DOES 1/3RD OCTAVE EQUALISATION IMPROVE THE SOUND IN A TYPICAL CINEMA?

Philip Newell	Acoustics Consultant, Moaña, Spain
Glenn Leembruggen	Acoustic Directions, ICE Design, Australia
Keith Holland	ISVR, University of Southampton, UK
Soledad Torres-Guijarro	University of Vigo, Spain
David Gilfillan	ICE Design, and Gilfillan SoundWork, Australia
Julius Newell	Electroacoustics Consultant, Blackburn, UK
David Santos Dominguez	University of Vigo, Spain
Sergio Castro	Reflexion Arts, Vigo, Spain

1 INTRODUCTION

This paper continues the investigation into the current poor state of sound in cinemas (1), (2), (3). Specifically, it investigates the viability of the specified Dolby equalisation of cinema sound systems and whether it enhances the aural experience. Much is still spoken and written about 'room equalisation', but, in reality, the concept is a myth. Rooms *cannot* be equalised. Sound waves expand three-dimensionally and interact with the boundaries of rooms in complex ways causing the frequency response at every point in any non-anechoic room to be different in both level and spectrum with a given source.

Using acoustic measurements conducted in a cinema style room using a loudspeaker, we demonstrate how attempts to equalise the response for a given position in a room will not necessarily produce improvements at the majority of other places within the room. Responses were measured with different time-window lengths to assess the changes in the received spectra over time.

Comparison is also made between two loudspeakers with different directivity characteristics, which show that the response at each location is highly dependent on the way in which the loudspeaker excites the room.

2 BACKGROUND

It has been long experienced by many professional sound engineers that attempting to equalise a system far back into a room using steady state measurements has resulted in poor and inconsistent results. As a result, it is often accepted in professional circles that an installed sound system should be frequency-corrected in the close field rather than the far field. Improved results have resulted using this method and industry practice has often followed this trend.

Many practitioners have therefore been dismissive of techniques used by the cinema industry over recent years, in which engineers and automated systems attempt to correct anomalies within auditoria by the comprehensive use of amplitude equalisation measured at a single position or a few positions. This approach of "one equalisation fits all seats" has never held much weight in live sound where debate has run for years about the benefits and pitfalls of mixing in the sweet spot. Live sound engineers frequently walk the auditorium during a show to ensure there are no gross spectral imbalances at positions away from the mix position.

2.1 Current Calibration Method

The current calibration method for Dolby certification of a cinema or dubbing suite involves the following process:

- play pink noise through each loudspeaker in turn
- measure response with 1/3rd octave analyser
- adjust 1/3rd octave equaliser until desired response is achieved
- the current target is the X-Curve is within ±3dB
- microphone position 2/3rd of distance from screen to back wall
- the microphone may be
  - multiplexed multiple spaced microphones
  - or a single microphone "waved" manually
  - or sometimes a single fixed microphone at ear height on room centreline

2.2 Problems with Current Method: Limited Frequency Resolution

The current Dolby specified method of 1/3rd octave analysis and equalisation is based on the understanding that human critical bands are approximately 1/3rd octave wide. However, the basis for this is how we perceive broadband noises.

However:

- human frequency resolution is much finer than 1/3rd octave
- loudspeaker and room response aberrations can be relatively narrow in frequency
- loudspeaker and room response aberrations usually don't fall neatly into the bands with fixed frequency centres
- 1/3rd octave filters with fixed centres are incapable of exactly matching almost all loudspeaker and room response aberrations

3 ACOUSTIC MEASUREMENTS

The following method was used:

a) A three-way loudspeaker was evaluated in the anechoic chamber of the university and its response noted. This represented a typical loudspeaker product that a cinema contractor would use.

b) The loudspeaker was set up where a centre-channel cinema loudspeaker speaker would be located in an auditorium with a reasonably good acoustics at Vigo University. Figure 1 shows the loudspeaker in situ.

c) The impulse response of the loudspeaker and room combination was recorded at with eight microphones at positions shown in Figure 2. The steady-state responses with pink noise were also examined in 1/3rd octave.

d) A basic attempt at equalisation was made using a 1/3 octave spectrum analyser and a 1/3 octave graphic equaliser to improve the response of the loudspeaker at a position approximately 10 m from the stage (1/3 of the distance towards the rear of the auditorium) and approximately 10 degrees off axis to the loudspeaker.

e) The eight impulse responses were re-measured again after this equalisation was applied.

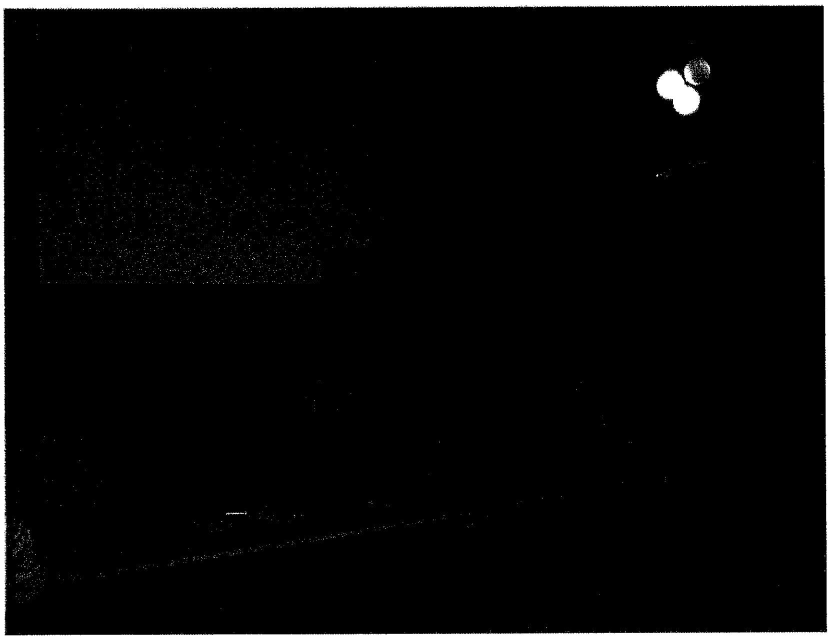

Figure 1 Loudspeaker in test auditorium as centre channel

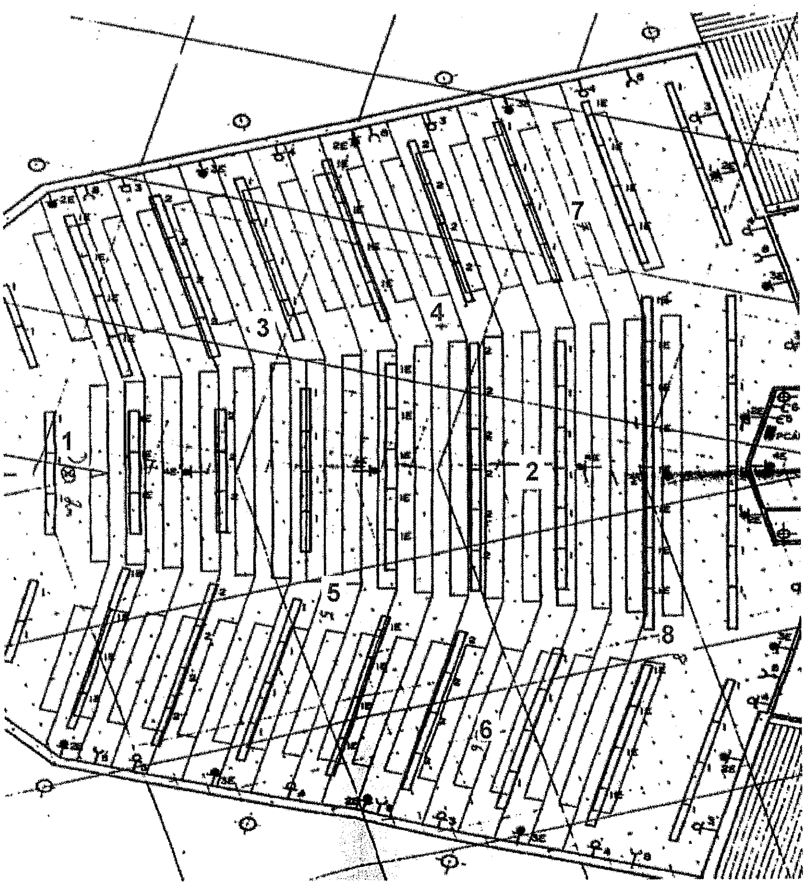

Figure 2 Seating positions in the room at Vigo University

3.1 Frequency Response Analysis

The frequency response of each measurement was computed from the impulse responses by the Fast Fourier Transform using a Tukey window of different lengths. The resulting frequency responses were then energy-averaged over a 1/15th octave bandwidth and the values assigned to the associated frequency at the centre of each bandwidth.

3.1.1 Time Windows

A Tukey window shape is also known as a "tapered cosine window" and can be regarded as a raised-cosine window which has been convolved with a rectangular window. An example of the half Tukey window is given in Figure 3. The flat top of the window allows equal weighting to all points within that section the impulse response (IR), while the half-cosine section reduces the leakage due to truncation of the data. The actual windows used consisted of a rectangular sections of length 10 ms, 50 ms, 80 ms and 400 ms followed by similar length half-cosine sections.

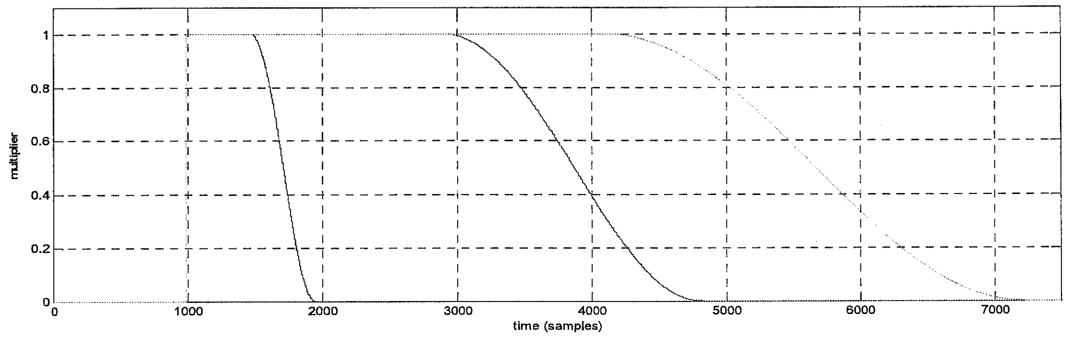

Figure 3 Example of half Tukey window

The rationale for the selected time window lengths is:

- The 10 ms window includes the loudspeaker's direct field at mid frequencies and above and represents the likely lower limit of the psycho-acoustic temporal integration time.
- The 50 ms and 80 ms windows are mirrors of the C_{50} and C_{80} acoustic metrics discussed below.
- The 400 ms window is a reasonable time to integrate the majority of the room's discrete reflections, and will include reflections that are not useful for clarity.
- All time data represents the steady-state condition, which would be measured with pink noise if sufficient measurements were made to average out the stochastic variations in the noise.

Each of these time windows makes a contribution to the subjectively perceived frequency response.

3.1.2 Parallels with Intelligibility Metrics

Measures of the ratio of early-arriving sound to late-arriving sound are used as reasonably reliable indicators of the ability of a sound/room system to deliver speech intelligibility. The C_{50} and C_{80} metrics are based on the principle that clarity is determined by the relative strengths of useful and detrimental sound energy. Useful sound is the combined energy of the direct and early-reflected sounds, while "detrimental" sounds are the combined energy of late reflected sound, reverberant sound and ambient noise. A duration of 50 ms for speech and 80 ms for music is generally used for the time period dividing these two types of sound field.

Both metrics are found by integrating appropriate portions of the room impulse response. It should also be recognised that the use of a sharp boundary division between early and late oversimplifies the situation.

The C_{50} is also loosely related to the direct to reverberant ratio (D/R) and includes the possible enhancement of speech sounds by strong early reflections.

4 RESULTS

4.1 Anechoic Response of Loudspeakers

Figure 4 shows the anechoic frequency responses of the loudspeakers A and B.

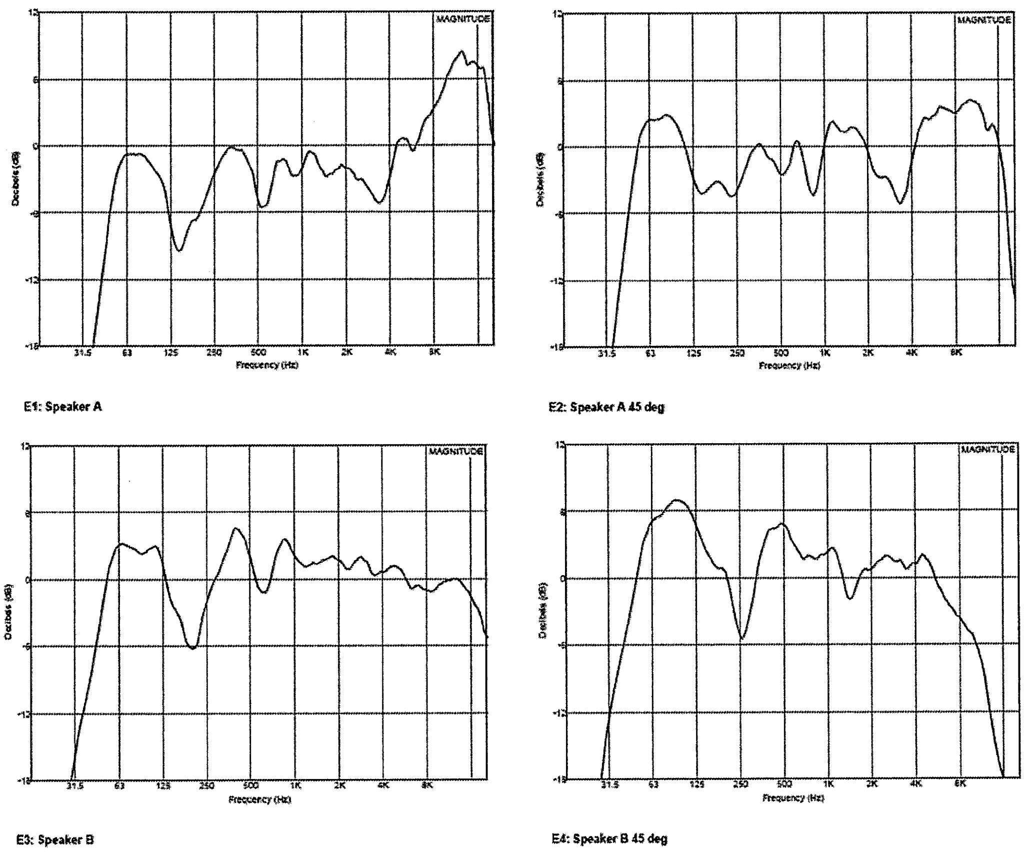

Figure 4 Anechoic frequency responses of the loudspeakers

4.2 Comparison of Third-Octave and High Resolution Responses

In Figure 5, the frequency response of the system at Position 7 measured by a 1/3rd octave real-time analyser with 800 ms wide Hanning window is compared with that computed from the impulse response with the 400 ms half-Tukey window.

Given the additional information that the response computed from the IR is so much more complete and the simplicity of modern IR analysers, it is hard to understand why cinema equalisation is specified to be measured in 1/3rd octave bands

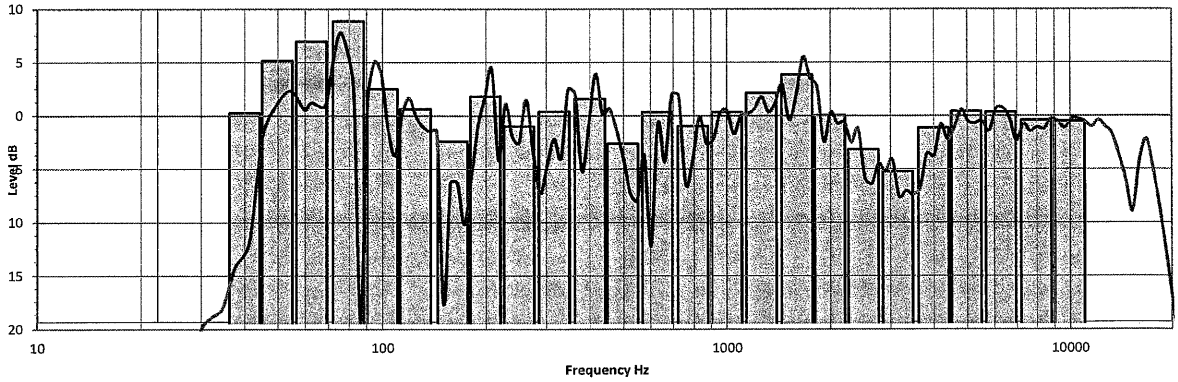

Figure 5 Comparison of frequency response measured by 1/3rd octave real time analyser with pink noise (800 ms Hanning window) with computed response from the impulse response using 400 ms half-Tukey window.

4.3 Responses at Each Position

Figure 6 shows a sample of the frequency responses at different locations with and without equalisations computed with different length windows. Although the response was equalised with a graphic equaliser at a specific position to be relatively flat as measured with 1/3rd octave steady-state spectrum analyser, Figure 6 shows that none of the responses is particularly flat.

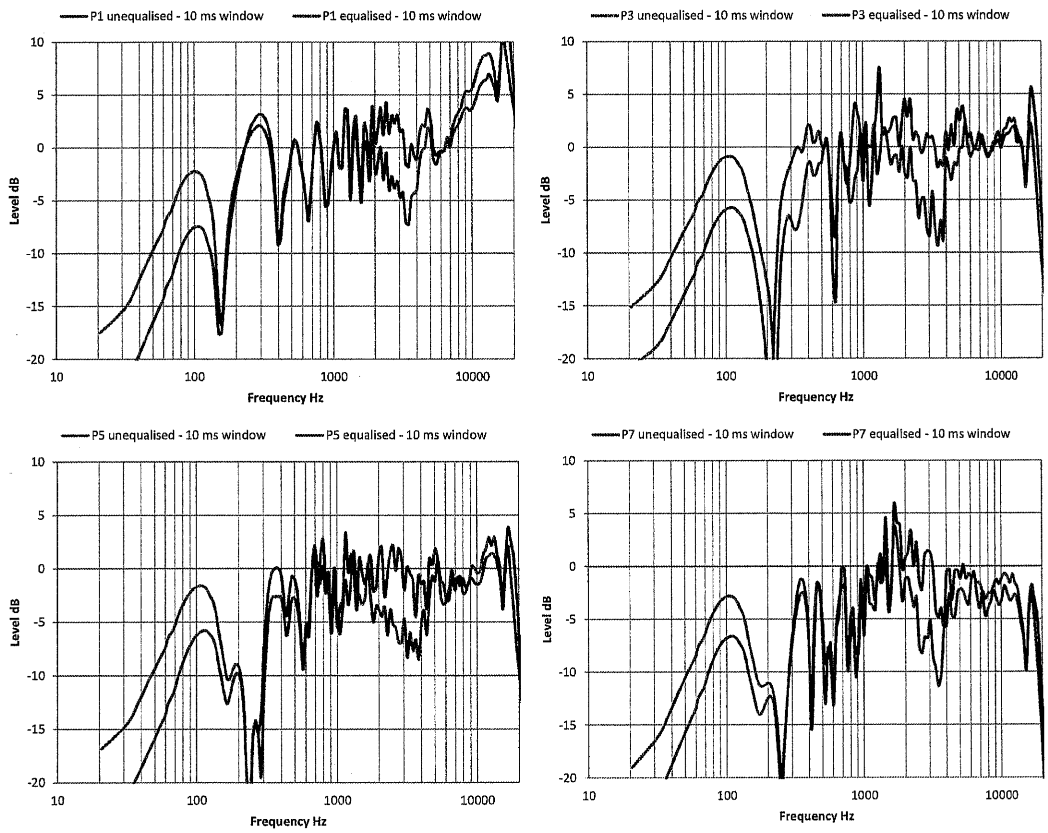

Figure 6 a Positions 1, 3, 5 and 7 with 10 ms window

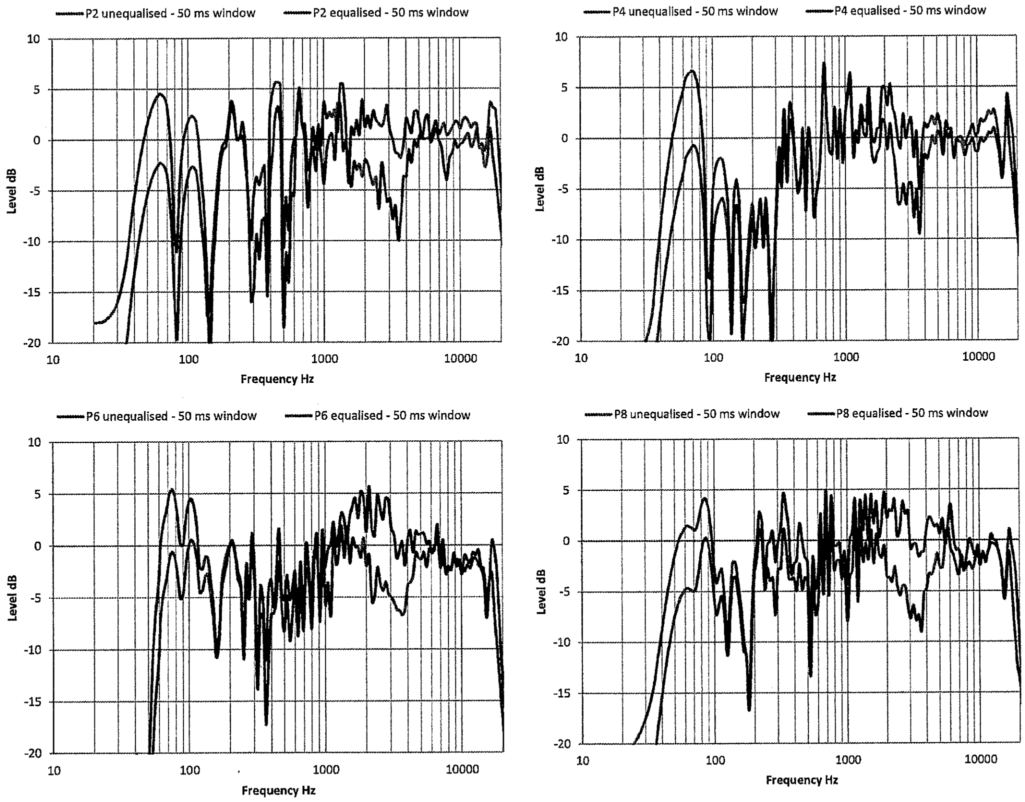

Figure 6b Positions 2, 4, 6 and 8 with 50 ms window

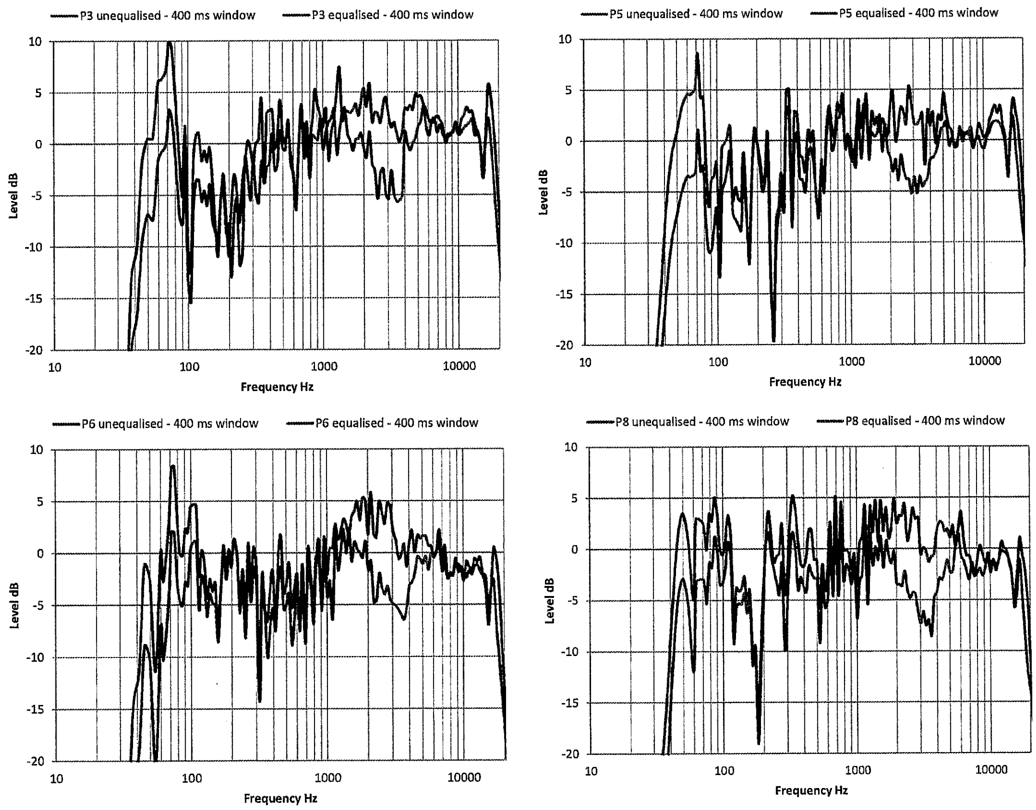

Figure 6c Positions 3, 5, 6 and 8 with 400 ms window

4.4 Consistency of Differences

A sound system is expected to be linear and time-invariant, with changes in the input signal producing corresponding changes in the output signal.
Figure 7 compares the smoothed differences at each position between unequalised and equalised for the four window lengths.

The only significant benefits that are apparent over the range of plots are the partial corrections of the inherent dip at around 3 kHz that was present in the anechoic near field measurement and the excessive energy below 100 Hz. However, these corrections were too coarse to properly compensate for these deficiencies.

Although the average trend of the equalisation is clearly present, there are narrow band variations above and below the overall trend. The reason for this is not understood properly, however it is expected to relate to time-variant effects in the room, and truncation by the half-Tukey window of inaudible but measurable low frequency noise resulting in the introduction of spurious spectral lines into the spectra.

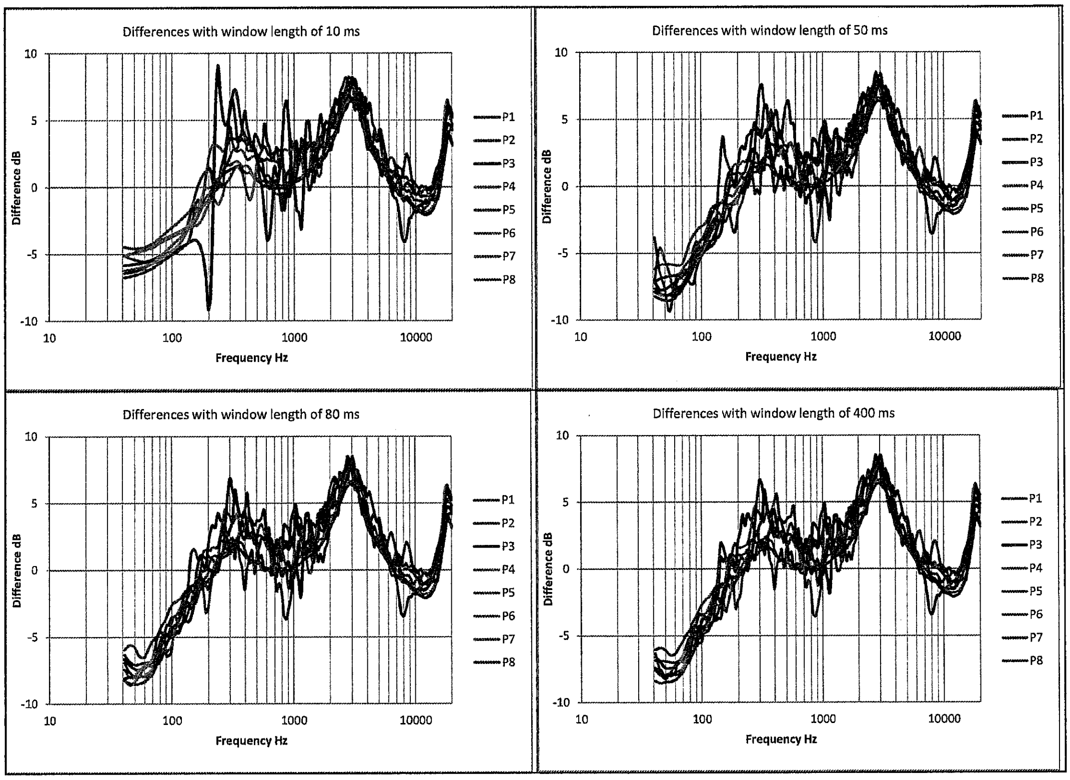

Figure 7 Differences between equalised and equalised responses at eight positions with the four window lengths.

4.5 Average Responses

The average effect of the equalisation was examined over the eight positions. Figure 8 shows the average over the eight positions of both the unequalised and equalised systems. It is clear that the equalisation undertaken at the single position has not produced a useful overall average response.

The poor overall frequency responses will also degrade dialogue intelligibility, due to the mechanism of psycho-acoustic upward masking (4), (5).

The effect of equalising the average response was then mathematically examined, by applying the predicted response of a series of parametric filters to the average response with each time window. Position 1 was excluded from the average, as its responses were sufficiently different from the other responses to possibly skew the result. Figure 9 shows the equalised average results for the four time windows, along with the response of the filters.

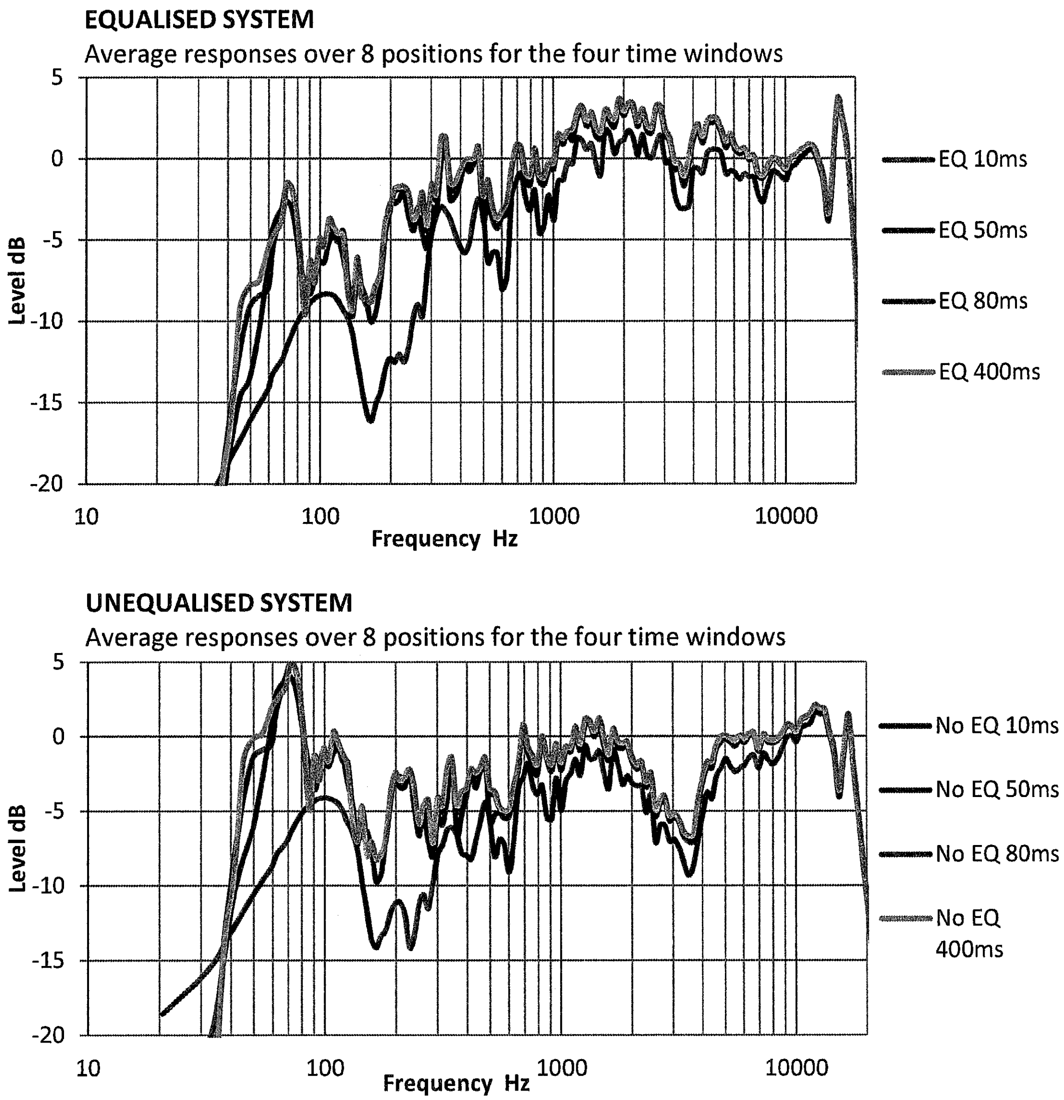

Figure 8 Unequalised and equalised responses averaged over the eight positions. Note that differences in overall level were not removed before computing the averages.

The following points are noted with respect to Figure 9:

- The responses of the 50, 80 and 400 ms windows are very similar.
- Compared to the other windows, the equalised response with the 10 ms window droops below 1 kHz, as substantially fewer reflections have arrived in this period. This is typical of many professional sound systems.
- The applied equalisation must only be considered as a starting point, as the 10 ms response has some validity to the subjectively perceived response, as it has considerable boost between 100 Hz and 200 Hz, this boost might cause problematic colouration.

Figure 10 compares the unequalised responses at all positions (80 ms window) with the responses when the simple parametric equalisation (shown in Figure 9) based on the average response is applied. It appears that worthwhile improvements result from the average equalisation process. For clarity, the responses have been smoothed over 1/3rd octave.

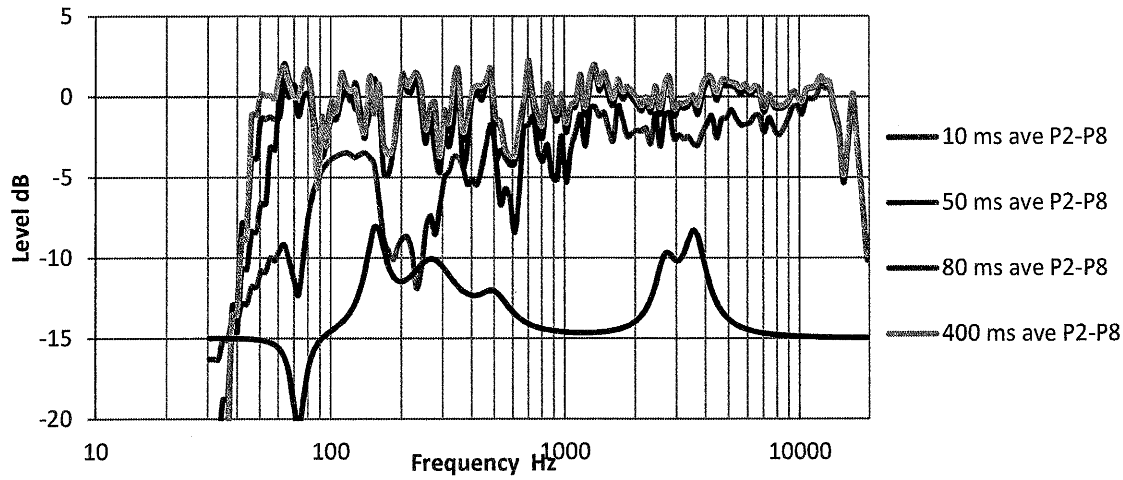

Figure 9 Effect of applying a set of parametric filters to each time-window average. The response of the filters is offset by 15 dB for clarity. Position 1 was excluded from average.

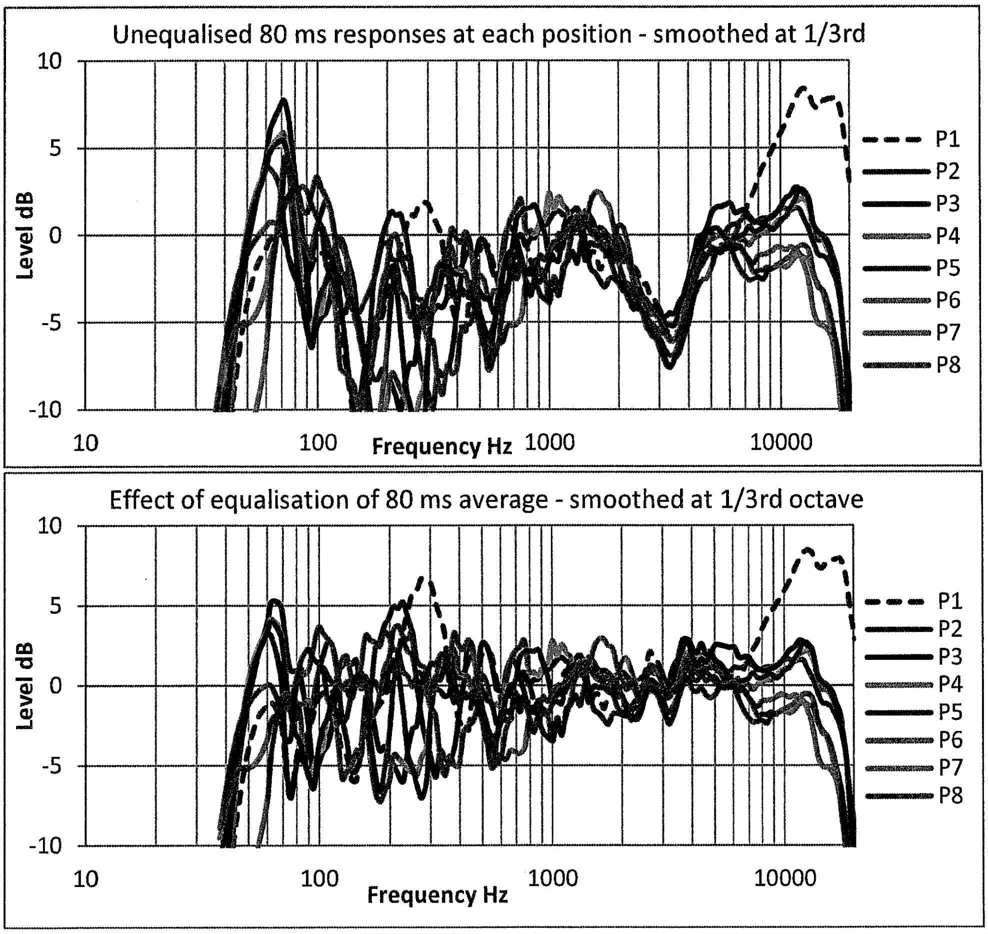

Figure 10 Effect of simple equalisation of average on each position (80 ms window, smoothed over 1/3rd octave)

4.6 Differences between A and B loudspeakers

The differences between the A and B loudspeakers were examined at each position by first normalising the response of each loudspeaker type to its response at Position 2. This position is in the centre of the room approximately 2/3rd towards the rear of the room, and therefore is representative of the Dolby calibration position. Normalisation is equivalent to perfect equalisation at a single location.

Figure 11 shows the difference between the 80 ms responses of the loudspeaker types at each position, after normalisation. For clarity, the responses were first smoothed over a 1/3rd octave before the difference was computed. Significant differences result between locations, which are directly attributable to differences in the directivity of the loudspeakers.

Figure 11 Differences between responses of Type A and B loudspeakers at the eight positions after normalisation to their responses at Position 2. (80 ms window)

5 DISCUSSION AND CONCLUSIONS

The results of this investigation illustrate issues concerning equalisation that many skilled audio professionals have been aware of for some time, but are not accounted for in the specified calibration process for cinema and dubbing suites.

a) Measurement of a system's frequency response with pink noise at a single calibration position tells us very little about the sound in the cinema. This process should be discontinued.

b) Compared to frequency response measurements made using the specified 1/3rd octave bandwidth spectrum analyser with pink noise, measurements computed from the acoustic impulse response with different length time-windows provide much greater insight into frequency and time domain behaviour.

c) Response calibrations based on 1/3rd octave bandwidth spectrum analyser using pink noise should be discontinued.

d) 1/3rd octave equalisers with fixed band centres lack the required precision for the process of frequency response correction. Their use should be discontinued.

e) While the current use of 3rd octave equalisation may be "better than nothing" in some circumstances, the practice is out-dated and the specified method of equalisation should be improved.

f) The results confirm that the coarseness of the Dolby specified measurement and equalisation process will allow two rooms which measure very similar to sound very different.

g) Steady state measurements made in the far-reverberant field lump together all the reflections, resonances, and direct sound. The ears can discriminate between all of these things, but this type of measurement cannot!

h) Equalising the average response over a number of widely spaced positions will yield substantially better subjective results compared to using only one position. However, this requires measurements derived from the impulse response. Care and skill and substantial critical listening must accompany this process to confirm that each equalisation filter produces an aural improvement.

   The more consistent that i) the loudspeaker's direct-field is over the audience and ii) the loudspeaker's power response is with frequency, the greater the benefits of this type of averaging.

i) Poor frequency response will degrade dialogue intelligibility, especially for listeners with non-native accents and during periods of Foley effects and background music.

j) The trend towards reduced cinema sizes and lower reverberation times allow more detail in the sound and render more obvious the effects of inappropriate equalisation.

k) Unless the loudspeaker's power response and directivity is exceptionally consistent with frequency, and the loudspeaker is located well away from surfaces that would create boundary-type image sources, it is likely that a single representative point cannot be used to formulate the equalisation of a system. Skill would be required to both recognise the presence of a response cancellation due to a floor reflection, and ignore it.

l) If response correction is to be *reliably* applied without listening by skilled practitioners, it should be done in the close field to the loudspeakers, otherwise it becomes significantly convolved with unequalisable, non-minimum-phase characteristics of the room acoustics, and 'correction' then becomes an inappropriate word to use for the process.

m) "Spatial averaging" by waving the microphone over a limited area cannot yield the required results as:

- The impulse response cannot be achieved as the system in time-variant.
- Temporal discrimination is not achieved.
- Errors due to cancelations resulting from floor reflections will be obscured

n) The specified 1/3rd octave equalisation process at the calibration position yields will most yield the following outcomes:

- poor room-to-room compatibility, especially over the range of listening positions
- poor dialogue intelligibility at many listening positions
- a harsh and tiring soundtrack

Equalising in the close field will improve these parameters as the direct field is optimised, and this directly equates to perception.

o) Toole gives a simple but useful treatise on the pitfalls of 1/3rd octave equalisation in (6)

p) We believe there is difficulty sourcing people of sufficient skill and understanding that are able to apply subtle adjustments to cinemas around the world.

6 REFERENCES

1. **Newell P, Holland K, et al.** New Proposals for the Calibration of Sound in Cinema Rooms. *130th Convention of the AES London.* 2011.

2. **Leembruggen G, Newell P, Newell J, Gilfillan D, Holland K, McCarty B.** Is the X Curve Damaging Our Enjoyment of Cinema? *SMPTE Convention.* 2011 Sydney.

3. **Newell, P; Holland, K; Torres-Guijarro, S; Castro, S; Valdigem, E;.** Cinema Sound: A New Look at Old Concepts. *Proc. of the Institute of Acoustics.* Nov 2010, Vols. 32, Part 5.

4. **Leembruggen, G, Stacy A.** Should the Matrix be Reloaded? *Proceedings of the IOA.* 2003.

5. **Leembruggen, G, Hippler, M, Mapp. P.** Further Investigations into Improving STI's Recognition of the Effects of Poor Frequency Response on Subjective Intelligibility. *Convention Paper 128th AES Convention London, UK.* 2010 May.

6. **Toole, Dr F.** *The Acoustical Design of Home Theaters.* s.l. : http://www.harmanaudio.com/all_about_audio/acoustical_design.pdf.