, given the handicap of the polar response degradation at mid and high frequencies due to the resonances inherent in the mechanical design. We have also found additional corrections for the encoder that try to minimize unwanted effects at very high frequencies due to the deterioration of the capsule polar patterns at those frequencies.

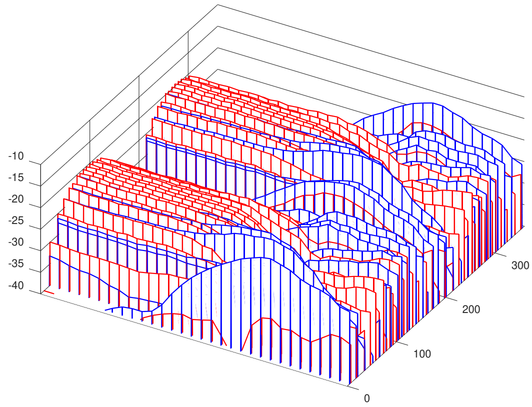

Figure 25. OctaSpHEAR v2 corrected Y waterfall; red traces: -10 to 10 degrees elevation, blue traces: 20 to 40 degrees, 5KHz to 20KHz

6. REFERENCES

[1] Michael Gerzon, "The Design of Precisely Coincident Microphone Arrays for Stereo and Surround Sound", 50th Audio Engineering Society Convention, Preprint L-20, London, 1975

[2] Eric Benjamin, "A second-order soundfield microphone with improved polar pattern shape", Audio Engineering Society Convention Paper 8728, 133rd Convention, San Francisco, 2012

[3] Fernando Lopez-Lezcano, "The *SpHEAR project, a family of parametric 3D printed soundfield microphone arrays", AES Conference on Sound Field Control, July 18–20 2016, Guildford, UK

[4] Fernando Lopez-Lezcano, "The *SpHEAR project update: the TinySpHEAR and Octathingy soundfield microphones", AES Conference on Audio for Virtual and Augmented Reality, August 20–22 2018, Redmond, Washington, USA

[5] Eric Benjamin, "Extending Quasi-Anechoic Electroacoustic Measurements to Low Frequencies", Audio Engineering Society Convention Paper 6128, 117 Convention, San Francisco, 2008

[6] R. H. Hardin and N. J. A. Sloane, "McLaren's Improved Snub Cube and Other New Spherical Designs in Three Dimensions", Discrete and Computational Geometry, 15 (1996), pp. 429-441

[7] Aaron Heller, "Derivation of the A-to-B matrix for a coincident array of first-order microphones", unpublished, 2007