# 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 |
| Julius Newell | Electroacoustics Consultant, Blackburn, UK |
| Soledad Torres Guijarro | University of Vigo, Spain |
| David Gilfillan | ICE Design, and Gilfillan SoundWork, Australia |
| 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) and the effects of specified calibration processes. We investigate the viability of the specified Dolby equalisation of cinema sound systems and whether it enhances the aural experience of listeners. 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 with a single, centre-front 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. Therefore, 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 encouraged the use of 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 theatre 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 to be within ±3dB of the X-Curve
- microphone position to be approximately 2/3rd of distance from screen to back wall
- the microphone(s) may be
  - multiplexed, multiple, spaced microphones
  - a single microphone "waved" manually
  - 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 critical bands in human hearing are approximately one-third octave wide. However, the basis for this is how we perceive broadband noises. In reality:

- 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 almost never able to exactly match loudspeaker and room response aberrations

# 3 ACOUSTIC MEASUREMENTS

Acoustic measurements were made using the following method:

a) The frequency response of a three-way loudspeaker was measured in a hemi-anechoic chamber. This represented a typical loudspeaker product that a cinema contractor would use. As the loudspeaker was lying on its back on the hard surface, the measurement included the effect of the image source behind the loudspeaker.

b) The loudspeaker was then set up close to where a centre-channel cinema loudspeaker would be located, being in an auditorium with reasonably good acoustics at Vigo University which is sometimes used as a cinema. Figure 1 shows the loudspeaker in situ.

c) The impulse response of the loudspeaker and room combination was recorded with a microphone located at eight positions shown in Figure 2. The steady-state responses with pink noise were also recorded in one-third octaves.

d) A basic attempt at equalisation was then made using a 1/3rd octave spectrum analyser and a one-third octave graphic equaliser to "improve" the response of the loudspeaker at a single position. That position was approximately 10 m from the stage, equating to 2/3 of the distance towards the rear of the auditorium, and approximately 10 degrees off axis to the loudspeaker.

e) The impulse responses at the eight positions were re-measured after this equalisation was applied. For these equalised measurements, the measurement microphone was carefully placed in the same positions as for the pre-equalisation measurements.


Figure 1  Loudspeaker in test auditorium as centre channel

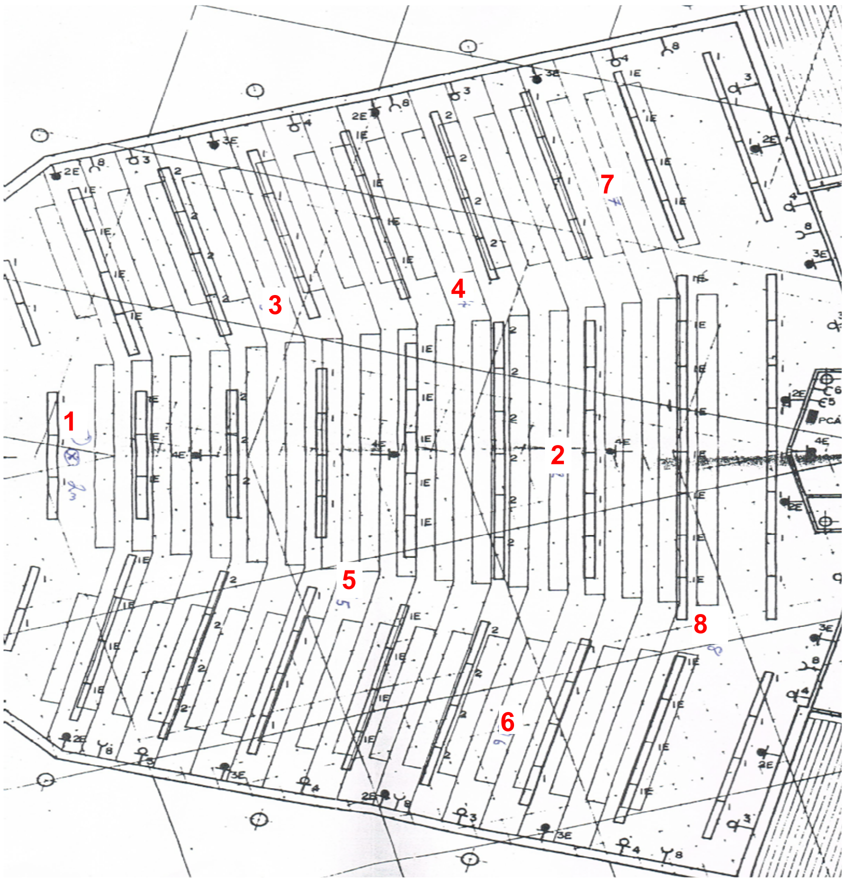


Figure 2  Microphone positions in the room at Vigo University.  The loudspeaker is forward and left of Position 1.

## 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 different window lengths are a simple attempt to consider the range of integration times that the human hearing process uses to assess frequency content. The resulting frequency responses were then energy-averaged over a 1/15$^{th}$ 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 of the impulse response (IR), while the half-cosine section reduces the leakage due to truncation of the data. The actual windows used consisted of rectangular sections of length 10 ms, 50 ms, 80 ms and 400 ms, followed by similar length half-cosine sections.

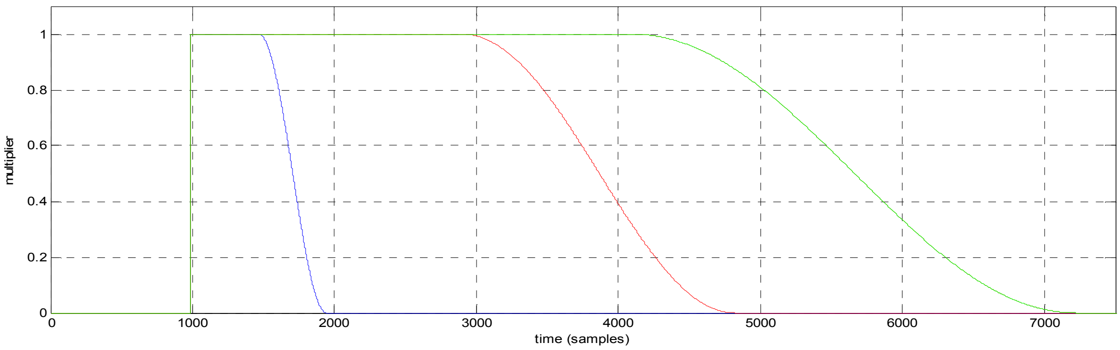


Figure 3 Example of half Tukey window

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 also 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 provides information relating to the frequency response that is subjectively perceived.

### 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 Hemi-Anechoic Response of Loudspeakers

Figure 4 shows the frequency responses of the loudspeakers A and B measured in the hemi-anechoic chamber. The dips in the frequency response in the range between 100 Hz and 300 Hz are mostly due to the presence of the image-sources behind each loudspeaker, and are a consequence of each loudspeaker being laid on its back. The high-frequency peak in the on-axis response of Loudspeaker A was only evident at angles very close to the axis.

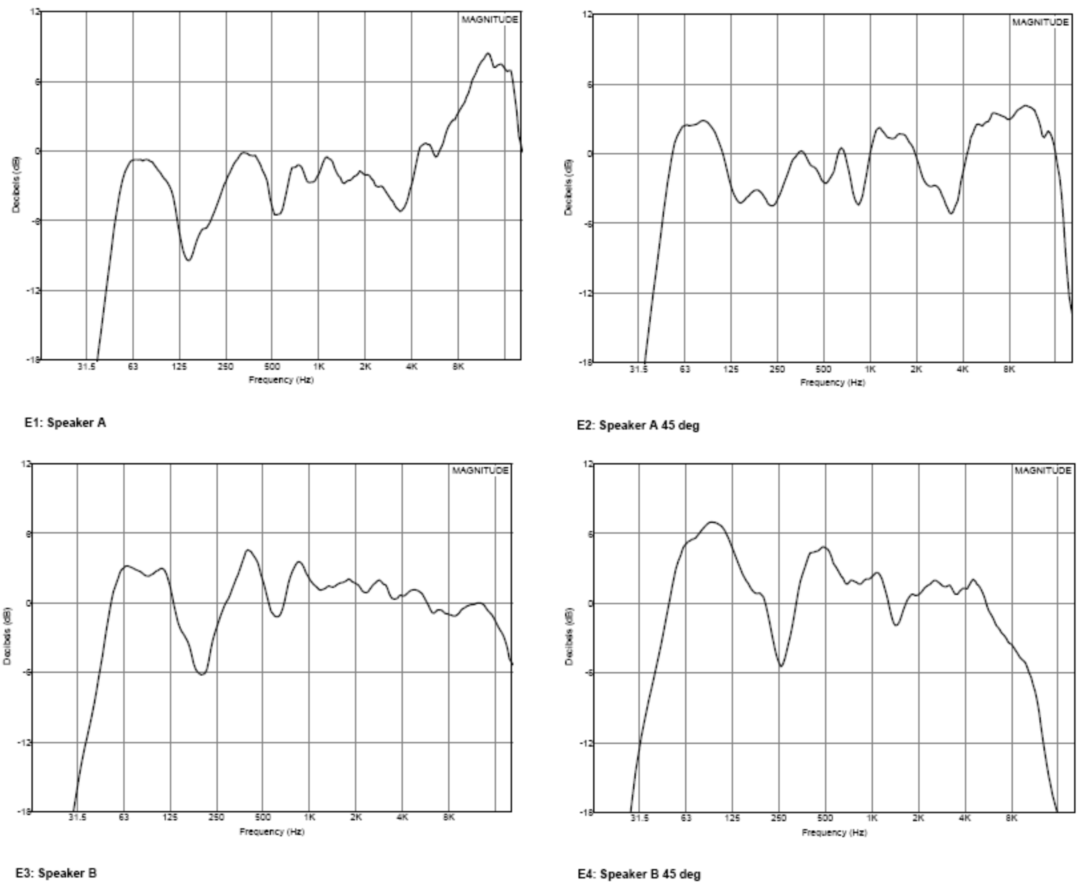


Figure 4 Frequency responses of the loudspeakers when laid on the floor of a hemi-anechoic chamber.

## 4.2 Comparison of Third-Octave and High Resolution Responses

Figure 5 compares the steady state frequency response of loudspeaker A at Position 7, measured with pink noise and a 800 ms Hanning window, and integrated into one-third octave bands with the response computed from the IR and the 400 ms half-Tukey window. As the IR with a 400 ms window essentially represents the system's steady state response, the expected broad agreement is present between both these measurement techniques.

Given that i) the response information computed from the IR is so much more complete, ii) the simplicity of modern IR analysers, and iii) the stochastic variations in pink noise result in level variations at low frequencies, it is hard to understand why measurements in one-third octave bands are still specified for cinema equalisation.

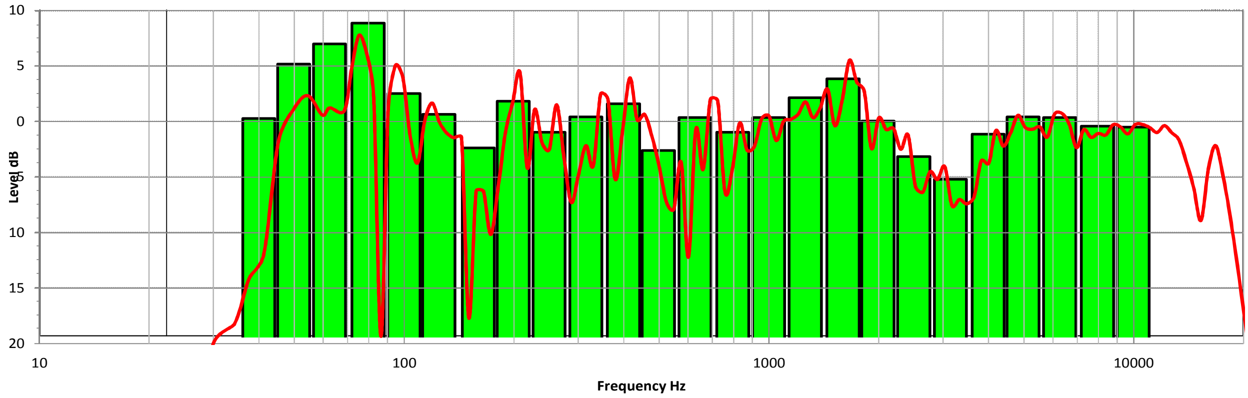


Figure 5 Comparison of frequency response measured with pink noise in 1/3$^{rd}$ octave bands (800 ms Hanning window) with computed response from the impulse response with 400 ms half-Tukey window.

## 4.3 Responses at Each Position

Figure 6 shows a sample of the frequency responses at different locations with and without equalisation, 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 the one-third octave steady-state spectrum analyser, Figure 6 shows that none of the responses is particularly flat.

It can be seen that the responses change over time, as the sound field builds up, due to reflections "filling in" the gaps resulting from floor or other early reflections and the arrival of all components of the impulse response that have been delayed by the system's group delay response.

When steady state responses are measured deep into the room, the variations in the steady-state low-frequency energy over the room due to boundary addition, modal effects and reverberation will mean there can be no reference position for measurement of the low frequency response. Accordingly, if the low-frequency balance and shape are equalised on the basis of these measurements, the direct sound in the room at low-frequencies is likely to be significantly degraded by that equalisation. This would lead to every room having a different direct sound, which does not bode well for uniformity.

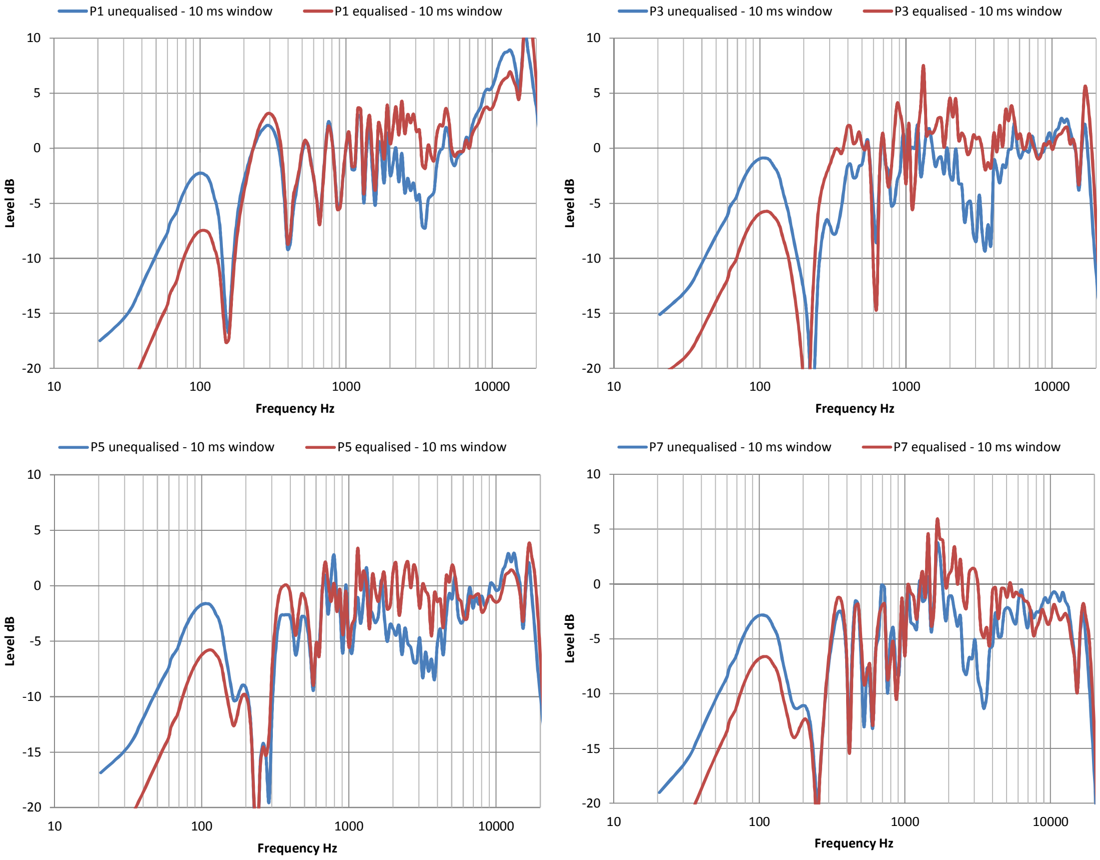


Figure 6 a  Frequency responses at Positions 1, 3, 5 and 7 with 10 ms window

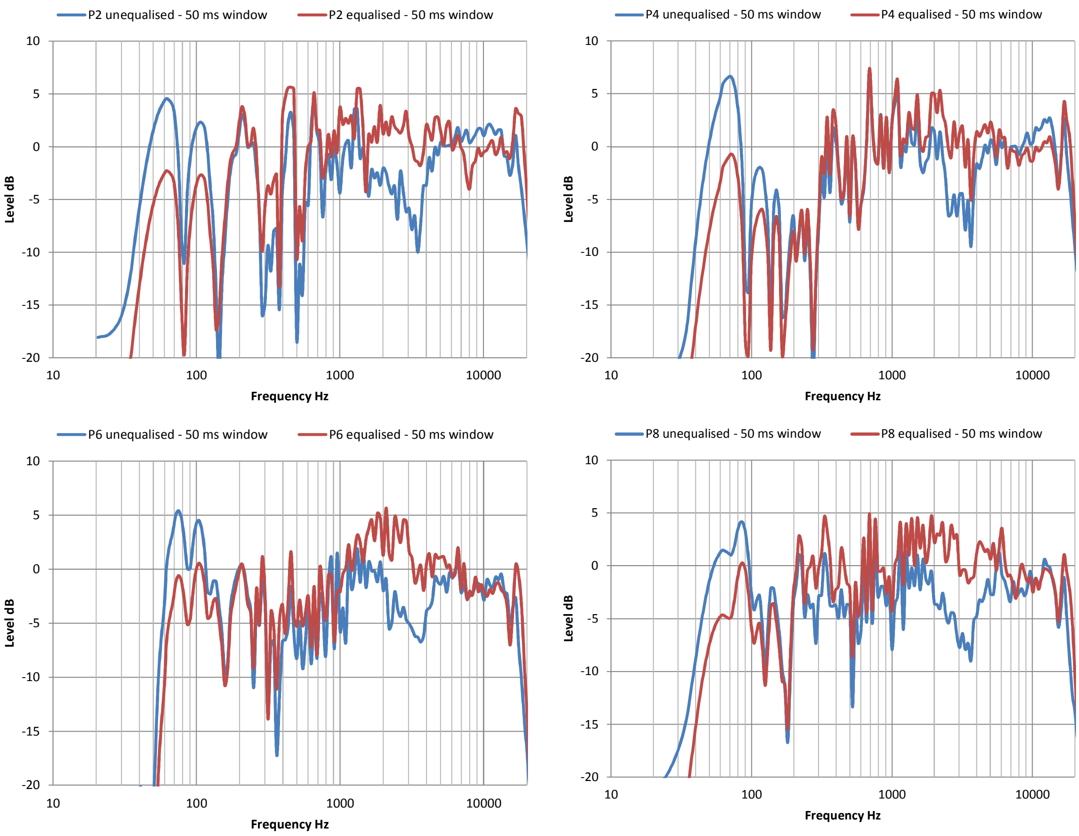


Figure 6b  Frequency responses at Positions 2, 4, 6 and 8 with 50 ms window

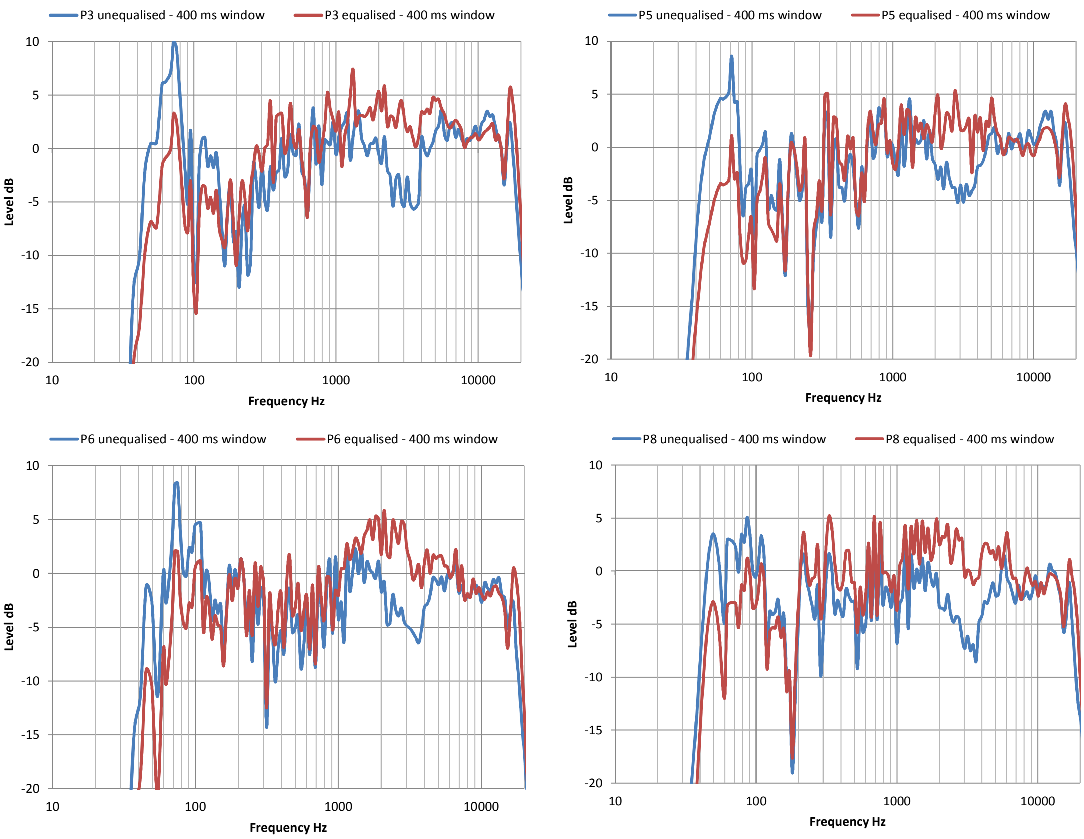


Figure 6c  Frequency responses at Positions 3, 5, 6 and 8 with 400 ms window

## 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 with 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.  These are expected to have primarily resulted from small differences between the microphone positions used for the un-equalised and equalised measurement sessions.  (An error necessitated these separate measurement sessions).  As the measurements were made in the seating area, reflections from near and far surfaces have produced comb-filtering in the frequency responses, and therefore small changes in microphone location have produced significant narrow-band changes in response.  These differences are particularly evident in some of the 10 ms responses.

Given that care was taken in the repositioning of the microphone at each location, the extent of the differences shows the i) fragility of this type of measurement when different people may be re-calibrating the same room from time to time and ii) the need for skilled interpretation of the measured responses.

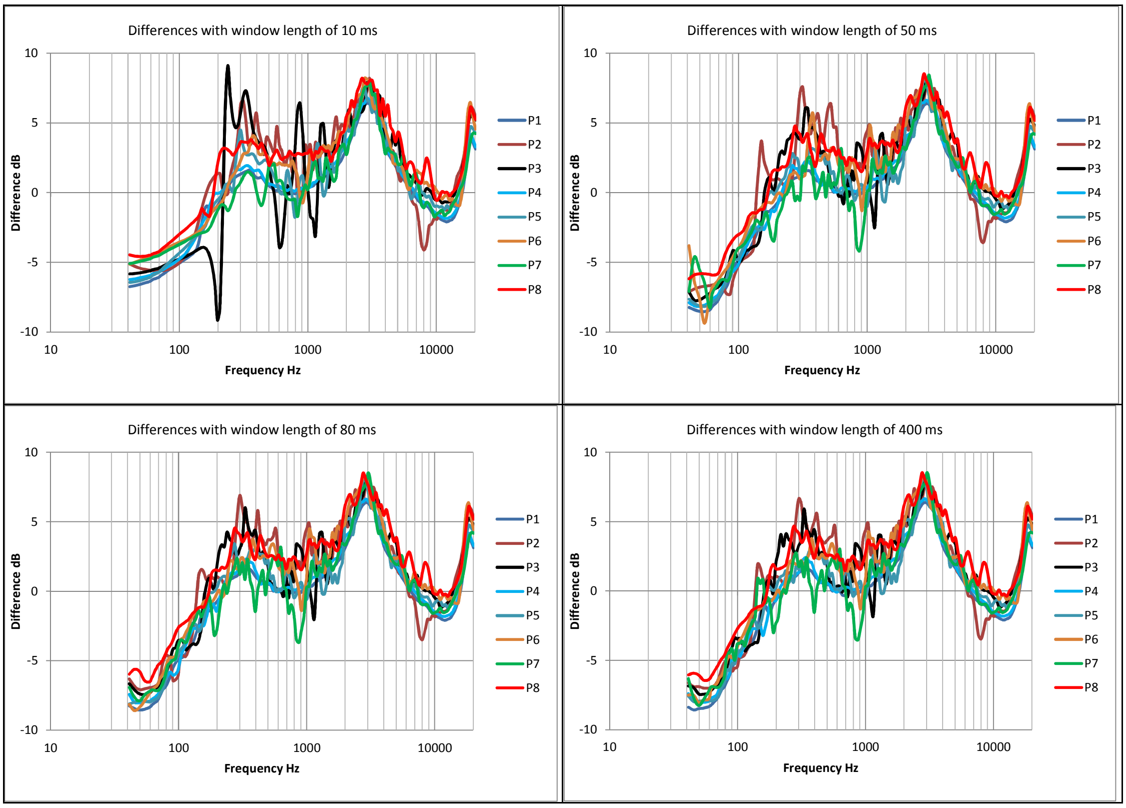


Figure 7 Differences between unequalised and equalised responses at eight positions with the four window lengths.

## 4.5 Average Responses

To examine the overall effect of the equalisation, the mathematical average of the responses at the eight positions was found. Figure 8 compares the response of both the unequalised and equalised systems for the four time windows after averaging over the eight positions. 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).

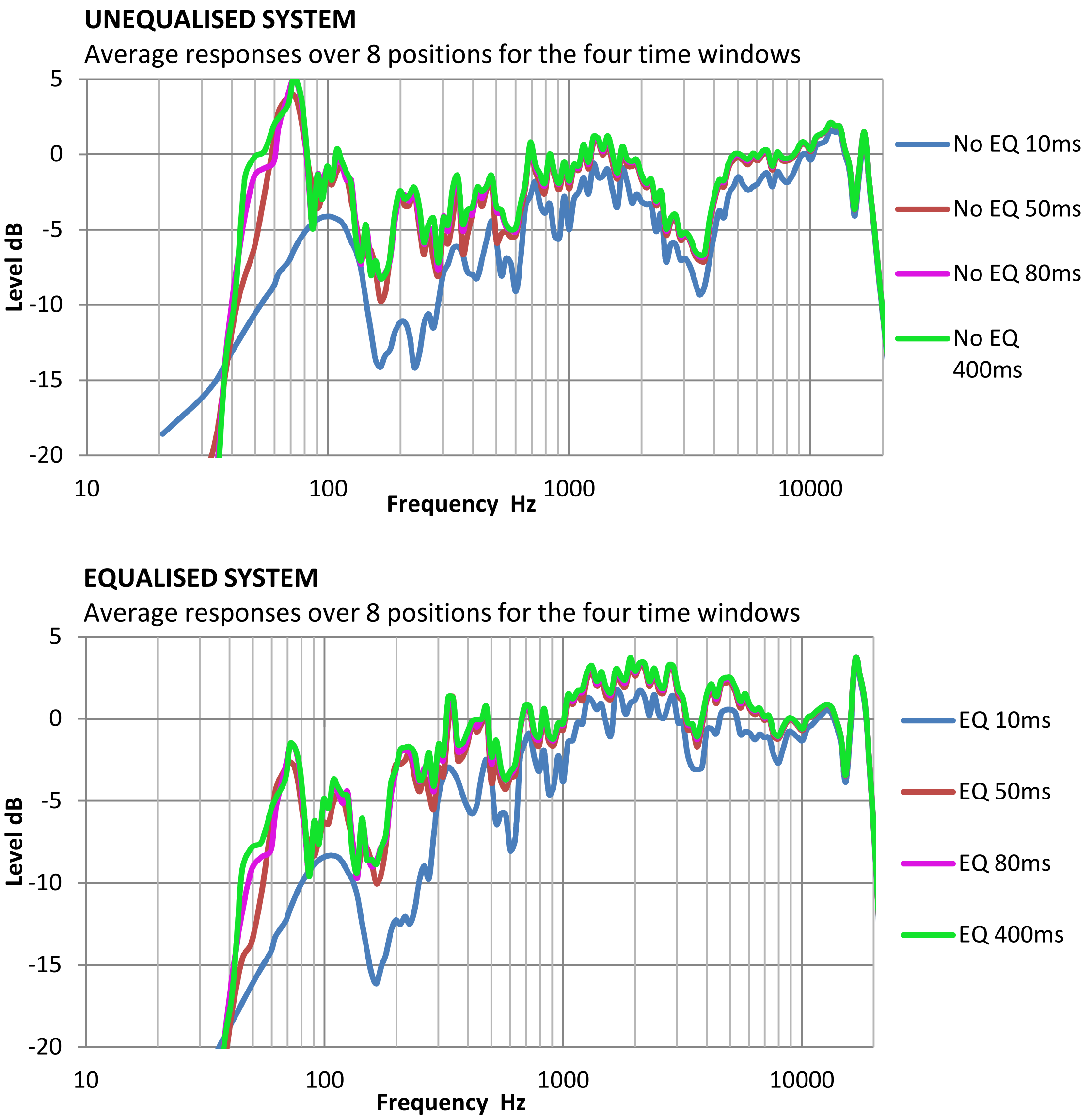


Figure 8 Unequalised and equalised responses averaged over the eight positions. Note that differences in overall level were not removed before computing the averages.

## 4.6 Equalisation of the Average Response

The effect of equalising the average unequalised response (over the eight positions) was then examined. Averaging was based on sound pressure levels in dB, rather than sound intensities (pressure squared), as this gives equal weight to all positions and better reflects the subjective differences between responses. The frequency response of a set of parametric filters was then computed and mathematically applied to the average response with each time window. Position 1 was excluded from the average, as its responses were significantly different from the other responses and would inappropriately skew the average.

Figure 9 shows the mathematically-equalised average response for the four time windows, along with the response of the filter set.

The following points are noted:

- 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, but does not necessarily reflect a problem, as i) the effect of cancellations due to floor reflections is more evident in the 10 ms, ii) group delay at low frequencies sometimes means that the impulse response of the system is still decaying at low frequencies at the truncation point of the window.
- The applied equalisation must only be considered as a starting point, and the effect of each filter must be aurally checked, particularly the filters that boost the response. For example, the 10 ms response does have some importance to subjective perception and as it shows considerable boost between 100 Hz and 200 Hz, this boost might cause problematic colouration.
- The equalisation process must also consider the cause of response troughs and peaks, so that it does not attempt to correct for response dips such as floor reflections which are spatially-variable.
- The optimum window length may depend on the reverberation time of the room; however experienced judgement is required for its selection. Toole (6) recommends that the anechoic responses of the loudspeakers should be used as a basis for the room measurements.

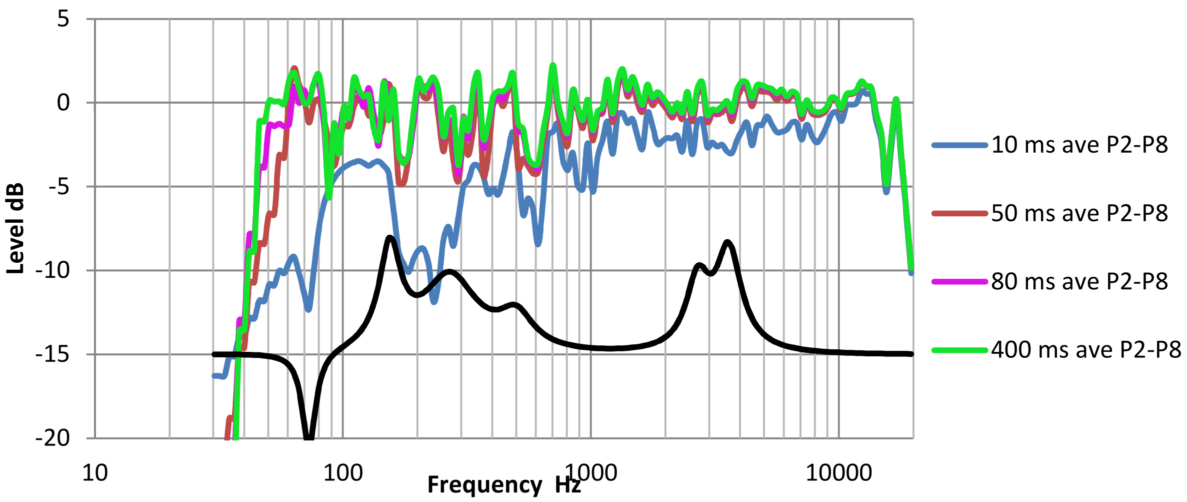


Figure 9 Effect of applying a set of parametric filters to each time-window average. The response of the filters is offset by 15 dB for clarity. Position 1 was excluded from average.

Figure 10 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. For clarity, the responses have been smoothed over one third octave.

Worthwhile overall improvements in the frequency responses have resulted from the process of equalising the average response. It is also noted that the average 10 ms response is substantially flatter with the averaging process than either of the 10 ms responses in Figure 8. The averaging process discounts local variations in responses, and the resulting equalisation is much more likely to flatten the frequency response of the direct sound.

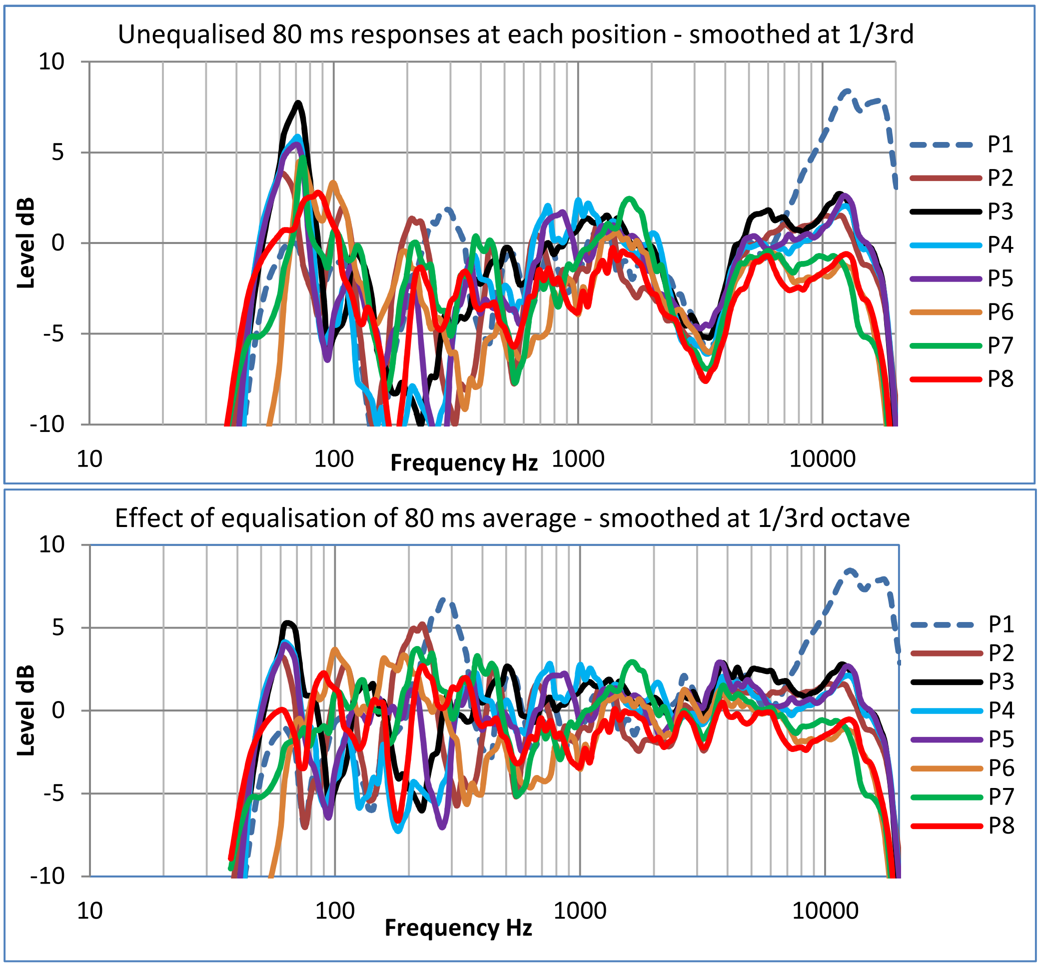


Figure 10 Effect of equalisation of the average response at each position (80 ms window, smoothed over 1/3$^{rd}$ octave)

Ideally, the averaging process would be based on the direct field, but it is not possible to measure the true direct field in a listening area as the phase interference of very early reflections causes strong significant changes in response. As inclusion of some reflections is unavoidable in almost all in-situ measurements, the choice of a suitable window length and shape is an important factor in the equalisation process.

A suitably short time window allows the direct field to be a dominant part of the average whilst reducing the effect on the measurement of these very-early reflections. In contrast, equalisation based on one-third octave integrated bands of pink noise and third-octave filters with measurements made a long way from the source cannot reliably address the direct field performance of the sound system.

From a field perspective, the authors have consistently found that the combination of the averaging process and critical listening always improves intelligibility, listener comfort and music enjoyment.

## 4.7 Differences between A and B loudspeakers

The differences between the A and B loudspeakers were examined at each position by first normalising the 80 ms response of each loudspeaker type to its response at Position 2. This position is in the centre of the room approximately 2/3$^{rd}$ 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. It is noted however, that the range of integration times that our hearing process uses with speech and music suggests that other window lengths may be just as appropriate as the 80 ms window.

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 one-third octave before the difference was computed. Significant differences result between locations, which are directly attributable to differences in the directivity of the loudspeakers and local phase interference effects.

The differences in the frequency responses between positions with the A and B loudspeakers resulting from this single location equalisation are clearly unacceptable.

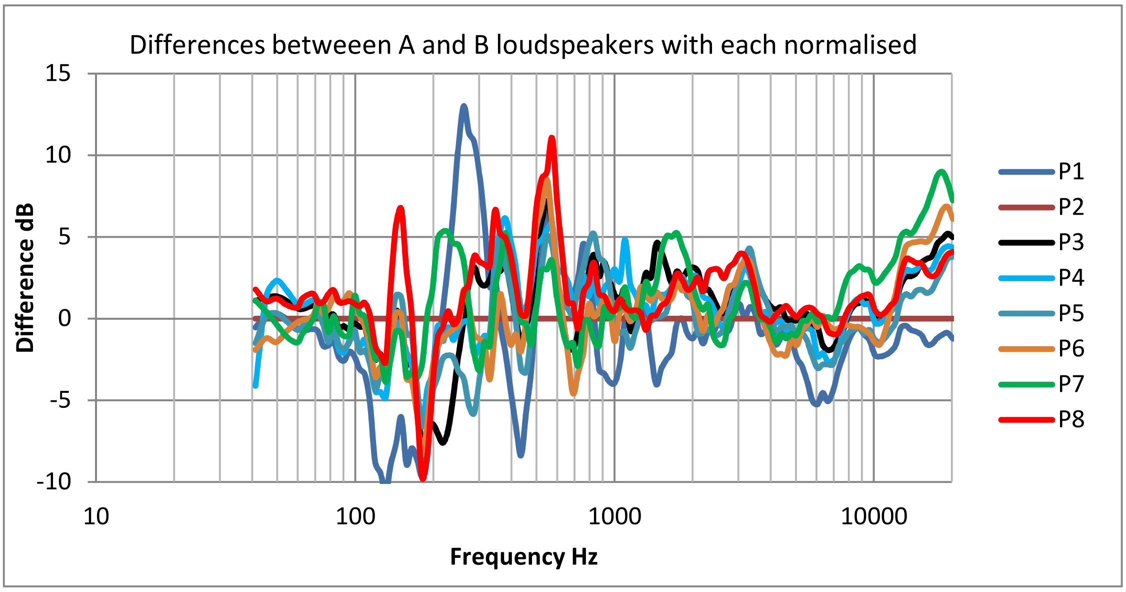


Figure 11 Differences between responses of Type A and B loudspeakers at the eight positions after normalisation to their respective responses at Position 2. (80 ms window)

# 5 DISCUSSION AND CONCLUSIONS

The results of this investigation illustrate various 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. Such issues include:

a) Measurement of a system's frequency response with pink noise at a single calibration position, or averaged over a small area, tells us very little about the general frequency response over the cinema. This process should be discontinued.

b) Compared to frequency response measurements made using the specified 1/3$^{rd}$ 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 one-third octave bandwidth spectrum analysis using pink noise should be discontinued.

d) One-third 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 one-third octave equalisation may be "better than nothing" in a few 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) The trend towards reduced cinema sizes and lower reverberation times allows more detail in the sound and renders the effects of inappropriate equalisation more obvious.

i) Poor frequency response, especially in the direct sound, will usually degrade dialogue intelligibility, particularly for listeners who are not familiar with the accents of the actors, or who have significant hearing loss. Such loss can cause problems with dialogue intelligibility during scenes with high levels of sound effects or background music.

j) Equalising the average response over a number of widely spaced positions will generally yield substantially better subjective results compared to using only one position. However, this requires measurements derived from the impulse response. Care, 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.

k) Compared to measurements made with one-third octave bandwidth, measurements with narrower bandwidths provide much more information about problematic peaks and troughs in the response. However, the use of narrow bandwidths introduces potential difficulties for future repeatability of calibration measurements in the far reverberant field, due to variability in the responses caused by small differences in the measurement positions. Those small positional differences can produce significant differences in response, due to different reflection patterns from surfaces near the microphone. Some skill would be necessary in the calibration process to discern the cause and importance of such differences.

l) It is our experience that optimum sounding equalisations are always based on information that is present in time durations ranging from the direct field (10 ms) through to steady state, and in this context, a flat response in any of these measurements may not represent the ideal subjective sound.

m) 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 neither a single representative point nor a measurement made over a small area can be used to formulate the optimal equalisation of a system. Skill would be required to both recognise the presence of a response cancellation due to a floor reflection, for example, and ignore it.

n) Items j), k), l) and m) above are not conducive to automatic equalisation processes, and therefore if response correction is to be reliably applied without listening by skilled practitioners, it should be done in the close field to the loudspeakers. If not done in the close field, the response becomes significantly convolved with un-equalisable, non-minimum-phase characteristics of the room acoustics, and 'correction' then becomes an inappropriate word to use for the process.

o) "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 is time-variant.
   - Temporal discrimination is not achieved, which is important to make judgements about response issues affecting intelligibility.
   - Errors due to cancelations resulting from floor reflections may be obscured.

p) The specified one-third octave equalisation process at the calibration position 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 has a strong bearing on perceived tonality.

q) Toole (7) gives a simple but useful treatise on the pitfalls of one-third octave equalisation. In (8), Linkwitz notes some pitfalls of equalisation for flat-steady state response, while Griesinger (9) re-iterates the importance of direct field for listener engagement.

r) We believe there is difficulty sourcing people of sufficient skill and understanding that are able to apply suitable adjustments to cinemas around the world. In this context, if a simple, reliable alignment standard the goal, equalisation at one or more reference positions in each loudspeaker's close field would provide substantially more accuracy and robustness against variation than the current method.

# 6 REFERENCES

1. **Newell P, Holland K, Newell J, Neskov, B.** New Proposals for the Calibration of Sound in Cinema Rooms. *130th Convention of the Audio Engineering Society, 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.* Reproduced Sound Conference Nov. 2010, Vols. 32, Part 5.

4. **Leembruggen, G, Stacy A.** Should the Matrix be Reloaded? *Proc. of the Institute of Acoustics.* Reproduced Sound Conference Nov. 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. *128th Convention of the Audio Engineering Society, London.* May 2010.

6. **Toole, F.** Private communication to the authors.

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

8. **Linkwitz, S.** The Challenge to Find the Optimum Radiation Pattern and Placement of Stereo Loudspeakers in a Room for the Creation of Phantom Sources and Simultaneous Masking of Real Sources. *127th Convention of the Audio Engineering Society, New York.* Oct. 2009.

9. **Griesinger, D.** Clarity, Cocktails and Concerts: Listening in Concert Halls. *Acoustics Today ASA.* Jan 2011, Vol. 7, No 1.