Towards a Better Integration of Room Acoustic and Sound System Design for Multipurpose Venues

James Heddle

Melbourne, Australia

Abstract - A recent landmark performance space experienced a subpar outcome for the installed sound system. Based on the changes that made a successful improvement to the sound and psychoacoustic aspects of room acoustic design, this paper attempts to outline desirable traits for sound systems that should be considered within the context of the acoustic and overall design during the collaborative design process.

1 INTRODUCTION

This paper was motivated by sound system problems in the US$274 million Walt Disney Concert Hall, a 2,265-seat hall with a fixed reverberation time of 2.0 seconds (occupied) intended as an acoustic venue for the performance of symphonic music (the home of the Los Angeles Philharmonic Orchestra).

In essence, the original sound system was designed for good coverage of the audience area whilst preserving sightlines to the stage and to the organ for seats in front of the stage. This resulted in the line arrays being flown 8 metres above the stage level. The arrays were later reconfigured to a central cluster above the stage (Hardiman, A., 2005).

It was found that in the relatively reverberant hall space, the sound system could readily overload the hall, described as ‘overwhelming’ and ‘unpleasant’, and there was a strong mismatch between the sound from the loudspeakers and that coming from the stage, particularly for the seats closest to the stage.

In an attempt to improve matters, Billy Woodman, the technical director and owner of ATC Loudspeaker Technology (manufacturers of recording studio monitors), was asked to have a look at the issue and, as an initial step, a pair of large studio monitors was trialled on the stage. A notable aspect of ATC monitors is the use of a 75mm diameter soft dome driver that has a very smooth frequency response and enables a wide dispersion up to around 10kHz (See Appendix A.4).

To summarise the findings, the significantly improved fidelity of the studio monitors compared to the concert line arrays was preferred and the co-location of the electro-acoustic and acoustic sources on the stage provided a more seamless blend for listeners. Much larger versions of the studio monitors were subsequently provided as the main option for reinforcement of live sound in the hall (see Figure 1).

Acoustics in the sun
Acoustics 2024 Gold Coast Queensland

Figure 1 – Disney Concert Hall Stage Located Sound

2 PSYCHOACOUSTIC ASPECTS

In discussing venues and sound systems it is important to underscore that the auditory system is highly sophisticated and adaptive. This means that there are a large range of less than ideal rooms and audio systems that listeners adapt to and even view as ‘normal’/’comfortably familiar’. However, given the opportunity for rapid comparison between the sound of two or more options, with sources of bias minimized, a truer picture of preference and desirable acoustic characteristics emerges.

2.1 Engagement and Precedence

In the case of an acoustic performance being augmented by a sound system, the sense of being strongly engaged with a performance, and minimally aware of the sound system and the room, is reliant on providing the senses with information consistent with the visual stage sources being *the* sound sources and the remaining sound being consistent with the stage sound as the primary source, minimizing clues that could detract from the illusion.

This means that for an ‘acoustic’ performance the direct sound from the stage should arrive at a listening position before other sound, preferably with the first reflections or sound system sound arriving more than 6msecs after the direct sound, to allow the auditory system to separate out the direct sound, but less than 30 msecs after (Heddle, J., 2016).

The spectrum and temporal envelope of the following sound should match that of the direct sound (Lokki, T. & Pätynen, J., 2020) in order that the following sound is suppressed in awareness and localization is to the direct sound source, the precedence effect. Additionally, the sound following the direct sound should not be more than ~15dB greater in level than the direct sound or there will be a shift in the apparent direction of the source to that of the louder sound (Moore, B.C., 1999).

The earliest and strongest non-direct sound energy should not arrive from above, such as may occur from an elevated sound system or due to ceiling or overhead reflectors, as this tends to cause colouration (Theile, G., 1980) and an increase in perceived distance to the source. It should arrive from the sides, ideally within a range of azimuth angles around 70° (±15°) and slightly elevated, 24°±10° for best binaural gain and benefit.

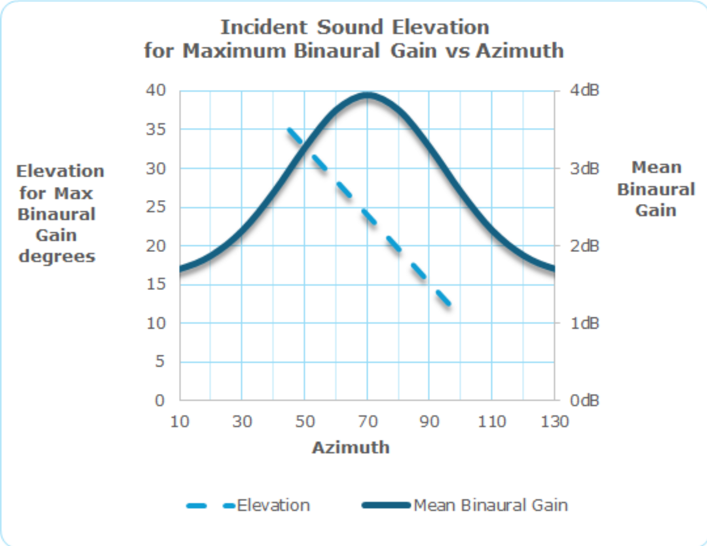


Figure 2 – Incident Sound Elevation for Maximal Binaural Gain vs Sound Incidence Azimuth (adapted from Fig 4., Lokki, T. & Pätynen, J., 2020)

2.2 Fidelity

In a study conducted at the Nordea (now Alexela) Concert Hall in Estonia using live music[1], the most influential factor on overall pleasantness was the fidelity of the sound system, the naturalness of sound (Tereping, A.-R., 2016). Overall pleasantness was defined as "the sound of this quality is pleasant to me, or I find sound of this quality to be unpleasant". Fidelity was defined as "how similar the presentation is to the original; natural sound".

"Therefore, when using a sound reinforcement system, one should focus on the fidelity of the sound in particular. Fidelity, not loudness, is also the best way to judge the quality of a sound reinforcement system and the competence level of a sound engineer" (Tereping, A.-R., 2016). The L_{Aeq} level at listeners found to be the most pleasant was 78 dBA (ranging somewhat with musical type and style between 73 dBA and 85 dBA) with higher levels perceived as less, not more, pleasant although the correlation was not high (r^2=0.297).

In a double-blind test involving 268 listeners, listeners preferred those loudspeakers that were perceived to be spectrally flat and this correlated with the average frequency response in a listening window with an angular range of ±10° elevation from on axis and ±30° from on-axis in the horizontal (Olive, S.E., 2004a,b), i.e. not just flat on-axis but flat and smooth off axis.

Other aspects that affect the fidelity of reproduced sound from loudspeakers are resonances which may degrade both the time response and the frequency response. Higher order modes in horns occur above a cut-on frequency and these have longer travel paths down the horn than that of the fundamental (plane wave) mode. This causes time smear of the signal and the characteristic horn sound (group delay distortion). Bass reflex alignments for bass are essentially resonance systems to obtain more bass for 'free', compromise transient performance and can mask detail (Harris, L., 2015).

"Impulsive signals will be smeared in time as bass notes arrive late compared to other signal content, and low level detail in the musical signal may be masked by the extended ringing on of one part of the response where the resonance occurs" (Harris, L., 2015).

[1] 146 participants. The sound level, LAeq between 60dBA and 100dBA, and the sound system configuration (7 variations) was randomized for each live music sample (7 different musical pieces). Refer to the reference for more details.

3 SOUND SYSTEM

3.1 Desirable Traits

The desirable traits for the sound system mainly relate to the system directivity, fidelity and location and include:

1. The system should be stage located such that the apparent sound source during performances appears to be from the stage rather than dislocated to somewhere else in space.
2. The sound system 'spill' to performers and performer microphones should be minimized and preferably not spectrally coloured, such as a low pass version of the direct sound.
3. The spectrum of the direct sound from the sound system should be consistent with the stage sound and be prioritised for fidelity, free of colouration and resonances, having a smooth and flat extended frequency response and transparent. For best fidelity, horns should not be used, and bass should be from sealed cabinets or open baffle loudspeakers[2] (Linkwitz, S., 1998) and not bass reflex/vented cabinets.
4. The direct sound from the sound system should not vary in spectrum between listening positions but may vary in level. The same requirement applies to the spectrum via earliest reflections from the sound system to listening positions. These aspects imply that the sound system should have both wide dispersion and frequency independent (constant) directivity, at least in the horizontal, and that with constant directivity room reflections of the sound system do not need to be minimized or avoided. Frequency independent directivity also means that the sound power and the direct sound have a similar frequency response. A spectrum shape that does not vary with listening position also implies that diffraction, such as from cabinet edges, is minimal and that the source is acoustically small relative to wavelength in the horizontal.
5. A constant sound level between near and far listening positions is not a priority. The system should be able to provide an LAeq level at the majority of listening positions of around 78dBA with 22 dBA headroom (classical music long term crest factor), i.e. a 100dBA maximum capability or thereabouts.

3.2 Feasible Solutions

There are three broad options for achieving broadband frequency independent directivity, i.e. from the bass through to the high frequencies. These are full bandwidth omnidirectional, full bandwidth cardioid and full bandwidth dipole. Maintaining a directivity at the highest frequencies consistent with that at the frequencies below is difficult to achieve as the transducer size becomes much larger than the wavelength. This means that there will be an upper frequency limit to the broadband constant directivity.

The omni option is readily achievable at low frequency but becomes much harder at higher frequencies and spill to the stage, particularly bass, is the highest of the options. Full bandwidth cardioid and dipole are the two other options with the cardioid requiring more drivers, extra complexity and with the cardioid directivity only approximated. Both have the distinct advantage of a better direct to reverberant level of 4.8dB due to their directivity and this is particularly advantageous in the bass and for reduced spill to the stage. A dipole response can be achieved very simply by not having an enclosure, with the dipole characteristic controlled by the effective open baffle size (no electronic processing required). Equalization and signal processing is then necessary to correct for the frequency response of the driver, the dipole response characteristic and to achieve the required crossover target.

[2] Compared to standard cabinet systems, open baffle/dipole loudspeaker designs have less linear distortion (no radiation from cabinet panels, no cabinet cavity resonances and energy re-radiation via the driver cone), excite the room modal response to a much lesser degree and better preserve the information content of a signal, the sonic detail, as evidence by the better Modulation Transfer Function performance (see reference).

Figure 3 shows an example 6m high dipole line array location and directivity in plan relative to the stage and audience (stage right shown) implemented in a 1,000 seat concert hall. The system was intended to augment live acoustic performance.

The array location and orientation (toed-in 24°) were chosen to achieve the dipole null being directed at the stage and coverage of the entire seating area (the on-axis direction the rear centreline of the seating). The array was also tilted back 5° to match the slope of the seating.

It can be seen that the early side wall reflection surfaces and the entire seating area all fall within the ±60° -6dB horizontal beamwidth of the dipole array whereas the sound energy to the stage is highly attenuated (at all frequencies). The vertical section is shown for context of the vertical array constant directivity discussion given in the Appendix. The origin, for the purposes of vertical directivity and apparent height, is taken back from the physical location of the array as a result of the radius of curvature of the array in the vertical (delay-based or actual) at stage height.

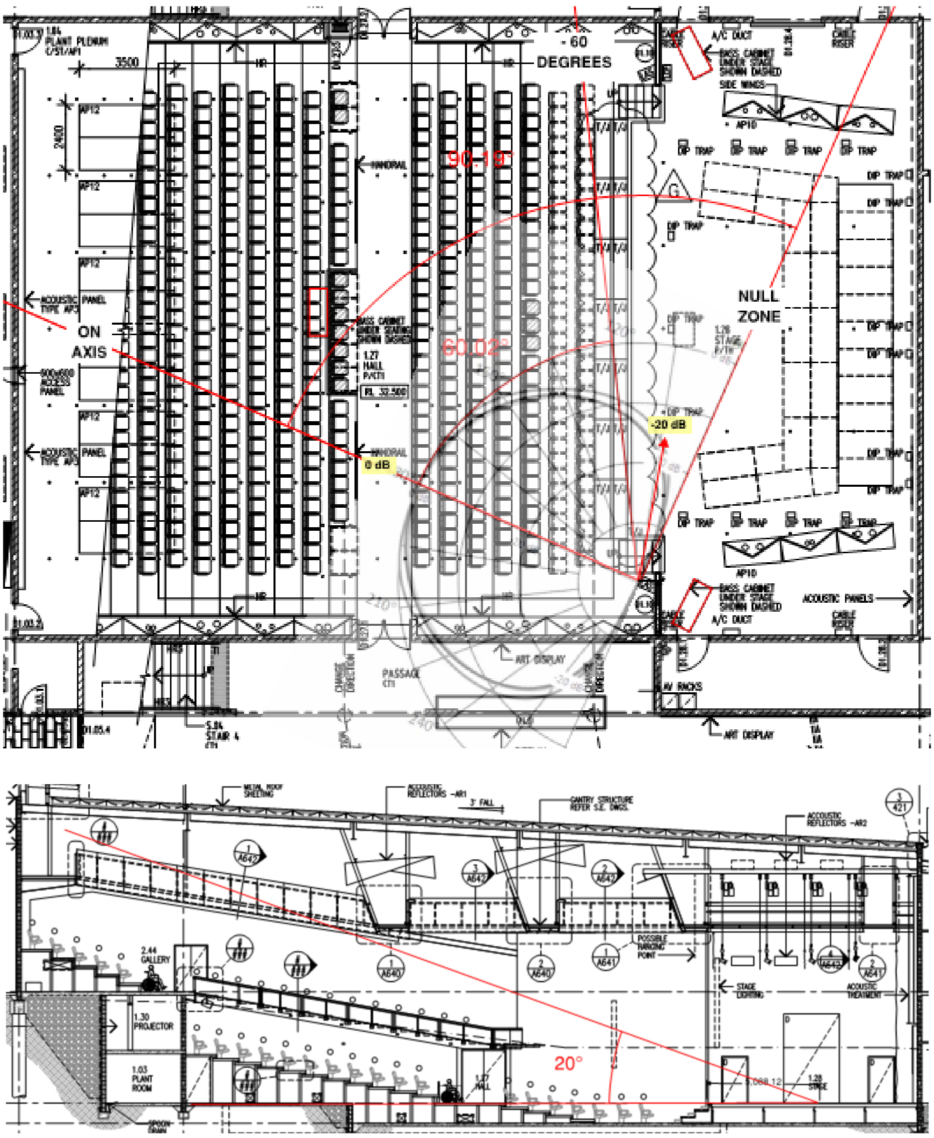


Figure 3 – Example Dipole System Location in Hall relative to Directivity

4 SUMMARY

Working back from the clues from the Walt Disney Concert Hall experience, desirable psycho-acoustical features and the goal of a more uniform experience for all of the audience:

1. The design of the venue should consider the requirements and constraints of the desirable sound system characteristics, particularly directivity, at an early design stage, not as a later add-on. The form of the venue or hall needs to be developed bearing this in mind from the outset. The shoe-box type hall shape is a suitable form, and successful acoustically. Other forms such as terraced surround/non-shoebox halls are more difficult geometries for uniform coverage and require more compromises.

2. The sound system should be stage floor or near stage floor located but with minimized spill to the stage.

3. The early lateral reflection surfaces should be designed for both the stage sources (performers) and stage floor zone located sound system to arrive at listeners ideally within a range of azimuth (horizontal) angles around 70° (±15°) and, preferably, slightly elevated. The delay and level of the sound system need to be adjusted so that the apparent audio source is from the stage performance.

4. The fidelity of the sound system should be given much higher priority. Particularly considering the very high cost of modern halls, the sound system should be recording studio quality with the desirable traits as given in section 3.1.

5. The spectrum off axis of system loudspeaker elements should mimic the on-axis spectrum, as far as possible, both in the horizontal and the vertical (not just in some limited angle or frequency range) and there should be a wide horizontal (~60°) -6dB beamwidth. This results in the audience area sound spectrum and the sound towards early reflection surfaces being spectrally consistent, no bad seats, and equalization becomes global rather than a compromise between listener locations. Most commercially available sound systems do not achieve the requirements.

6. A full range dipole line array sound system is one feasible option to achieve the requirements. It is noted that a dipole array has the feature of being bidirectional which may be of use in some audience layouts. The requirements to achieve frequency independent directivity in the vertical for a dipole array are given in the Appendix A.1, A.2.

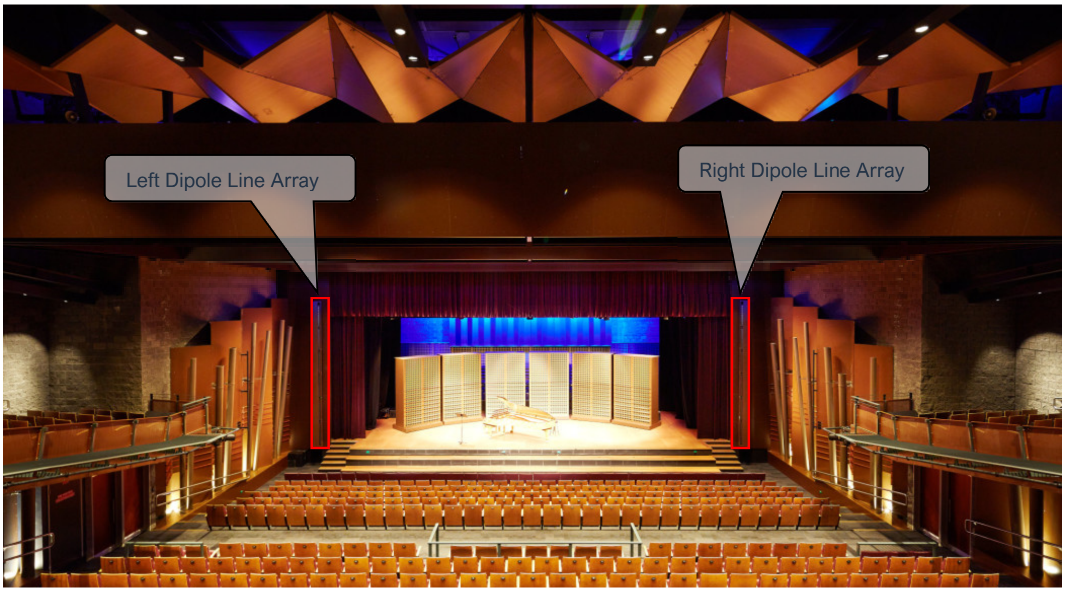


Figure 4 – View to Stage in Hall with Dipole Array Stage Left and Right

ACKNOWLEDGEMENTS

The author thanks Tapio Lokki for useful comments on the draft of the paper and is grateful for the contributions to, and communications on, the design of dipole loudspeakers by Siegried Linkwitz, now deceased. A wealth of information for those curious on loudspeaker design, and dipole loudspeakers in particular, is available at the Linkwitzlab website, https://linkwitzlab.com/models.htm

REFERENCES

Hardiman, A., 2005. *Revisiting the Performance Sound System.* Lighting & Sound America March 2005

Harris, L., 2015. *Development of a modulation transfer function-based method for evaluating bass reproduction accuracy in professional monitoring loudspeakers*. University of Southampton, ISVR PhD Thesis

Heddle, J., 2016. *Room Acoustics for Listening*. Proceedings of Acoustics2016, Brisbane, Australia.

Linkwitz, S., 1998*. Investigation of sound quality differences between monopolar and dipolar woofers in small rooms*. In Audio Engineering Society Convention 105.

Lokki, T. & Pätynen, J., 2020. *Auditory Spatial Impression in Concert Halls*. In The Technology of Binaural Understanding. Springer, pp. 173–202.

Manke, K. et al., 2017. *Implementation of a Dipole Constant Directivity Circular-Arc Array*. In Audio Engineering Society Convention 143.

Moore, B.C., 1999. *Controversies and mysteries in spatial hearing*. In Audio Engineering Society Conference: 16th International Conference: Spatial Sound Reproduction.

Olive, S.E., 2004a. *A multiple regression model for predicting loudspeaker preference using objective measurements: Part I-Listening Test Results*. In Audio Engineering Society Convention 117.

Olive, S.E., 2004b. *A multiple regression model for predicting loudspeaker preference using objective measurements: Part II-Development of the Model*. In Audio Engineering Society Convention 117.

Taylor, R., Manke, K. & Keele Jr, D., 2017. *Constant directivity circular-arc arrays of dipole elements*. In Audio Engineering Society Convention 143.

Tereping, A.-R., 2016. *Listener preference for concert sound levels: Do louder performances sound better?* Journal of the Audio Engineering Society, 64(3), pp.138–146.

Further information on the test hall can be found at https://www.kontserdimaja.ee/korraldajale/venue-sheet/

Theile, G., 1980*. On the localisation in the superimposed soundfield.* Technische Universität Berlin.

APPENDIX

A.1 Dipole Array Shading for Constant Directivity

For a sound array arc of radius a from a given origin point, composed of radially oriented dipoles, it can be shown that the array will have a far field frequency independent (constant) vertical directivity above a minimum frequency if a specific angle dependent attenuation of output level (shading) of the array elements can be achieved (Taylor, R. et. al., 2017). The shading varies as a function of the elevation angle α, S(α), and can be expressed as a Fourier cosine series.

$$S(\alpha) = \sum_{n=0}^{\infty} a_n \cos n\alpha = a_0 + a_1 \cos\alpha + a_2 \cos 2\alpha + a_3 \cos 3\alpha + a_4 \cos 4\alpha \ldots \qquad \text{(A.1)}$$

, with the requirement that the absolute value of the sum of the even n components is equal to the absolute value of the sum of the odd n components, $|S_{even}(\alpha)| = |S_{odd}(\alpha)|$ for all α. The directivity control applies above a minimum frequency given by

$$f_{min} = \frac{n_{max} \cdot c}{2\pi a \cos\varphi} \qquad \text{(A.2)}$$

below which, in theory, the array acts as a single dipole at the origin (centre of curvature, apparent source location) of the array. The n_{max} in the formula is the largest n for which the cosine series coefficients are non-negligible, i.e. if the coefficients a_6 and above are negligible then $n_{max} = 5$. Φ is the angle from on-axis laterally, c is the speed of sound in metres per second, a is the radius of the array.

One approach to finding these shading functions is via algorithmic searching for a_n coefficients meeting the requirements and there are number of shading functions that meet the requirement for constant directivity. As n_{max} increases f_{min} increases, the matching of the even and odd components becomes more exact, the array beamwidth may reduce and, with this, the output from the sound system relative to its potential maximum reduces as there is an overall greater attenuation of the array elements.

A characteristic of constant directivity array shading is minimized lobing together with a 3dB/oct rise in the frequency response above f_{min}, which needs to be equalized.

A.2 Shading Examples

Array coefficients that meet the requirements for n_{max} = 3, 4 and 6, found by genetic algorithm optimization[3], are given below. For convenience and accuracy, the coefficients are expressed as their inverse, i.e. $b_n = 1/a_n$.

The graph shows the required shadings, $20\text{Log}_{10}[S(\alpha)]$, derived using the coefficients with cosine shading shown for comparison. Note that the shadings also define the array directivities.

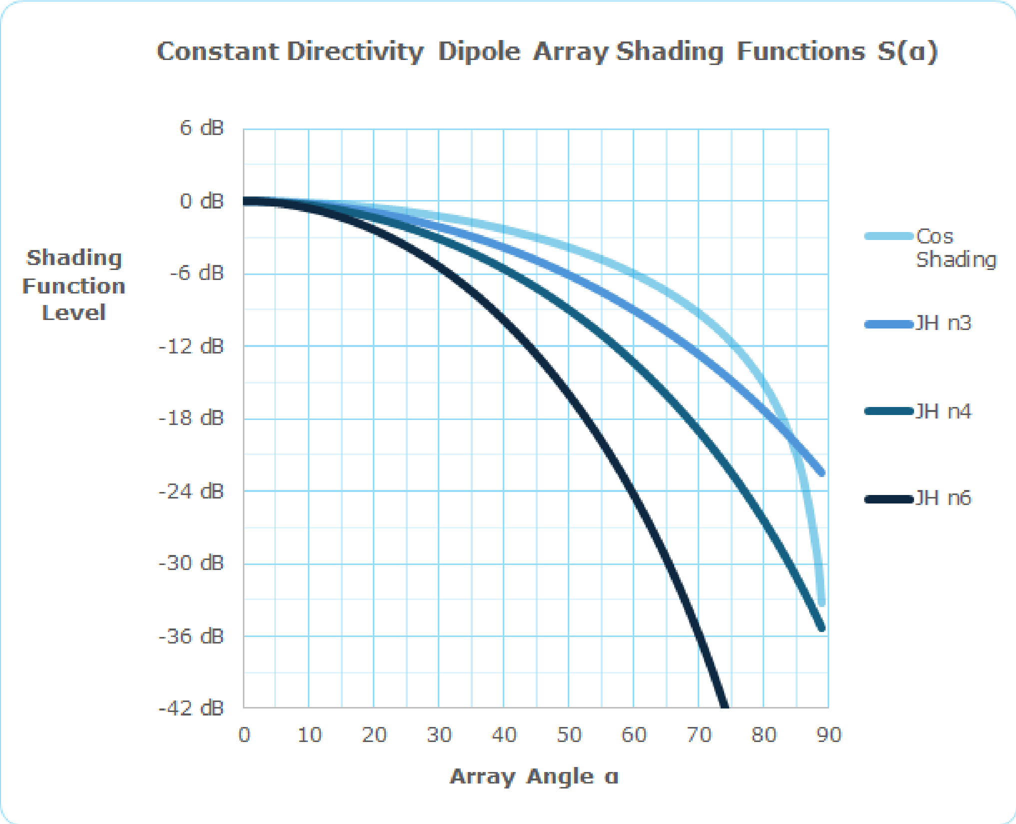
For an array radius of 5.5 metres, the n_{max} =4 shading would have an f_{min} of 40Hz on axis.

n_{max} =3: b_0 = 3.5207 [Constant];
b_1 = 2.2147 [Cosα];
b_2 = 4.6380 [Cos2α];
b_3 = 20.4775 [Cos3α]

n_{max} =4: b_0 = 4.1917 [Constant];
b_1 = 2.4674 [Cosα];
b_2 = 4.1095 [Cos2α];
b_3 = 10.4686 [Cos3α];
b_4 = 51.8709 [Cos4α];

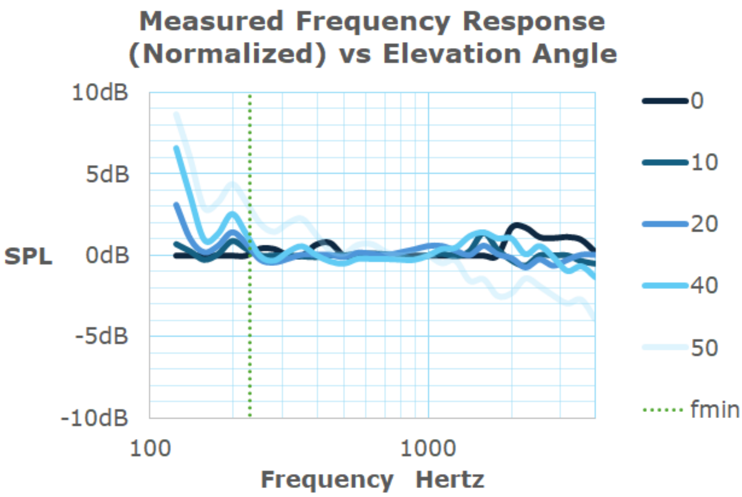
n_{max} =6: b_0 = 5.4799 [Constant];
b_1 = 3.0205 [Cosα];
b_2 = 4.0701 [Cos2α];
b_3 = 6.8344 [Cos3α];
b_4 = 14.9025 [Cos4α];
b_5 = 46.0240 [Cos5α];
b_6 = 255.425 [Cos6α]

[3] Some variation in the coefficients, and therefore the shading function, will occur depending on the cost function. The cost function used had decreased weights for far off axis angles where the array is much more attenuated in level.



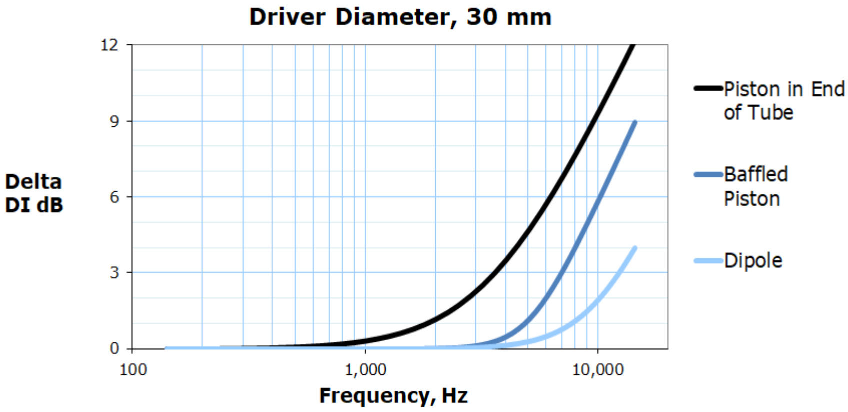
Measurements, in the useable frequency range of the dipole driver, of the variation with elevation angle for a 10-element circular dipole array with constant directivity shading are shown in the graph (adapted from Manke, K. et al., 2017).

The level normalized frequency response variation, with the dipole and loudspeaker response equalized, in the range of elevations between ±40° is within ±2dB above the design frequency of 230Hz (fmin), validating the theoretical result.



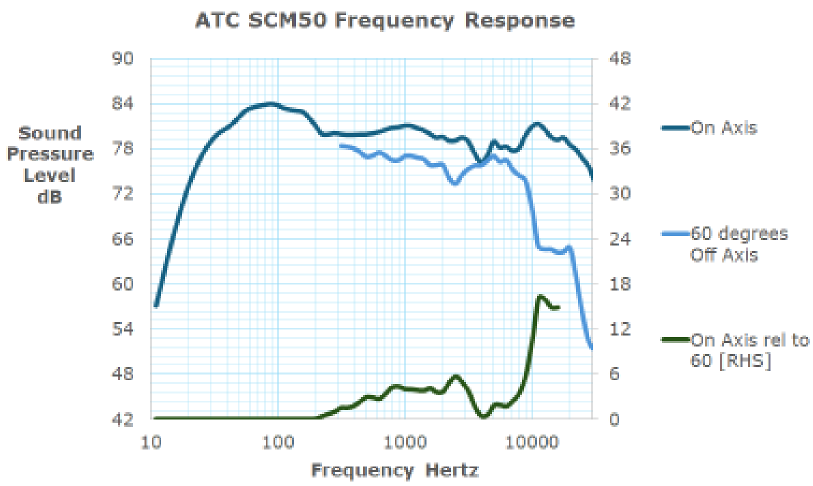
A.3 Loudspeaker Directivity for Different Configuration of the Driver

The graph shows the change in Directivity Index with frequency for the same driver in three different configurations. This shows that the driver used in a dipole configuration has a broader frequency range in which its directivity has minimal change. At low frequency its absolute DI = 3 giving a significant advantage over omni-directional bass (+4.8dB direct to reverberant improvement, as does cardioid bass).



A.4 ATC SCM50 Studio Loudspeaker Directivity

The frequency response curves for on and 60 degrees off axis of an ATC SCM50 Loudspeaker. The variation in the received spectrum in the angle range ± 60 degrees from on axis is within 6dB up to around 10kHz, that is, the on and off axis spectra are similar over a broad range of receiver angles up to high frequency – a reasonably constant directivity source.



A.5 Commercial Loudspeaker Example Directivity

The majority of commercial/professional audio products are omnidirectional in the bass but tend to be paired with midrange to high frequency elements that are significantly more directional. The loudspeaker frequency response, (0° to 90° in 7.5° degree steps) shown in the graph is an example. Although described as constant directivity it is far from what is required as it can be seen that the spectrum at a listener changes markedly with angle to the loudspeaker. Equalization to a flat response for all listeners or reflection surfaces within a ± 60° azimuth range is not possible.