17.0000 Akustische Holografie und Holofonie LU

Measurement Report

Patrick Fahrngruber
Thomas Röck

Supervisor: Univ.Prof. DI Dr.rer.nat. Franz Zotter

Graz, 21.11.2024

kunst uni graz

institut für elektronische musik und akustik

Contents

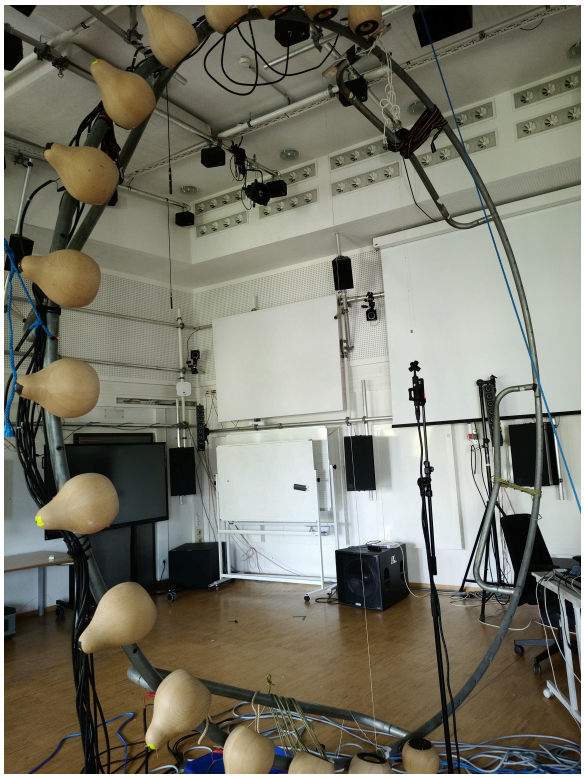
1. Introduction

This report deals with directivity measurements on first-order Ambisonics (FOA) microphones (*Soundfield ST450* and *Røde NT-SF1*), which helps to characterize differences in their properties and directivities. To this end, we measure exponential sine sweeps for different loudspeaker-microphone combinations in the *IEM CUBE* and determine an extensive set of directional impulse responses (IRs). Sixteen loudspeakers are mounted on a vertical, circular metal frame and face towards the center of the circle, where the microphone is located. The microphone stand is positioned on a remote-controlled turntable. The effects of the measurement room, the microphone's frequency response and the slightly unequal source-receiver distances on the IRs are minimized by taking only the direct sound into account and equalizing and time-aligning the IRs with a reference microphone (*Earthworks* Measurement Microphone). These processed IRs are then transformed into the spherical-harmonics domain to interpolate and investigate the microphone's directivity performance regarding a simple beamforming application. The measured directivity data of the *Røde NT-SF1* is available online[1].

2. Measurement

The directivity measurement of the *Røde NT-SF1* was conducted in the *IEM CUBE*. 16 loudspeakers are mounted on a circular frame with a radius of $r = 1.5\text{m}$. The frame is hung vertically from the ceiling as can be seen in Figure 2.1. The microphone is placed in the center point of the frame on a turntable which is set to turn for 10° after each sweep burst for each of the 16 loudspeakers. The 16 loudspeakers result in a elevation resolution of $\Delta\zeta = \frac{180°}{16} = 11.25°$. The first loudspeaker however is placed at $\zeta_0 = \frac{\Delta\zeta}{2} = 5.625°$ as otherwise this would result in redundant measurements for the zenith and nadir respectively. The azimuthal resolution is $\Delta\varphi = 10°$. The excitation signal is an exponential sweep that offers a high Signal to Noise Ratio (SNR) and insensitivity to non-harmonic distortions.

[1] https://phaidra.kug.ac.at/o:135269



a) The measurement setup with the *Røde NT-SF1* as the specimen.

b) An *Earthworks* measurement microphone is used for the loudspeaker equalization.

Figure 2.1: Sweep measurement setup with 16 loudspeakers mounted on a circular frame with a radius of $r = 1.5\text{m}$.

The time-domain signal of the excitation sweep signal is defined as

$$s(t) = \sin\left(\frac{\omega_{\text{start}}\left(e^{tb} - 1\right)}{b}\right) \tag{2.1}$$

where

$$b = \frac{\ln\left(\frac{\omega_{\text{stop}}}{\omega_{\text{start}}}\right)}{N}. \tag{2.2}$$

The start- and stop frequency is chosen as

$$\begin{aligned} f_{\text{start}} &= 19.866 \text{ Hz} \\ f_{\text{stop}} &= 24\,000 \text{ Hz} \end{aligned} \tag{2.3}$$

and the duration of the sweep is set to

$$N_{\text{sweep}} = f_{s_{\text{sweep}}} = 48\,000. \tag{2.4}$$

In order to shorten the duration of the measurement, the multiple exponential sweep method described in [1] is used, in which the sweeps are played back sequentially with

an overlap, which amounts to 16 000 samples, here. Accordingly, the interval in which the sweeps were started was $H = 32\ 000$ samples. For $\frac{360°}{\Delta\varphi} = 36$ azimuthal positions and 16 elevational positions, a total of $4 \cdot 36 \cdot 16 = 2304$ responses could be recorded. To get the time-domain representation of the impulse response, the inverse FFT of the spectral division of the spectrum of the recordings and the spectrum of the sweep is computed:

$$h(t) = \text{IFFT}\left\{\frac{X(e^{j\theta})}{S(e^{j\theta})}\right\} \tag{2.5}$$

which deconvolves the time-domain signals.

The IRs are truncated at $N = 300$ samples to cut off most of the reflections and extract primarily the direct sound. A single IR for all four capsules of the *Røde NT-SF1* is shown in Figure 2.2. The reference measurements with the *Earthworks Measurement Microphone* are used for equalizing and time-aligning the *Røde NT-SF1* IRs. Therefore, a one-sided 64-sample Hann window is applied to the IRs of the reference measurements, starting 64 samples after the initial peak, to ensure a gradual fade-out. Moreover, the IRs are set to zero 64 samples before the initial peak and immediately after the Hann window. To equalize the frequency response of the *Røde NT-SF1* with the *Earthworks Measurement Microphone*, their FFTs have to be calculated and a division leads to:

$$h_{\text{eq}}(t) = \text{IFFT}\left\{\frac{H_{\text{nt-sf1}}(e^{j\theta})}{H_{\text{ref}}(e^{j\theta})}\right\}. \tag{2.6}$$

If some magnitude bins of the reference contain very small values, the magnitude of the equalized IR explodes due to the division. By regularizing the FFT of the reference, by adding a small value (e.g. 0.05) to the magnitude of the FFT, this can be avoided:

$$H_{\text{ref,reg}}(e^{j\theta}) = \left|H_{\text{ref}}(e^{j\theta}) + 0.05\right| \cdot e^{j\angle H_{\text{ref}}(e^{j\theta})}. \tag{2.7}$$

Inserting Equation (2.7) for $H_{\text{ref}}(e^{j\theta})$ in Equation (2.6) leads to our equalized and regularized IRs:

$$h_{\text{eq,reg}}(t) = \text{IFFT}\left\{\frac{H_{\text{nt-sf1}}(e^{j\theta})}{H_{\text{ref,reg}}(e^{j\theta})}\right\}. \tag{2.8}$$

Furthermore, we generate SOFA-files [2] with these IRs and visualize the data with *balloon_holo*[2].

[2] https://git.iem.at/p2774/balloon_holo

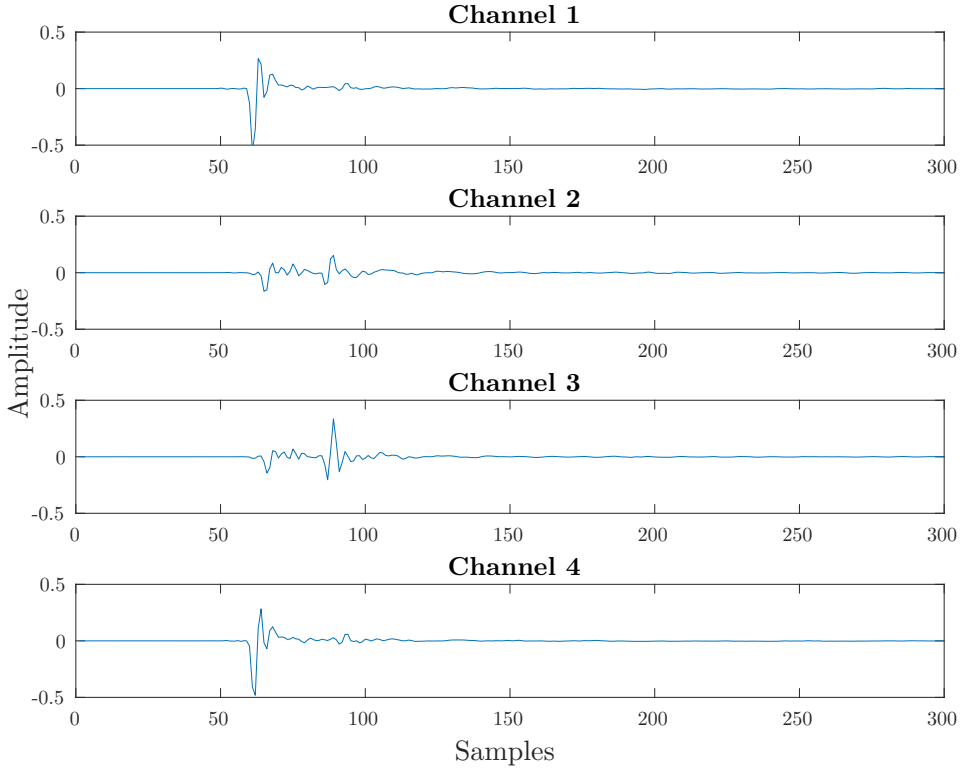


Figure 2.2: IRs of all four channels of *Røde NT-SF1* for the source at azimuth- $\varphi = 0°$ and zenith angle $\zeta = 5.625°$

3. Evaluation Metrics

3.1. Directivity Index (DI)

The directivity index (DI) measures the ratio of the beamformers frequency response at the direction it is steered towards versus the frequency response of all directions. If $H_{0,0}(\omega)$ is the frequency response of the beamformer at the steering direction, then the DI spectrum (in dB) is

$$\mathrm{DI}(\omega) = 10 \cdot \log_{10} \frac{|H_{0,0}(\omega)|^2}{\frac{1}{4\pi} \int_{\varphi=0}^{2\pi} \int_{\zeta=0}^{\pi} |H(\omega, \varphi, \zeta)|^2 \sin(\zeta) d\zeta d\varphi}. \tag{3.1}$$

For a set of N azimuthal measurements with resolution $\Delta\varphi$ and associated M zenith measurements with resolution $\Delta\zeta$, the DI can be approximated with

$$\mathrm{DI}(\omega) = 10 \cdot \log_{10} \frac{|H_{0,0}(\omega)|^2}{\frac{1}{4\pi} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} |H(\omega, \varphi_n, \zeta_m)|^2 \sin(\zeta) \Delta\zeta \Delta\varphi} \tag{3.2}$$

where

$$\begin{aligned} \varphi_n &= n\Delta\varphi \qquad \text{for } 0 \leq n < N, \\ \zeta_n &= m\Delta\zeta + \zeta_0 \quad \text{for } 0 \leq m < M. \end{aligned} \tag{3.3}$$

4. Results

This section deals with the results of the analysis of the two microphones, the *Røde NT-SF1* and the *Soundfield ST-450*, using *balloon_holo*. The data for the *Soundfield ST-450* is provided within the *DirPat* repository by Brandner et. al. [3] which is available online[3].

A 1-st order beam is steered to six different directions $(\varphi, \zeta) \in \{(0°, 0°), (0°, 90°), (0°, 180°), (90°, 90°), (-90°, 90°), (180°, 90°)\}$ and the polar pattern response of the beamformer is analyzed. The matrix to encode the tetrahedral signals of the *Røde NT-SF1* into spherical harmonics domain is

$$\begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ B_3 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ \sqrt{3} \cdot \begin{pmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{pmatrix} \end{bmatrix} \begin{bmatrix} \text{FLU} \\ \text{FRD} \\ \text{BLD} \\ \text{BRU} \end{bmatrix} \tag{4.1}$$

where the subscript i of B_i denotes the Ambisonics Channel Number (ACN) and the vector on the right side of Equation (4.1) contains the four tetrahedral microphone channels.

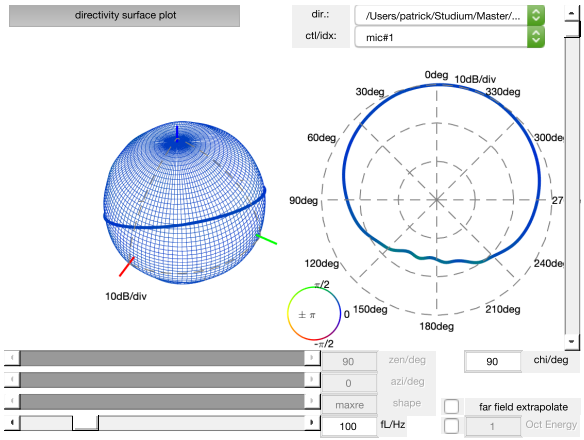
Additionally, the directivity is plotted over a frequency range from 100 Hz to 16000 Hz for a slice orthogonal to the beamformer axis. Finally, the Directivity Index is plotted and compared over the frequency range in question.

4.1. Polar pattern of the microphones

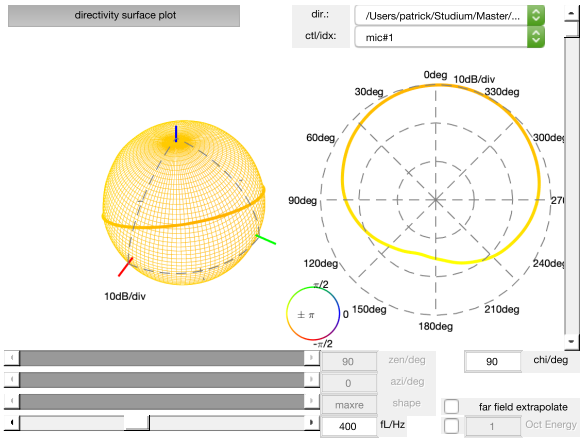
The polar pattern for the first microphone (FLU) of the *Røde NT-SF1* for six distinct frequencies is shown in Figure 4.1. According to its datasheet, the polar pattern should exhibit a perfect cardioid pattern. In our measurements, however, we notice more of a "wide cardioid" or subcardioid with roughly 15 to 18 dB attenuation at the rejective axis.

The output of the *Soundfield ST-450* is already encoded into B-format and the polar plots of its channels are shown in Figure 4.2 at 1000 Hz. As expected, they correspond to the first four spherical harmonics of 1st order Ambisonics and show a well defined directivity pattern. There are small ripples recognizable at directions that were affected by the turntable reflection that was not removed, most notably for the vertical cuts depicted for the W-, X-, and Z-channels in Figure 4.2.

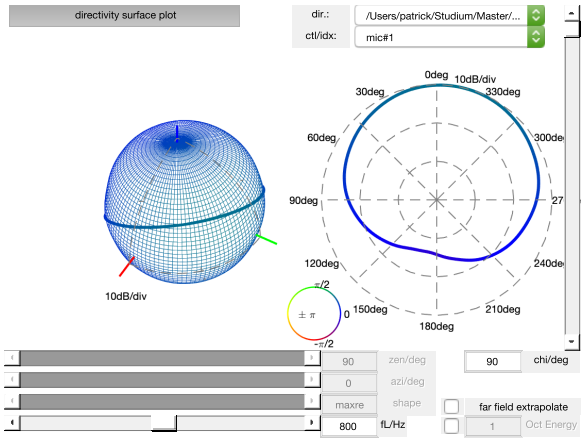
[3] https://phaidra.kug.ac.at/view/o:68229



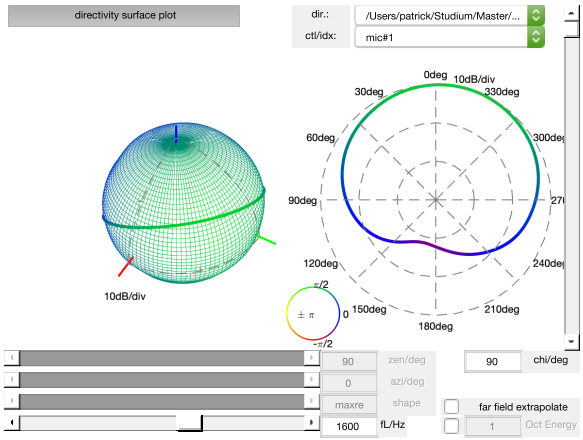
a) 100 Hz



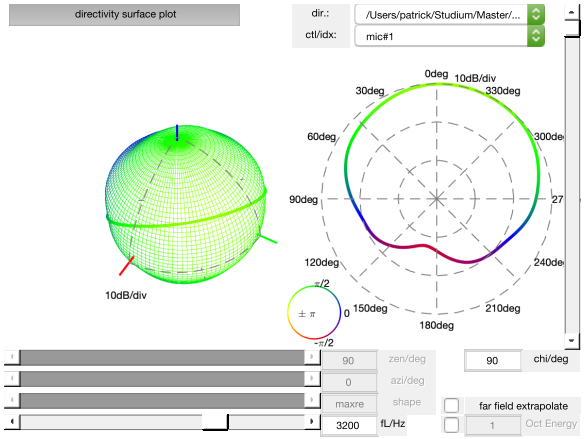
b) 400 Hz



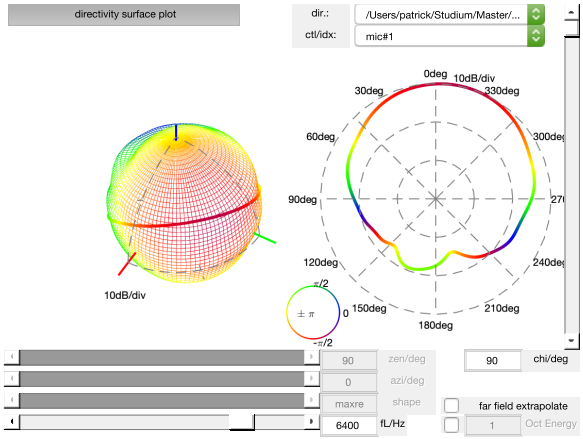
c) 800 Hz



d) 1600 Hz

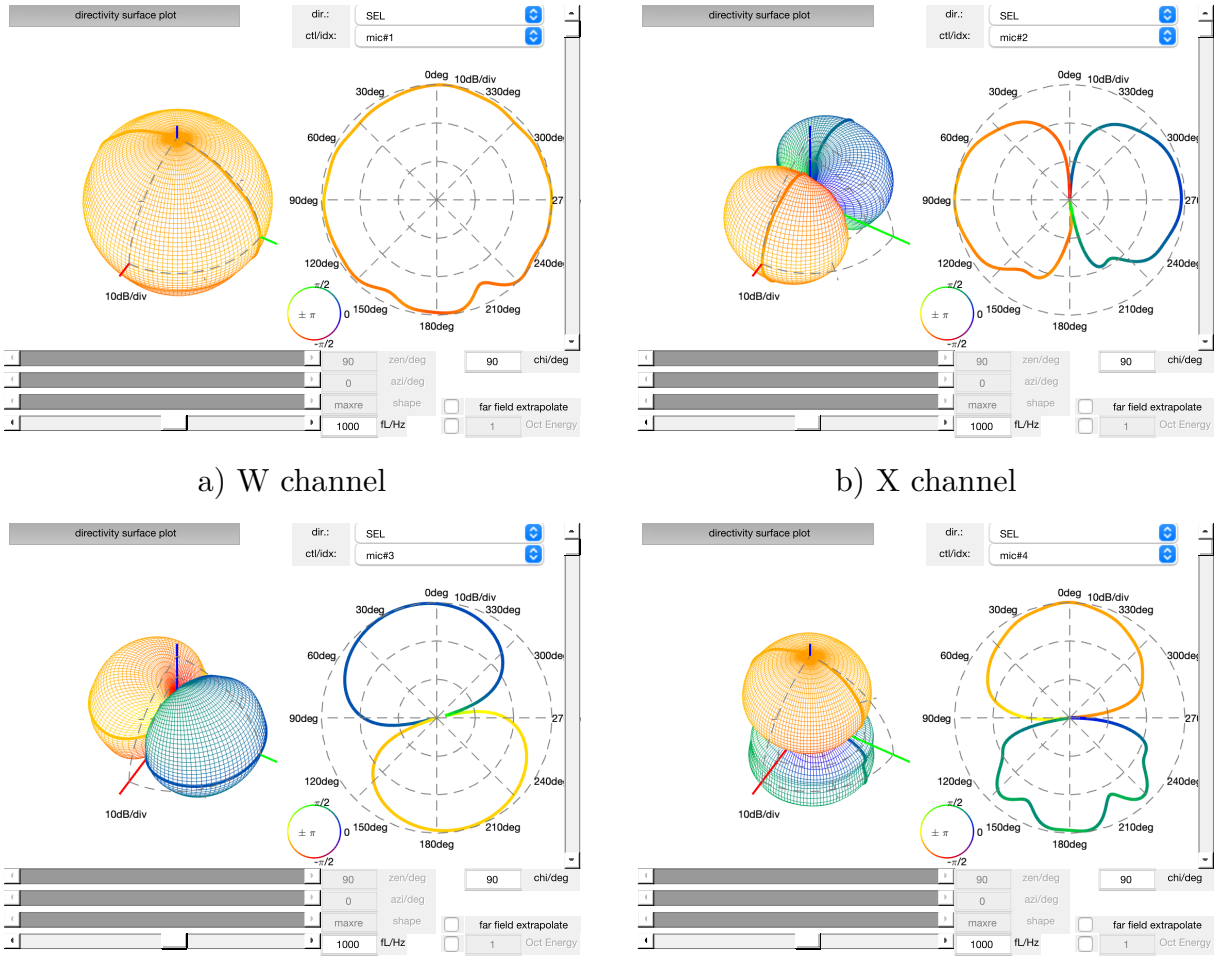


e) 3200 Hz



d) 6400 Hz

Figure 4.1: Polar pattern of channel 1 of *Røde NT-SF1*.



a) W channel b) X channel

c) Y channel d) Z channel

Figure 4.2: Polar patterns of the four B-format channels of the *Soundfield ST-450* at 1000 Hz.

4.2. Shape of the Beamformer

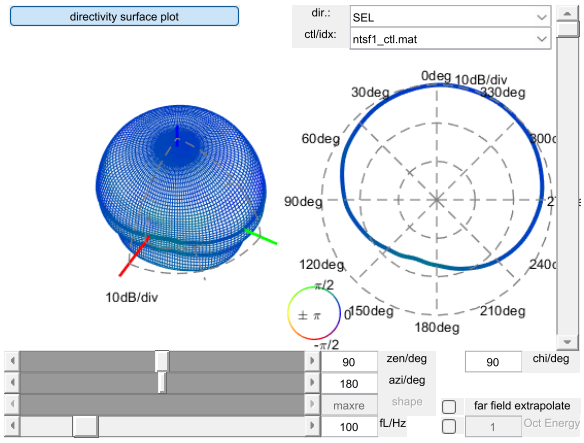
Investigating the balloon plots of the beamformer in Figure 4.3 and Figure 4.4 respectively, it is already noticeable that the *Soundfield ST-450*'s beamformed directivity resembles much more a cardioid than the beamformed directivity of the *Røde NT-SF1*. The former has excellent rejection from the opposite side, especially at 800 Hz, and the directional pattern of a cardioid can be seen very clearly in the view of the slice. The general shape of the beamformer of the *Røde NT-SF1* can be linked more to a subcardioid than a cardioid, as the individual capsules of this microphone already show this characteristic, as can be seen in Figure 4.1. Furthermore, it can be noticed that the *Soundfield ST-450* maintains its beampattern for higher frequencies than the *Røde NT-SF1*, which is especially recognizable in the balloon plots in Figure 4.3 e) and Figure 4.4 e) respectively. We can observe a collapsing directivity pattern at approximately 3200 Hz in Figure 4.5 a) and at 4200 Hz in Figure 4.5 b). The cancellations at 350 Hz and 1050 Hz in Figure 4.7 a) could be related to reflections from the microphone's housing or holder, which could be observable in channel two and three of Figure 2.2

at about 90 samples. Such comb filters are not present up to 3200 Hz in Figure 4.7 b) because of the different geometry of the *Soundfield ST-450*.

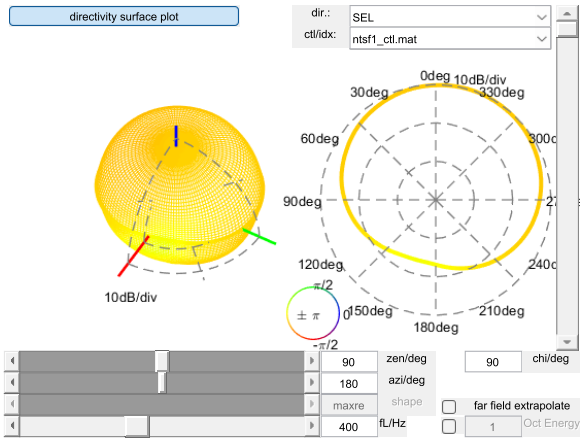
Theoretically, sound cancellation occurs at all multiples of the frequency $f = \frac{c}{2 \cdot a}$, where c is the speed of sound and a is the distance between two microphone capsules. According to [4], the microphone radius of the *Soundfield ST-450* is $r_{\text{ST-450}} = 1.47$ cm , hence the distance $a_{\text{ST-450}}$ between the microphones is $a_{\text{ST-450}} = \sqrt{\frac{2}{3}} \cdot 2r_{\text{ST-450}} = 2.4$ cm. For the *Røde NT-SF1*, the distance $a_{\text{NT-SF1}}$ between the microphones was measured to be $a_{\text{NT-SF1}} = 4.2$ cm and its cancellation frequency would hence be $f_{\text{NT-SF1}} = \frac{343}{2 \cdot a_{\text{NT-SF1}}} = 4083$ Hz. The *Soundfield ST-450* should show comb filtering effects at multiples of $f_{\text{ST-450}} = \frac{343}{2 \cdot a_{\text{ST-450}}} = 7146$ Hz. In the system of microphones, this does not directly yield a comb pattern, but we can observe cancellation phenomena at such frequencies for the *Røde NT-SF1* in Figure 4.5 a) at 0° or the *Soundfield ST-450* in Figure 4.5 b) at 0°. A strong similarity can be noticed between Figure 4.5 and Figure 4.9, which reveals the front-back symmetry of the microphones. Steering the beamformer downwards leads to comb filters caused by floor reflections as in Figure 4.11 a) and to cancellations at ±30° as in Figure 4.11 b). Figure 4.13 and Figure 4.15 show the beamformers steered to $\varphi = 90°$ and $\varphi = 270°$ with a fixed $\zeta = 90°$, which are expected to look similar. This is the case for the *Soundfield ST-450*. The deviation of the *Røde NT-SF1* beamformers could originate from the distortion caused by the 3 dB-step color scale.

4.3. Directivity Index

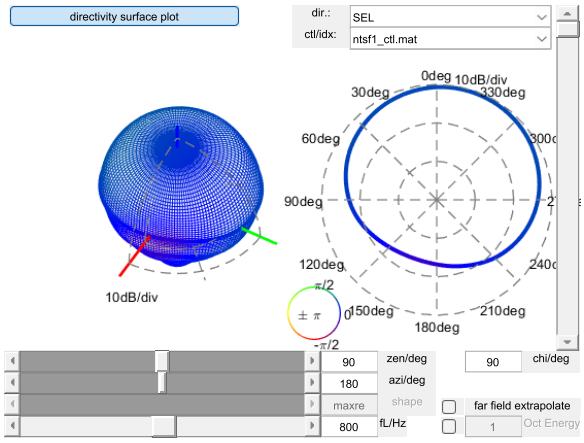
The DI is computed according to Equation (3.2) with $N = 36$, $M = 16$, $\Delta\varphi = \frac{2\pi}{36}$, $\Delta\zeta = \frac{\pi}{16}$ and $\zeta_0 = \frac{\pi}{32}$. The DIs in Figure 4.6, Figure 4.8, Figure 4.10, Figure 4.14, Figure 4.16 show a similar behavior. The DIs increase towards higher frequencies and stay in the range from 3.5 to 6 below 1600 Hz, which is typical for cardioid patterns (a flawless cardioid has a DI of 4.8 dB). Due to its more directional capsules, the DIs are slightly higher for the *Soundfield ST-450*. The oscillating pattern is caused by interference effects from spatial aliasing corresponding to the distance between microphone capsules, reflections and shadowing e.g. by the microphone enclosure, and the directivity of the loudspeakers. The DIs in Figure 4.12 show the major influence of the microphone enclosure, the microphone stand, the rotating platform, and the floor, especially on the *Røde NT-SF1* but with less impact on the *Soundfield ST-450*.



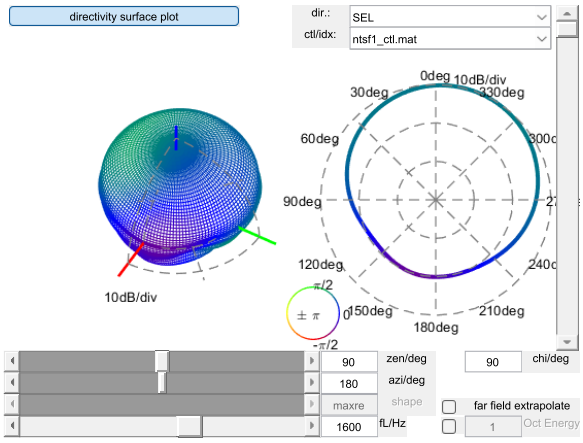
a) 100 Hz



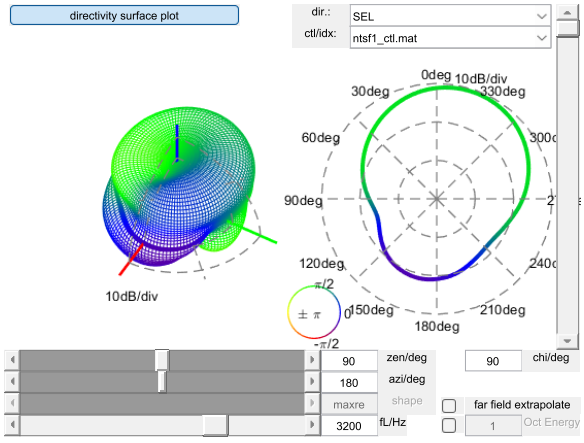
b) 400 Hz



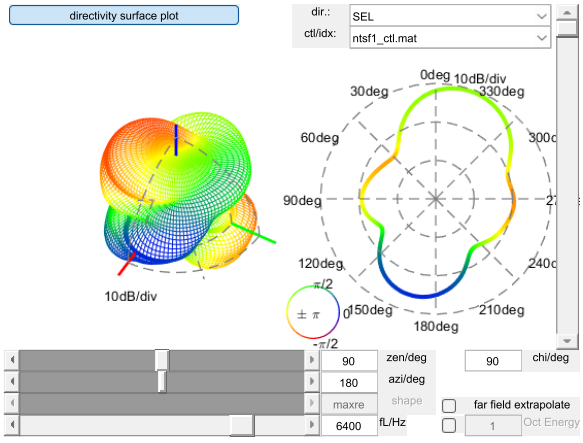
c) 800 Hz



d) 1600 Hz

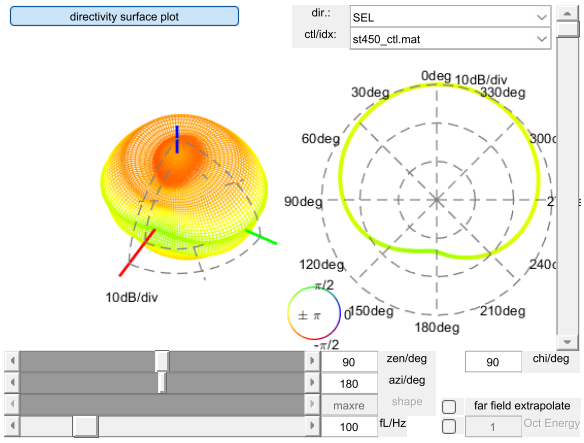


e) 3200 Hz

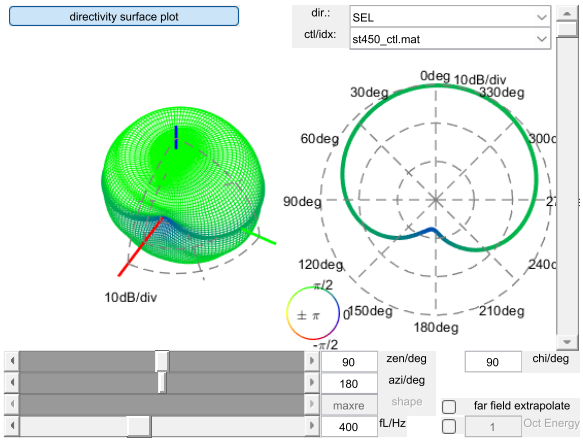


d) 6400 Hz

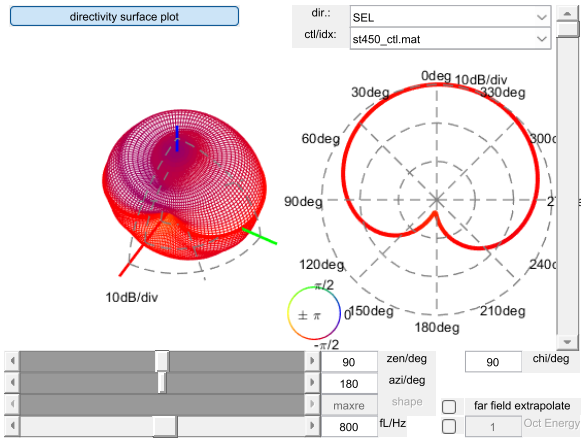
Figure 4.3: *Røde NT-SF1* balloon plots for the beamformer steered to $\varphi = 180°, \zeta = 90°$ at six different frequencies.



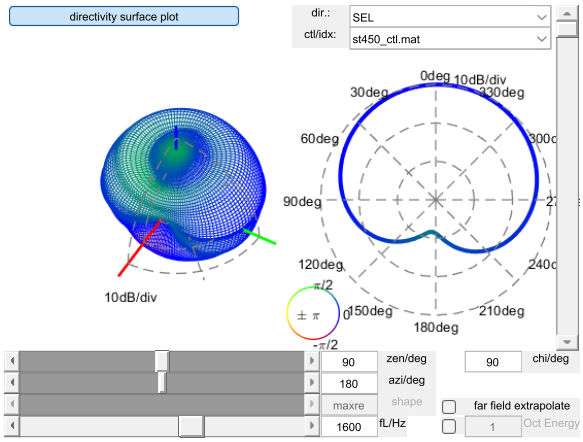
a) 100 Hz



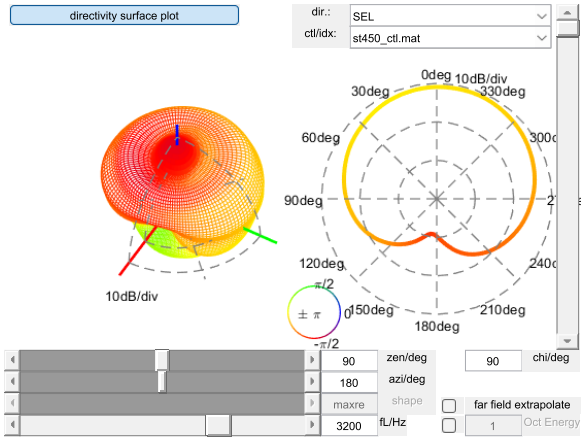
b) 400 Hz



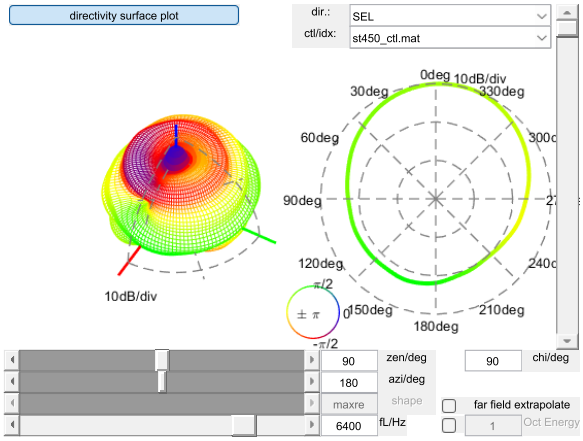
c) 800 Hz



d) 1600 Hz

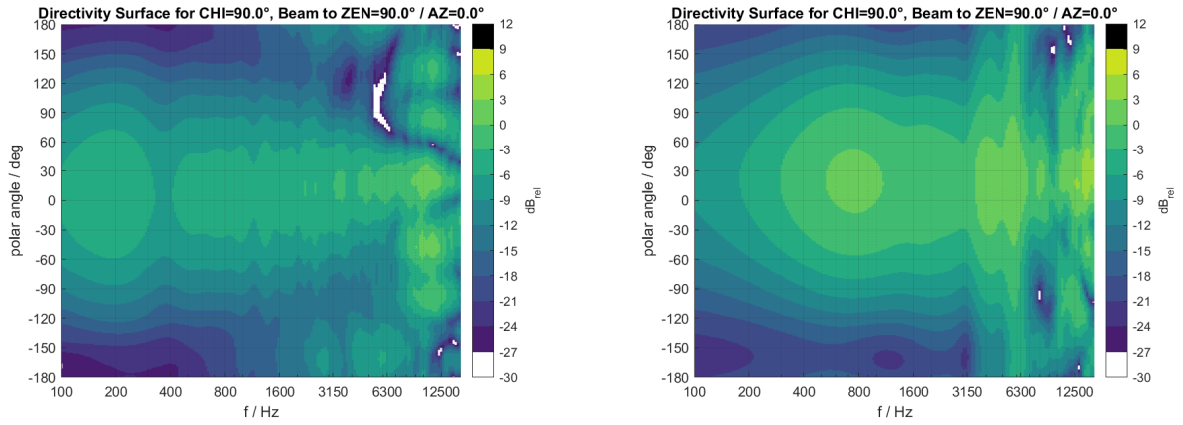


e) 3200 Hz



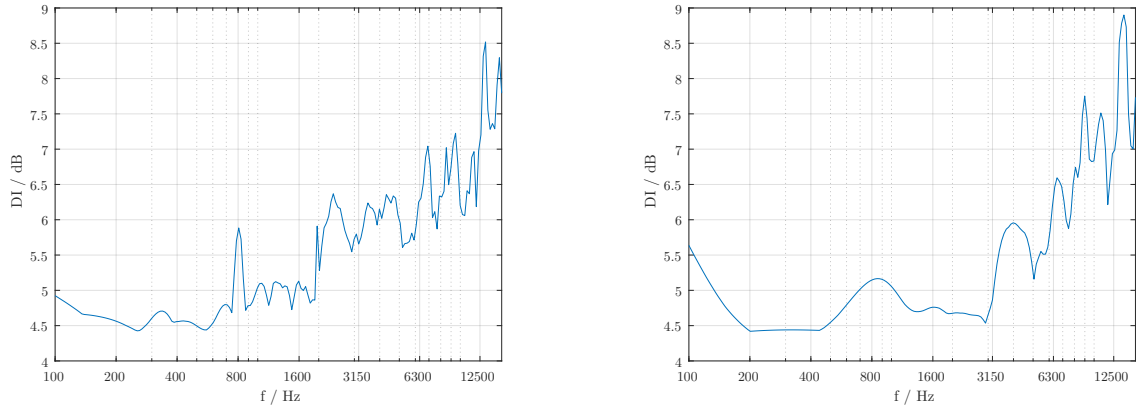
d) 6400 Hz

Figure 4.4: *Soundfield ST-450* balloon plots for the beamformer steered to $\varphi = 180°, \zeta = 90°$ at four different frequencies.



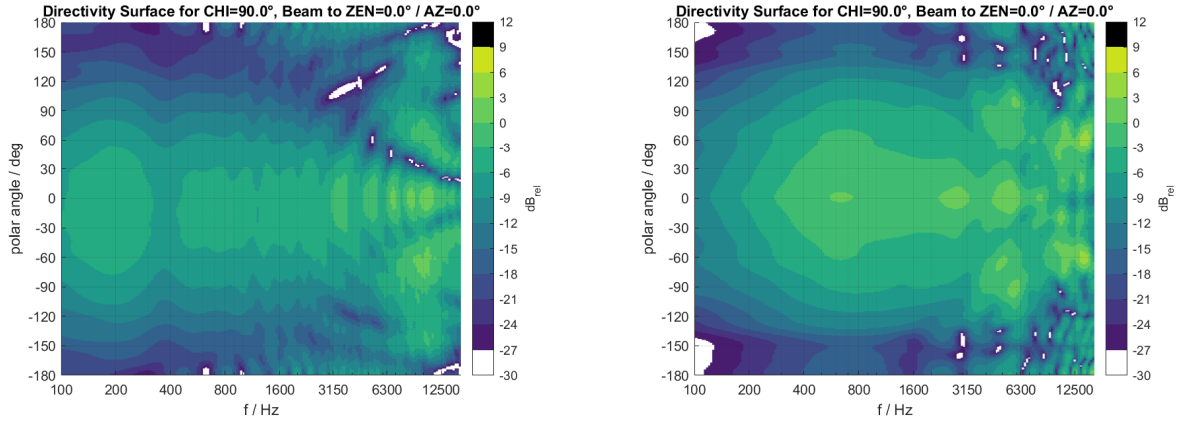
a) *Røde NT-SF1* b) *Soundfield ST-450*

Figure 4.5: Directivity of the beamformer steered to $\varphi = 0°$, $\zeta = 90°$.



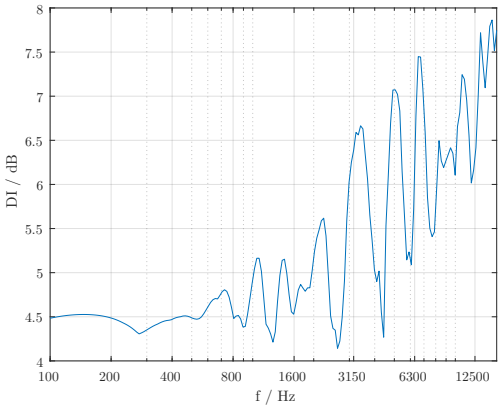
a) *Røde NT-SF1* b) *Soundfield ST-450*

Figure 4.6: DI of the beamformer steered to $\varphi = 0°$, $\zeta = 90°$.

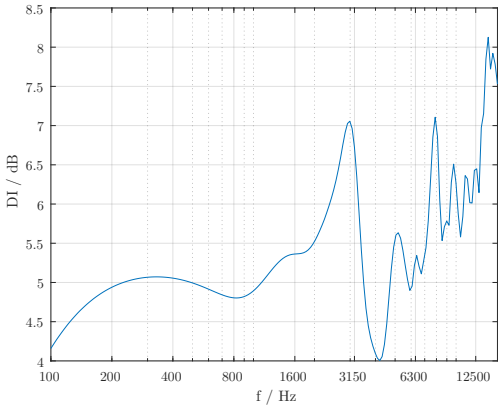


a) *Røde NT-SF1* b) *Soundfield ST-450*

Figure 4.7: Directivity of the beamformer steered to $\varphi = 0°$, $\zeta = 0°$.

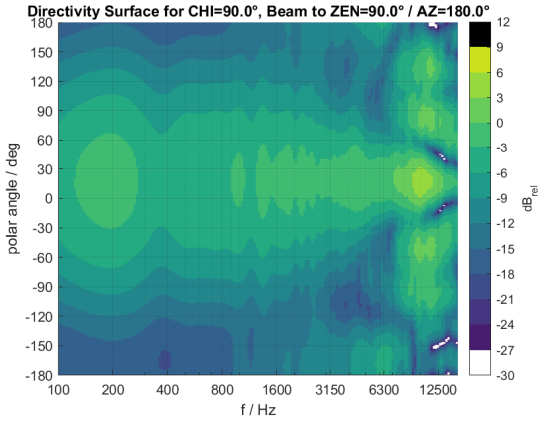


a) *Røde NT-SF1*

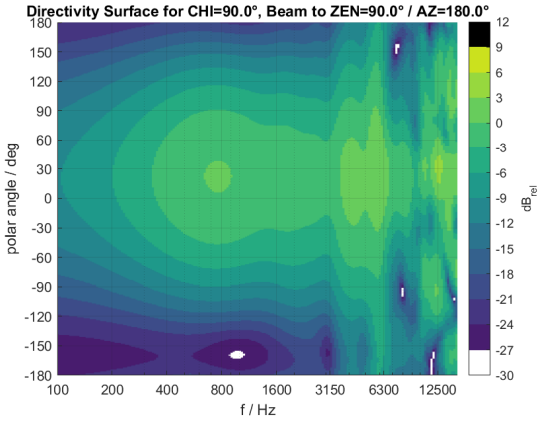


b) *Soundfield ST-450*

Figure 4.8: DI of the beamformer steered to $\varphi = 0°$, $\zeta = 0°$.

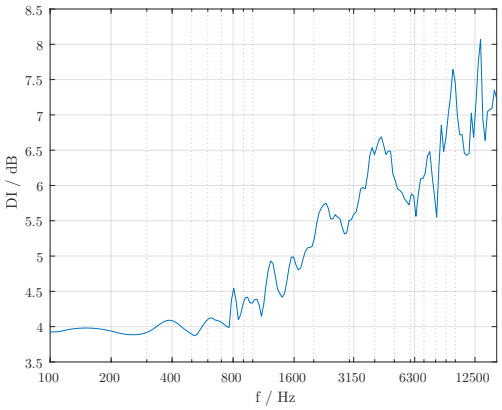


a) *Røde NT-SF1*

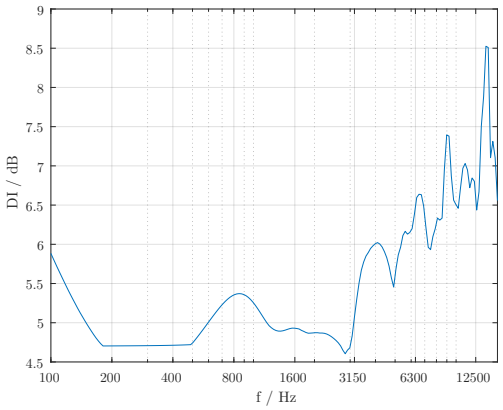


b) *Soundfield ST-450*

Figure 4.9: Directivity of the beamformer steered to $\varphi = 180°$, $\zeta = 90°$.

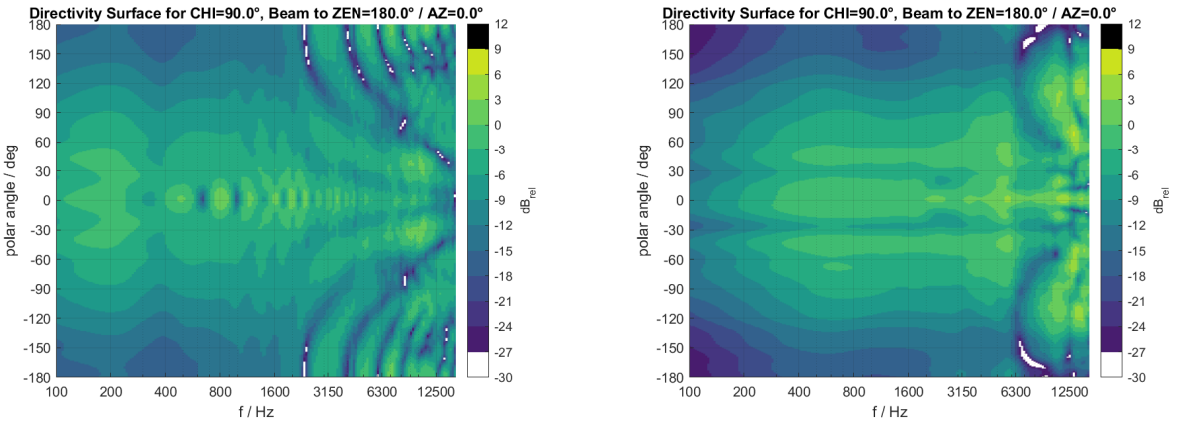


a) *Røde NT-SF1*



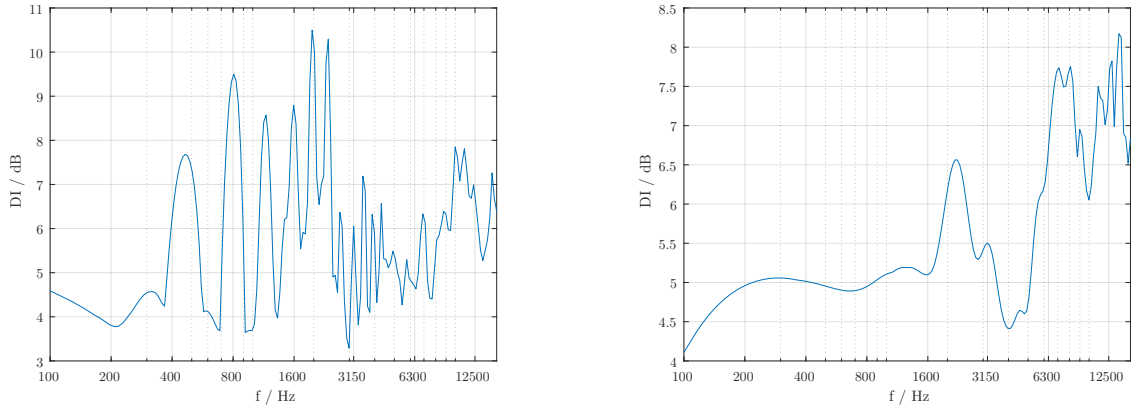
b) *Soundfield ST-450*

Figure 4.10: DI of the beamformer steered to $\varphi = 180°$, $\zeta = 90°$.



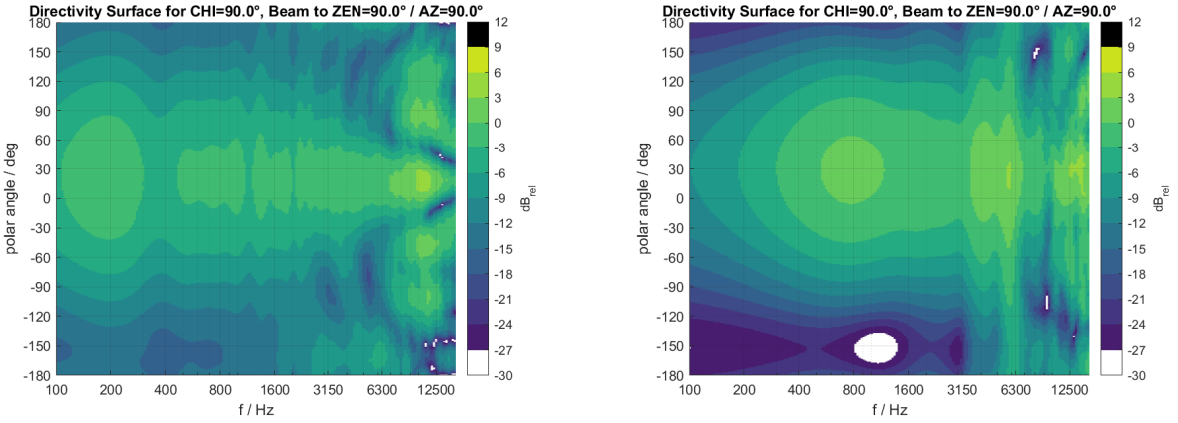
a) *Røde NT-SF1* b) *Soundfield ST-450*

Figure 4.11: Directivity of the beamformer steered to $\varphi = 0°$, $\zeta = 180°$.



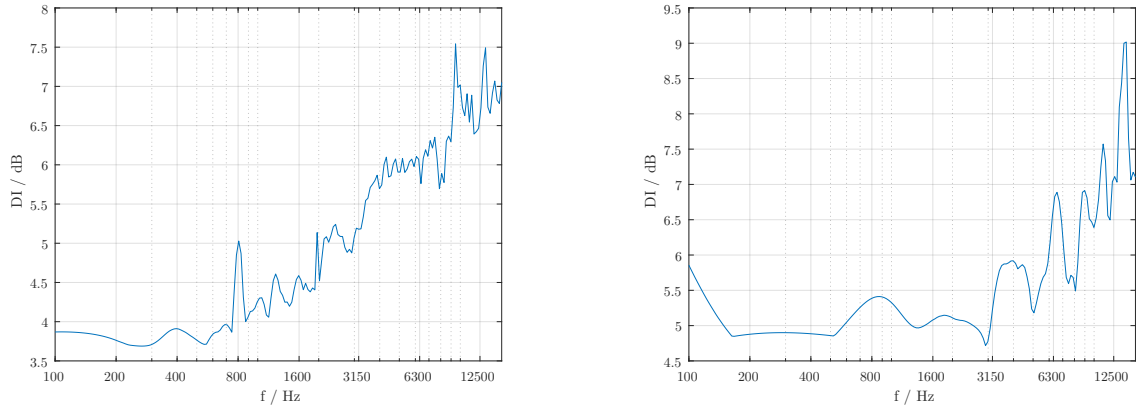
a) *Røde NT-SF1* b) *Soundfield ST-450*

Figure 4.12: DI of the beamformer steered to $\varphi = 0°$, $\zeta = 180°$.



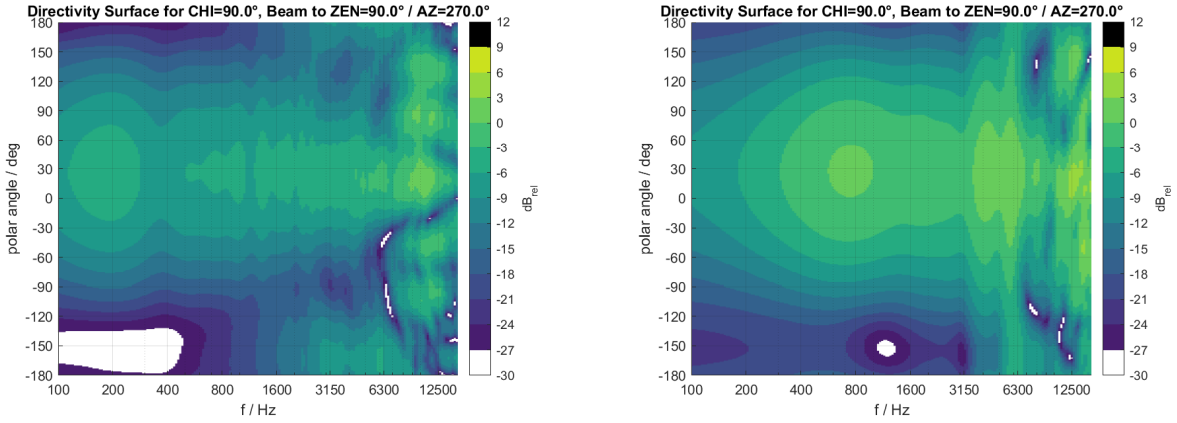
a) *Røde NT-SF1* b) *Soundfield ST-450*

Figure 4.13: Directivity of the beamformer steered to $\varphi = 90°$, $\zeta = 90°$.



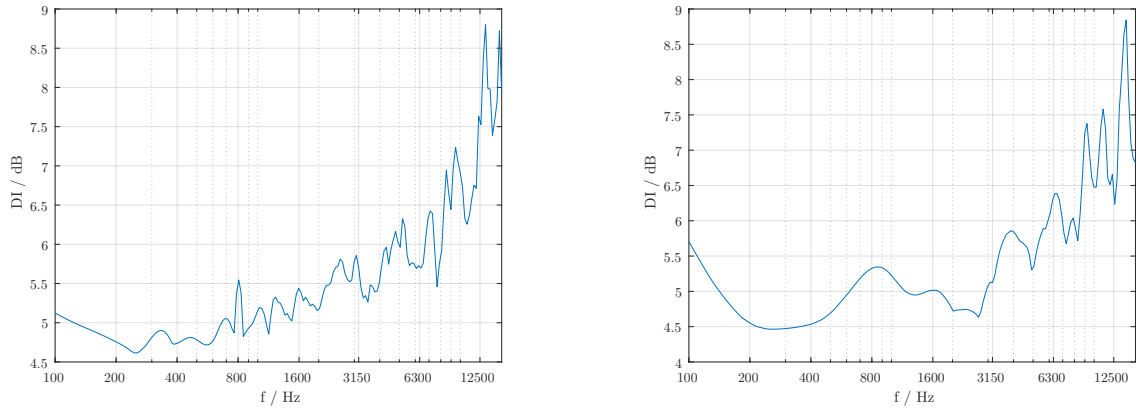
a) *Røde NT-SF1* b) *Soundfield ST-450*

Figure 4.14: DI of the beamformer steered to $\varphi = 90°$, $\zeta = 90°$.



a) *Røde NT-SF1* b) *Soundfield ST-450*

Figure 4.15: Directivity of the beamformer steered to $\varphi = 270°$, $\zeta = 90°$.



a) *Røde NT-SF1* b) *Soundfield ST-450*

Figure 4.16: DI of the beamformer steered to $\varphi = 270°$, $\zeta = 90°$.

5. Conclusion

As part of this practical measurement exercise, the directivity of a tetrahedral *Røde NT-SF1* FOA microphone is measured and compared to already existing directivity data of a *Soundfield ST-450* FOA microphone. The frequency range in which the beampattern is maintained is smaller for the *Røde NT-SF1* than for the *Soundfield ST-450*, the former already shows interference phenomena at 3200 Hz while the latter still exhibits its beampattern shape for frequencies up to 4200 Hz. We can observe that the directivity of the *Soundfield ST-450* is in all cases better than that of the *Røde NT-SF1*, which is reflected both in their polar plots, as well as the DI used as a quality measure. The differences can be traced back primarily to the spacing and characteristics of the capsules, which, for the *Røde NT-SF1* might rather exhibit a slight sub-cardioid than a cardioid pattern. While both microphones can be utilized as immersive 3D audio microphones, the DOA detection capacities of the *Soundfield ST-450* will possibly be more accurate than those of the *Røde NT-SF1*.

Bibliography

[1] P. Majdak, P. Balazs, and B. Laback, "Multiple exponential sweep method for fast measurement of head-related transfer functions," *Journal of the Audio Engineering Society*, vol. 55, no. 7/8, pp. 623–637, 2007.

[2] "AES Standard for File Exchange - Spatial Acoustic Data File Format," no. AES69–2020. 2020.

[3] M. Brandner, M. Frank, and D. Rudrich, "DirPat—Database and viewer of 2D/3D directivity patterns of sound sources and receivers," in *Audio Engineering Society Convention 144*, 2018.

[4] J.-M. Batke, "The B-format microphone revised," *choice*, vol. 2, no. 35.26, p. 9, 2009.