DISTRUBUTED LOUDSPEAKER ARRAY MEASUREMENT AND CORRECTION TECHNIQUES

Julius Newell, Electroacoustics consultant, Lisbon, Portugal
Keith Holland, ISVR, University of Southampton, UK
Philip Newell, Acoustic Consultant, Moaña, Spain

1 INTRODUCTION

It has been common practice in cinema calibration (and some other areas of audio) to use a distributed-source loudspeaker-array for the ambient sound-field reproduction. As a part of the *cinema* loudspeaker-alignment process, systems engineers have been required to analyse and equalise these arrays to a recommended target response by means of the averaging of between four and eight microphones distributed within the calibration area.[1, 2] Ostensibly, this procedure would make the array subjectively match the sound of an individual screen-channel that had been calibrated to the same response, but the evidence that this would actually be the case was, at best, tenuous. Other experts had argued that more-compatible results could be achieved if the individual loudspeakers in the arrays were timbrally matched to the screen channel. This paper will present evidence of the finer details and pitfalls encountered when attempting to analyse a complex sound-field created by what is often more than five individual distributed sources, and whether such analysis for the purpose of system correction is even possible.

The system under test was a working, 'Dolby Features'-certified dubbing theatre, with a low decay-time and a high degree of acoustic control. While not exactly anechoic, the room was defined as 'non-environment', with excellent acoustic control, even at very low frequencies. A fundamental reason for having such a low decay-time is so that the natural acoustics of the room will never dominate the intended ambience of the soundtrack, such as in a scene in an open field, for example. The decay time (RT_{60}) in the test room was nominally around 120 ms, as shown in Figure 1 and verified by the waterfall decay-plot shown in Figure 2.

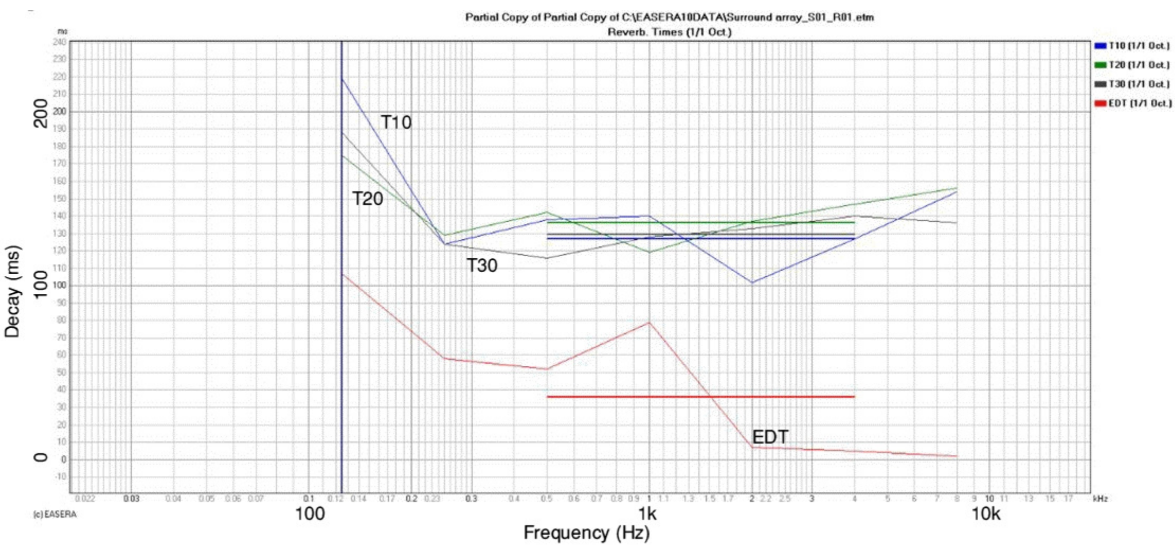

Figure 1. *Room decay time at reference (a) position*

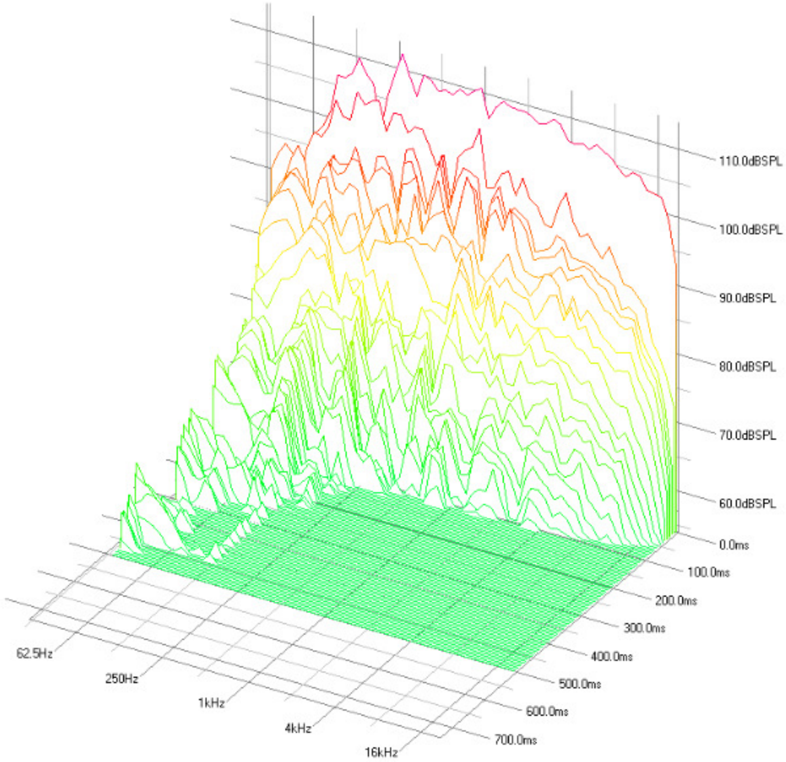

Figure 2. *Room decay waterfall plot at reference position*

It can be seen from Figure 2 that the room under test was a very clean, 'acoustically sterile' environment, where loudspeaker interaction could easily be measured without undue disturbance from room effects.

2 THE TEST SET-UP

The system under test consisted of five individually-amplified and processed surround elements (loudspeakers). The loudspeakers (Electromotive Laboratories, Satellite prototypes) were especially designed for immersive surround systems. They are specified individually to work down to 80 Hz, and have a polar response with -6 dB at around 120° between 500 Hz and 10 kHz, in all axes. For general surround use in a Dolby Atmos system (for which the room was equipped), they normally work in conjunction with a bass-extension system, although it was disconnected for the purposes of this test.

Each loudspeaker was verified at manufacture to be no more than ±1 dB from the reference production plot. Unlike later versions of the Satellite loudspeaker, the prototypes were passively crossed over and remote amplified, which is actually typical of most cinema systems. All channels were controlled by a network of BSS Soundweb processors, with each loudspeaker having its own processing channel.

During the test being reported here, all system-processing for Dolby Atmos was deleted from the Soundweb processors. The only processing left on each channel was a simple HF roll-off filter to tailor the elements to the specified X-Curve response. No loudspeaker corrective EQ was active, or in the signal path, and all channels were identically processed.

All five loudspeakers in the test group were mounted on the same wall in the same horizontal plane. Each element was identical, and all were verified to have the same on-axis responses. As the walls were highly absorptive, even to quite low frequencies, each element could reasonably be considered to be operating in free space. Initial analysis of the on-axis response at 2.5 metres showed a notable reflection from the floor, which was the only significantly reflective surface in the room. The primary reflection-dip was noted to be at approximately 180 Hz, with a secondary dip at 360 Hz at a reduced level. These reflections can be seen in the unsmoothed response plot shown in Figure 3.

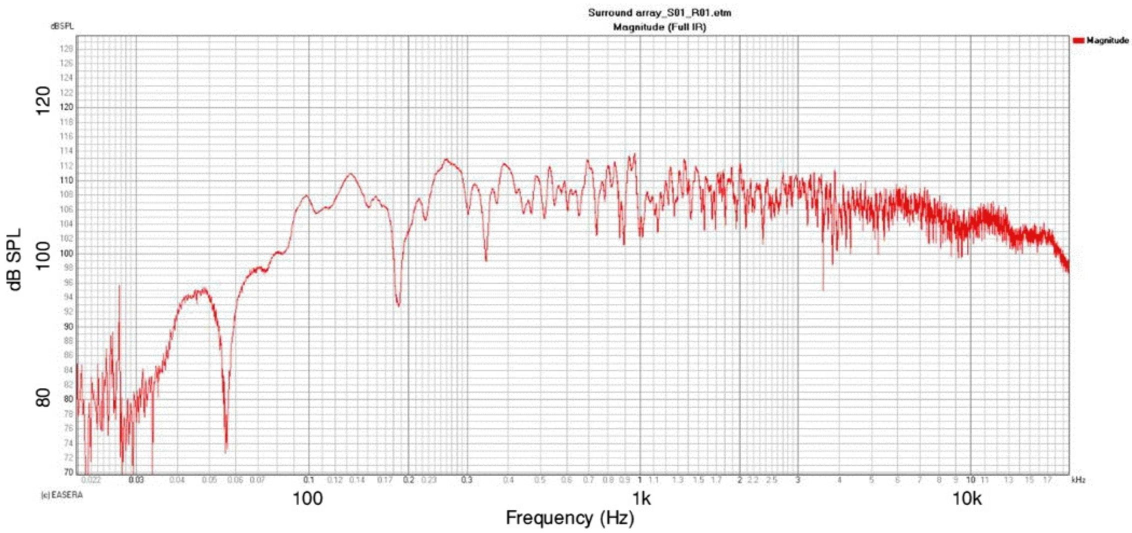

Figure 3. *Unsmoothed on axis response of one single array element*

3 INITIAL MEASUREMENTS

A reference measurement was initially taken of one, single, unequalised array-element, and the result is shown in Figure 4.

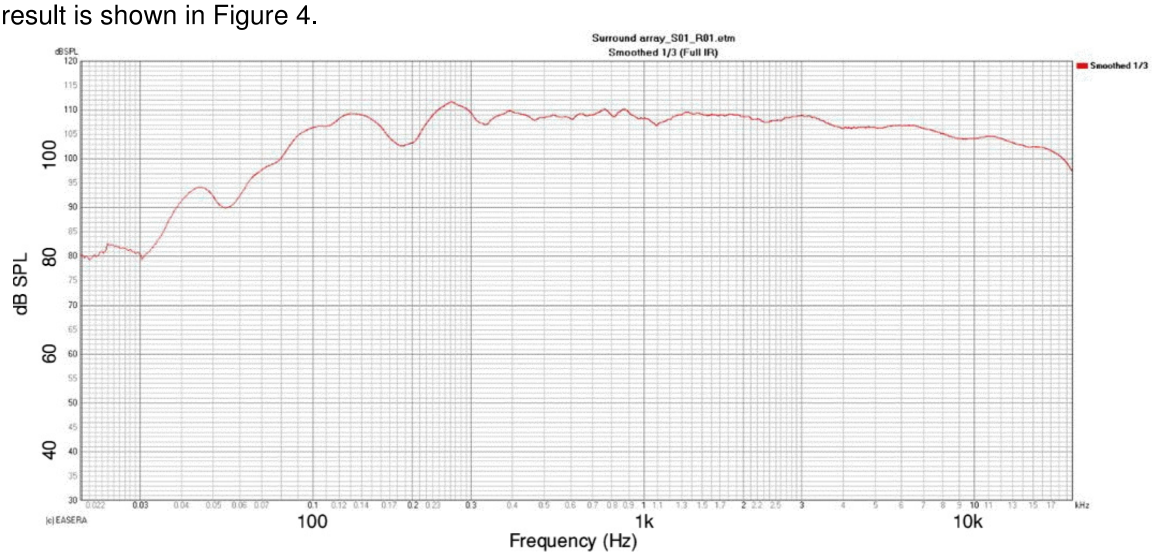

Figure 4. *Third octave smoothed on axis response of one single array element (with a simple HF roll-off filter to comply with the X-curve requirement)*

It can be seen from Figure 4 that the unequalised response of the single array-element is well within the target response the certification process, the only noticeable anomalies are the floor reflections, which are, in any case, *not* correctable with equalisation. All array elements were verified to follow this response individually, without any post-processing window. As an array element, this response would be deemed to pass any certification process.

3.1 Multiple Elements

The next stage was to look at how a second, identically-performing element (loudspeaker), alongside the reference element, would affect the measured response taken on the axis of the initial (reference) element at a distance of about 1.5 metres. The result is shown in Figure 5.

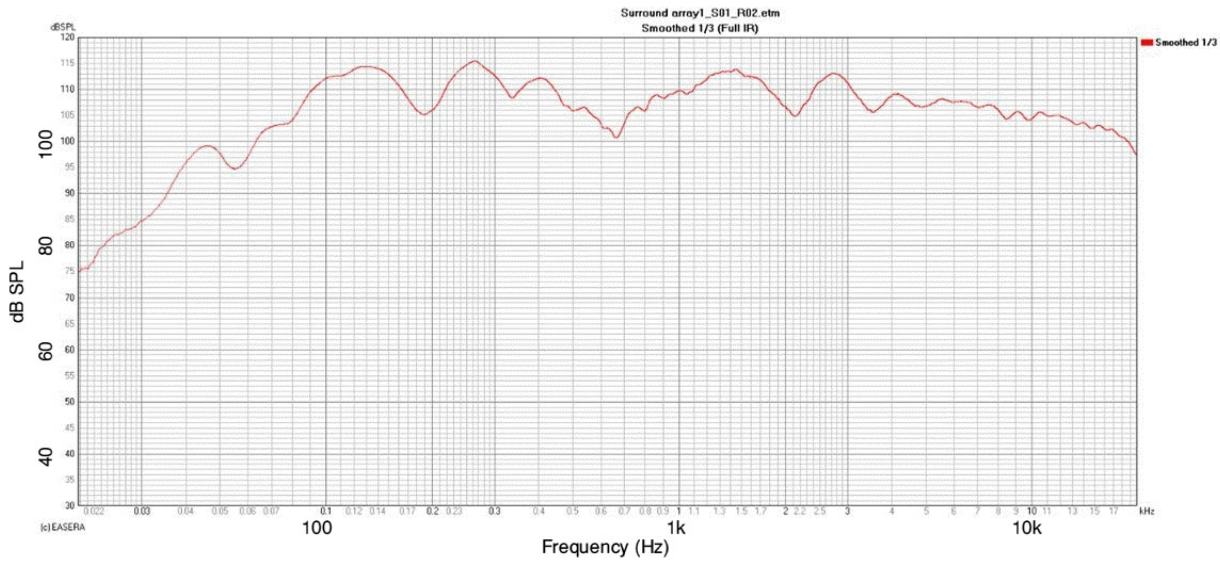

Figure 5. *Two array elements measured on-axis to one array element*

It can clearly be seen that the measured, combined spectral-response has significantly worsened in its general uniformity. Although both sources are identical in polar and frequency responses, their physical displacement has led to significant interference issues due to the different arrival times, which can clearly be seen in Figure 6. It can also be seen that the sound from the second, more distant loudspeaker arrives at lower level than that from the first, as does its associated floor reflection.

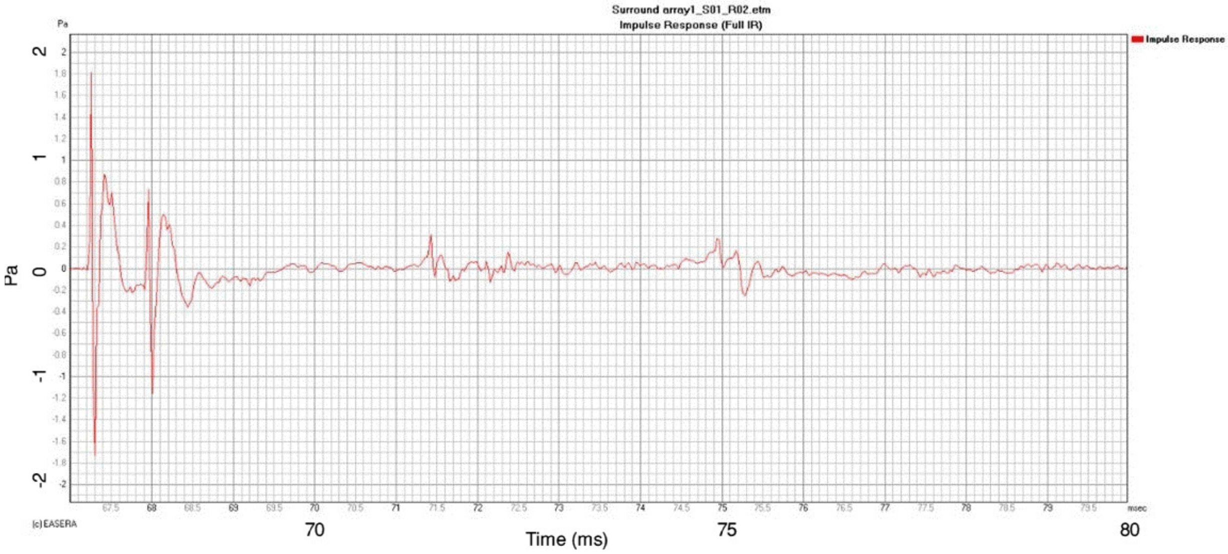

Figure 6. *Impulse response at measurement position A, with two array elements running*

By contrast, Figure 7 shows the measured result after moving the microphone to a position roughly equidistant from each loudspeaker, although not precisely so. From this position, the levels arriving at the microphone were substantially similar, although a small offset in arrival-time remained. A significant response-dip is clearly apparent at about 1.1 kHz, which would be sufficient to tempt a large proportion of systems engineer to try to equalise it. However, the response shown in Figure 4 still represents that which was emanating from each of the loudspeakers, and suggests that no action should be taken. So; Figure 5 suggests that there may be minor issues at 650 Hz and 2.2 kHz, while Figure 7 suggests very strongly that there is a problem that demands attention at 1.1 kHz. There is conflicting information here, yet a simple walk-by test would suggest to the listener that the sound was still more or less similar to that represented by Figure 4, albeit with a slight ‘hollow’ sound in the region of the measurement positions where position-dependent dips occurred.

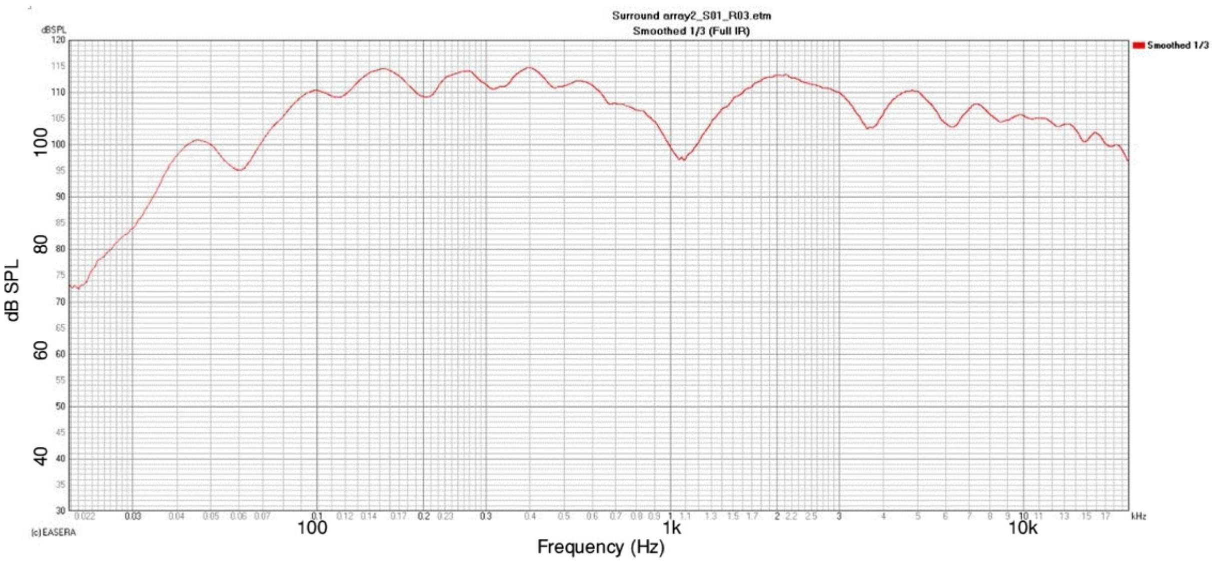

Figure 7. *Measurement microphone, placed roughly between two array elements*

Nevertheless, no equalisation could correct the audible ‘hollowness’. Equalisation could only *add* to the colouration, because any equalisation that was applied would inevitably change the individual responses away from that shown in Figure 4, which had previously been deemed to be correct. Furthermore, no matter what a spectrum analyser might show, the arrival times would still be different with *any* asymmetrical measurements, and so, the cancellations would still occur.

4 FURTHER INVESTIGATION

For this test, a group of three loudspeakers were used, consisting of the reference loudspeaker and one loudspeaker at either side of it. This simulated a section of an array with a typical pattern of interaction in a ‘normal’ cinema, although slightly simplified. Figure 8 shows how the three loudspeakers responded at an on-axis position, 2.5 metres from the centre loudspeaker (position A) The general trend follows the response of the single element, as was shown in Figure 4, but there is a less-even mid-range, with a wide and flat ‘hole’ from about 350 Hz to 1.1 kHz. Despite this, the response does not appear particularly troubling, and would suggest that a broad equalisation-boost in the lower mid-range would be beneficial to the sound quality.

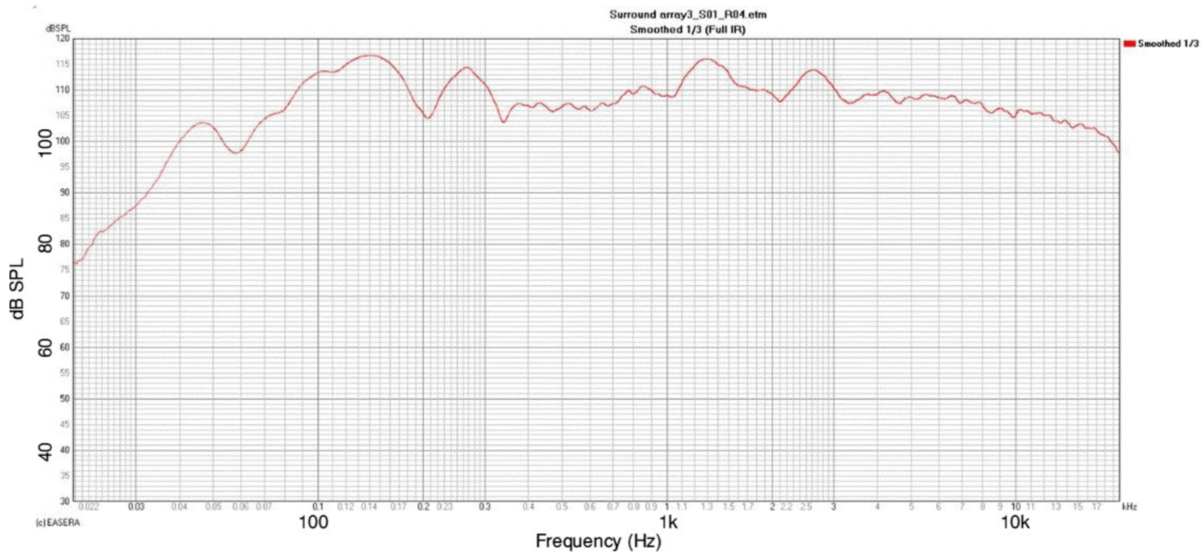

***Figure 8**. Three loudspeakers measured at position A, perpendicular to the centre loudspeaker*

Measurements taken at two other positions (B and C – see Figures 9 and 10) show a different picture. [The loudspeaker and microphone positions are shown in the Appendix at the end of the paper.] What would immediately look like a good correction for position A (in Figure 8) would lead to a serious problem at position C. In reality, there is no common correction in the mid-range that would result in an improvement at all three positions. Despite this, as already stated, many cinema-calibration technicians are prone to try to equalise the responses as seen, because they have been taught that dips like these should not be left uncorrected. Two of the three plots (in Figures 8 to 10) have a dip at 1 kHz to 1.1 kHz, and even a prudent technician may try to apply a little wide-and-gentle equalisation, such as shown in Figure 11.

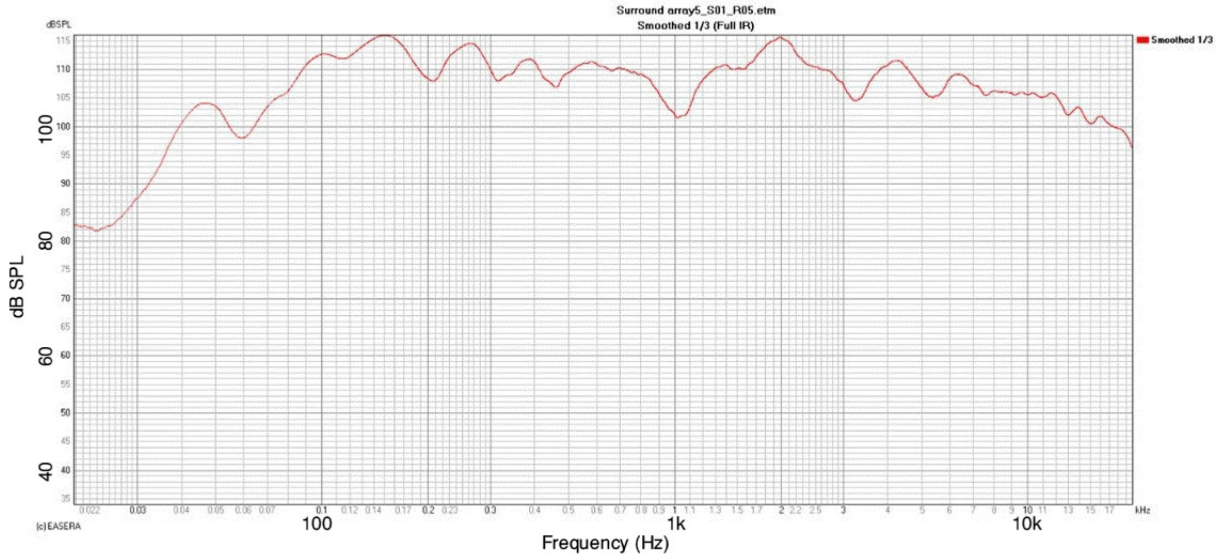

Figure 9. *Three loudspeakers at position B, equidistant from loudspeakers 2 and 3*

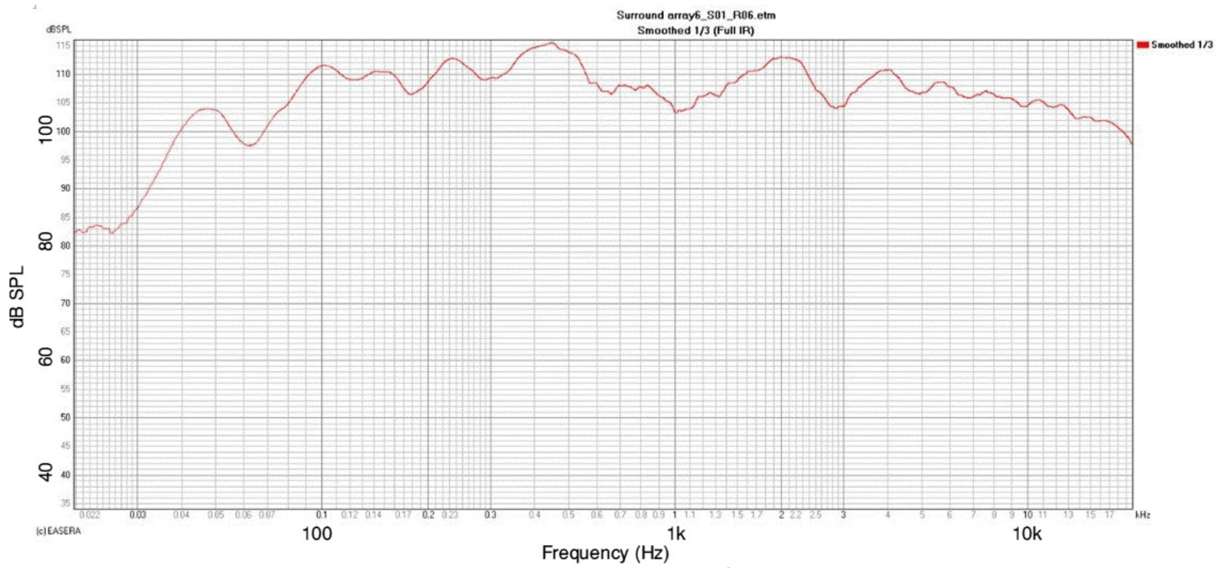

Figure 10. *Three loudspeakers measured at position C, perpendicular to loudspeaker 3*

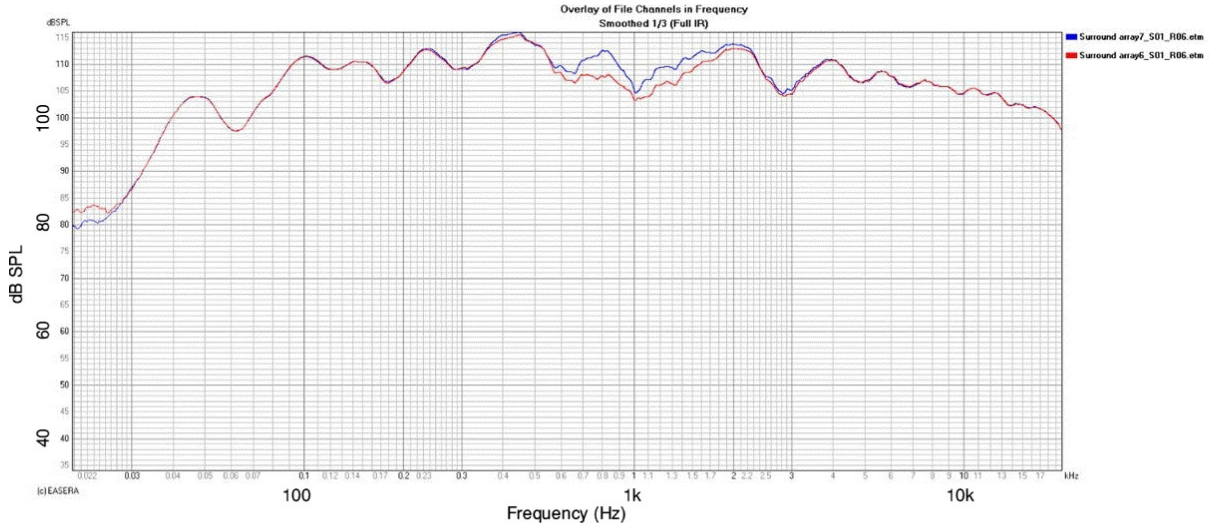

Figure 11*. Small correction (blue line) applied at 1 kHz, one-octave wide, +4 dB measured at position C*

As can be seen from Figure 11, the application of one, simple parametric-filter to the response shown in Figure 10 would have some apparently-beneficial effect, but there would be an alarm bell ringing for an experienced systems-engineer because the applied equalisation has not created a correction that is proportional to the equalisation applied. Something is evidently cancelling at 1 kHz, and the equalisation cannot bring it back, even though the surrounding response is rising. This is more likely to be a function of the one-third-octave-band smoothing, rather than a real, wide dip, but people do tend to insist on making measurements with smoothing applied. [Historical, legacy procedure, perhaps.]

Looking at Figure 12, many systems-engineers would be happy with the ‘correction’ done at position C. In fact, it has also translated well to position B, and has perhaps even worked *better* there, but there is still something untoward in the 1 kHz region that does not seem to be right, even though things appear to have significantly improved around it.

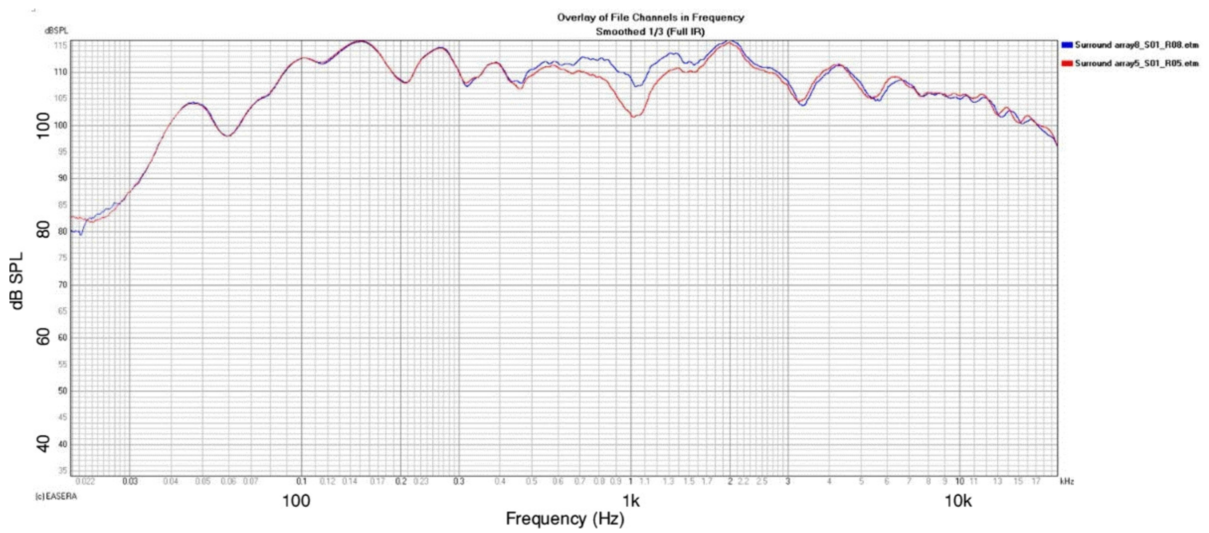

Figure 12. *Correction from Figure 11 measured at position B*

Nevertheless, however well the 'corrections' seem to have responded elsewhere, Figure 13 shows that they have resulted in a significant peak in the direct responses which would give a significant "honk" to the mid-range, and which was *not* an intended consequence of the equalisation. In the direct field from each loudspeaker, the acoustic cancellation issues at 1 kHz were not there, so the applied equalisation had a proportionate effect upon the response. The 4 dB of 'excessive' boost is clearly visible in the octave around 1 kHz, so a conscientious systems-engineer would not want to apply this, even though the *combined* response in Figure 8 would appear to require it.

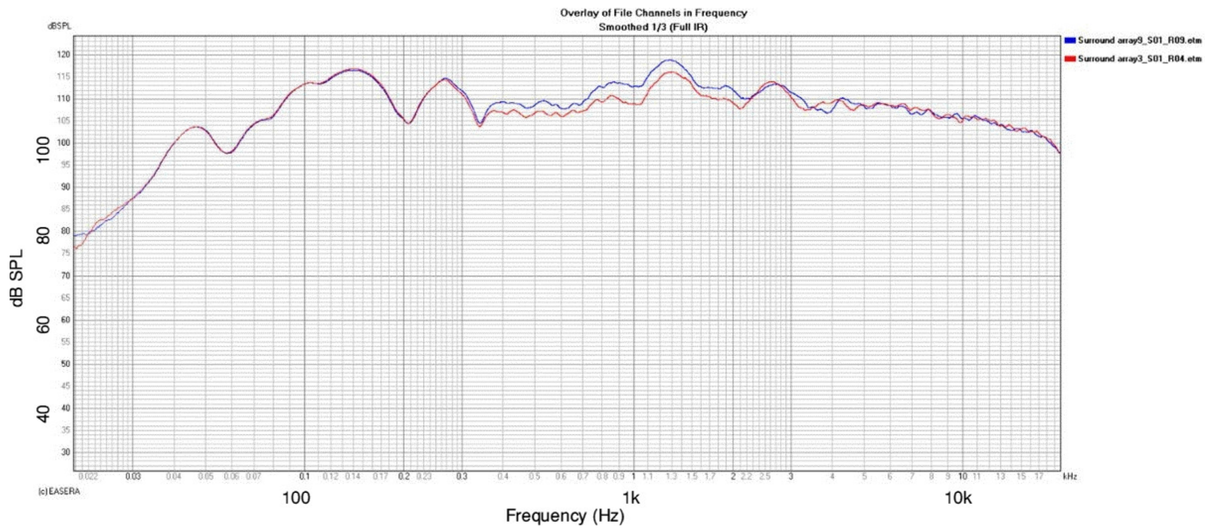

Figure 13. *Equalisation from Figure 11 measured at position A*

Even the *averaging* of the responses at the three measurement positions would not help. Figure 14 clearly suggests that a boost would be required around 1 kHz, but this would be very detrimental to the response at position A, as shown in Figure 13. It would suggest a one-third-octave cut of 5 dB around 1.3 kHz, but most definitely *not* a 5 dB boost at 1 kHz, yet even that cut would not work on the plots shown in Figures 11 and 12. The average therefore does not proportionally represent the individual positions.

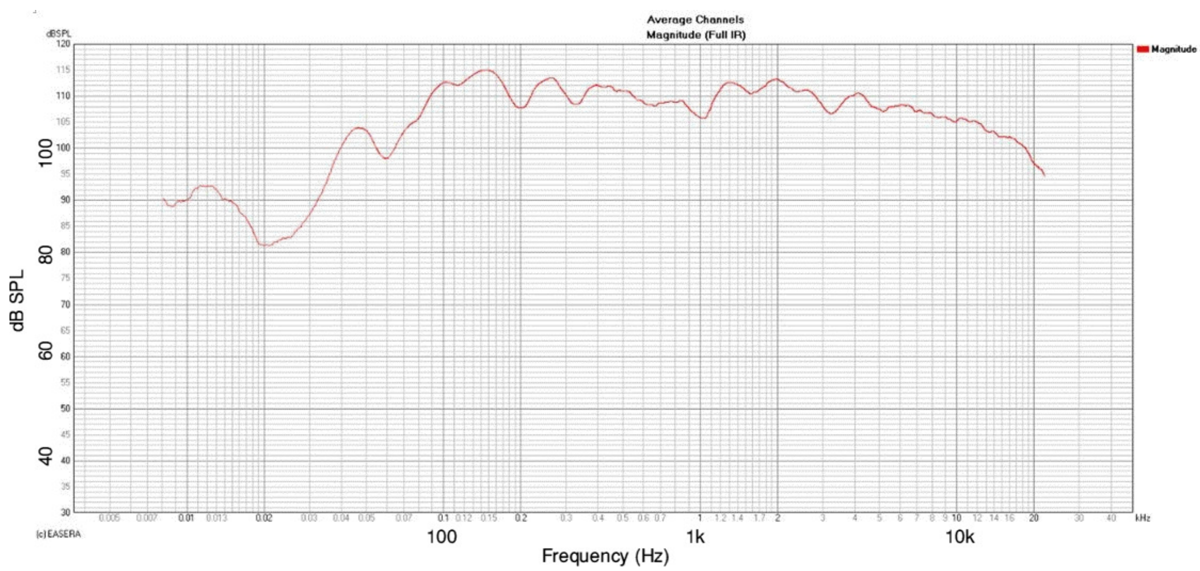

Figure 14*. Energy sum average of positions A, B, and C before equalisation*

The reality of the situation is that none of the information from the microphone positions is particularly useful regarding how anything should be corrected in any overall beneficial manner. Even the averaged response does not relate to all the listening positions because the microphones are in an interference field which is highly position-dependent. Figure 15 clearly shows how, at measurement positions A, B, and C, there is a completely different arrival-time from each source. In each case, the summation of the sources will create a very different interference pattern, and hence a different and unique frequency response plot on the analyser: as shown in Figures 8, 9, and 10.

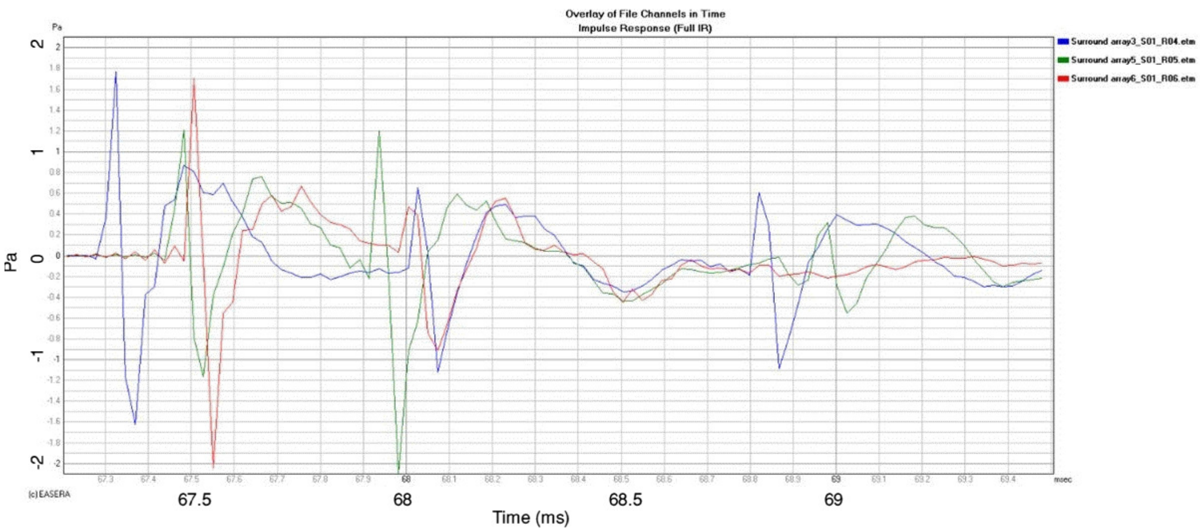

Figure 15*, Overlaid impulse response of measurements from figures 8, 9, and 10.*

No points in the room will ever have the same arrival characteristics, and thus would never have the same frequency response. Nevertheless, the ears and brain can make much more sense out of this than any analyser can. We cannot 'average' the time-dispersed responses to create a '*representative*' frequency response for the average listener because the concept of an average listening position is unrealistic. Any attempt to find a point (or points) from which to measure and equalise a distributed-source array is a highly improbable goal to achieve, yet despite this, many cinema-alignment processes still require the system-technicians to perform this very task.[1, 2]

As Floyd Toole pointed out in a communication to the SMPTE (August 4th, 2010), and similarly in his 2008 book, 'Sound Reproduction', '*Unlike a human, the microphone does not take any note of the angle of incidence of the direct and reflected sounds; nor does it make any allowance for the time of arrival of those sounds; nor does it acknowledge spectral variations among any of the sounds. The microphone simply adds them together. .....It is well known that two ears and a brain are vastly more analytical than a microphone and an analyser. Humans respond differently to sounds arriving from different directions at different times'*. Microphones, by contrast, do not.

As was said Section 3.1, a simple walk-by test would suggest to a listener that the general sound of the spectral balance of the array would be more or less similar to that of an individual loudspeaker. From the sound alone, it is probable that an experienced calibration technician would *not* choose to apply as much equalisation as would be indicated on a spectrum analyser, yet the temptation usually seems to be there to 'appease the screen'.[3]

5 A MORE REALISTIC CINEMA ARRANGEMENT

So far, the measurements described here have been based on the somewhat idealised responses of only three positions and three loudspeakers in a relatively dead room. Some schools of thought suggested that a greater number of loudspeakers in a more reverberant room can smooth out the responses, but the room in which the above measurements were made complies with (and is certified to) the Dolby studio requirements and is Atmos equipped. It is a highly-rated room for mixing the soundtracks of feature films. Furthermore, one, standard alignment procedure is applied for all sizes of rooms and for all numbers of surround loudspeakers, so whatever technique is used needs to be robust.

A further test was carried out using a full ‘side-surround’ channel, known as Rss (or Right, Side, Surround) – (see Appendix). In this case, the set-up consisted of five individual elements, although in some cinemas the Rss array can consist of ten or more elements. In practice, in most cases, it would be hard to find a seat where more than any five of the elements were really covering the listener, although in more lively room they could all be contributing to the reverberation. However, in *modern* cinema rooms, the decay time is usually so low that no real reverberation exists, so the more distant elements would be too far away from any given position to significantly affect the responses of the nearer elements.

Figure 16 shows what happens when the five surround elements in the test were radiating the same signal and being measured at five individual listening positions. Up to 14 dB of variation is evident from one microphone position to another, which, should the analyser display truly represent what the listener hears, would suggest that there would be an unacceptable amount of tonal difference from seat to seat. However, in reality, such is not the case, because rather little tonal difference is actually perceived from seat to seat.

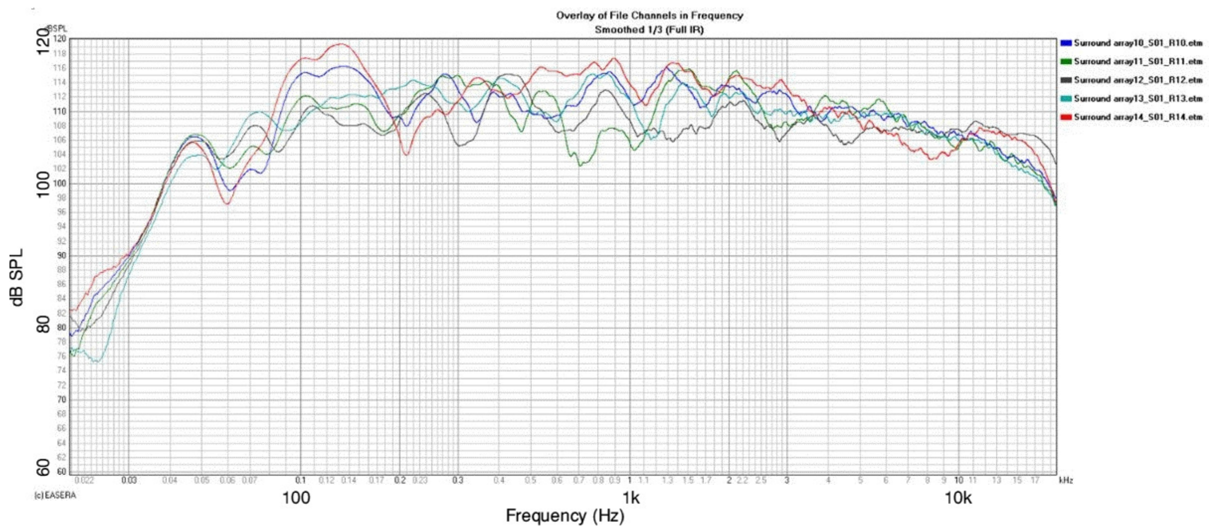

Figure 16. *Five loudspeakers radiating simultaneously, measured at five locations*

During the test, a walk-around in front of the array showed a reasonable consistency in tonality, albeit with some noticeable ‘hollowness’ as the listener passed between close-by elements, but the overall tonal balance was considered to be very acceptable. Of course, each of the five array elements was still responding exactly as shown in Figure 4, which was deemed to be a very good response, so given Floyd Toole’s comments, quoted above, it is not surprising that the overall sound was still very acceptable to the listener. It was merely a combination of five correctly-responding elements, which the ear perceived as actually being a combination of five, tonally appropriate sources.

5.1 Further Tests

In order to further investigate whether a generally-corrective equalisation could be found, an average response of the five positions was used as a reference for equalisation in the mid-band. Positions were chosen which seemed to each have a noticeable problem when assessed by measurement alone, evident on the five-position averaged plots. However, in many cases, this correction had a negative effect on the smoothness of the frequency balance when assessed at the individual listening positions. Also, in every case where a ‘problem’ in the response of the multiple, interfering loudspeakers was evident from the measurements, it was found that attempting to correct the problem using analysis and equalisation resulted in the worsening of the measured responses elsewhere, with no improvement in the audibly-perceived response.

As can be seen in Figure 16, no common deviation across all responses could be found. Where one response would seem far away from the general pattern, the only general pattern that *could* be seen tended to loosely follow the responses of the individual array element.

Figure 17 shows the average response over five positions. A broad dip is apparent between about 500 Hz and 1.2 kHz. Somewhat like the case shown in Figure 8, this would suggest to many technicians that a corrective equalisation should be applied.

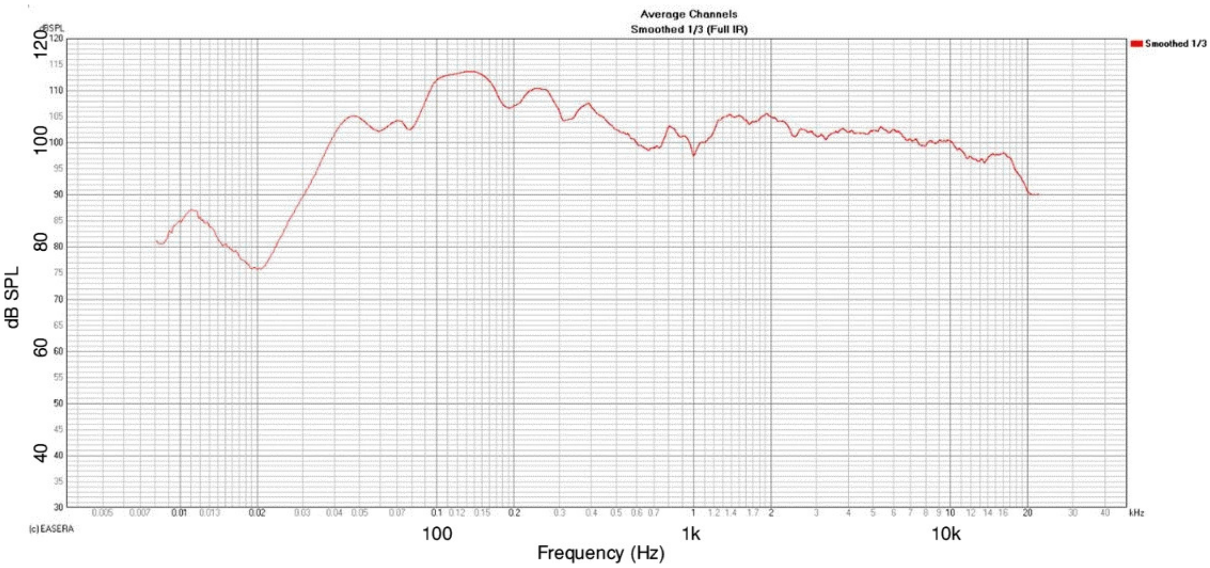

***Figure 17**. Average sum showing dip between 500 Hz and 1.2 kHz, which would suggest that correction was needed*

Figure 18 shows the measured response at position B after a ‘correction’ had been applied to the averaged response. Compared to the red, unequalised trace in Figure 18, it could be said that the correction applied to Figure 17 was also, on balance, beneficial at position B, despite the peak around 900 Hz, although some people may argue that this peak could be potentially more colouring than the broader dip that was present before equalisation.

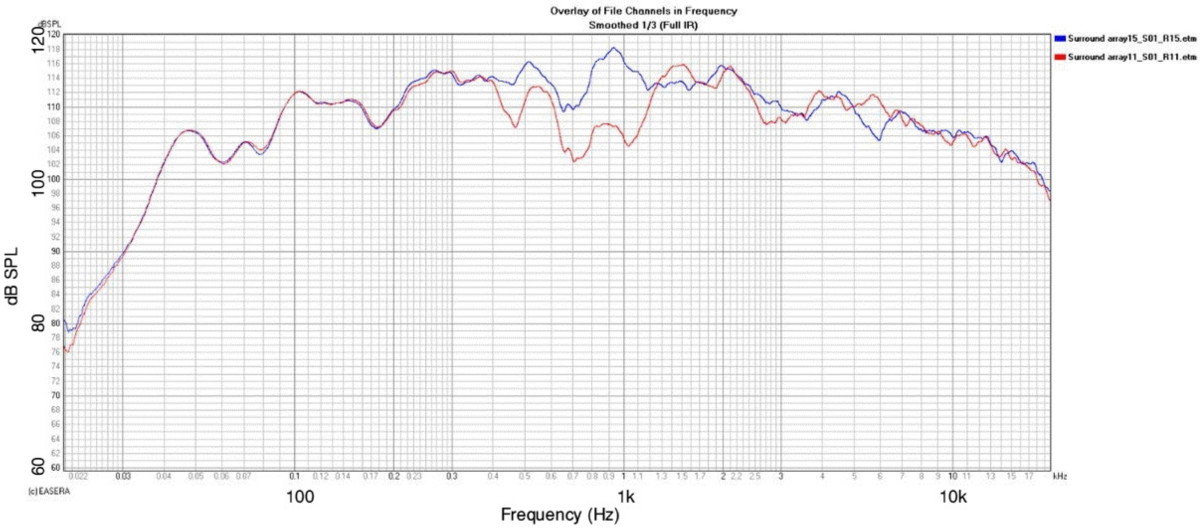

Figure 18. *Corrected five-loudspeaker array, measured and corrected to single microphone at position B. Blue = equalised response*

By contrast, at position D, Figure 19 shows a very different set of circumstances Before equalisation, the dips in the mid-range response were at significantly different frequencies from those evident at position B. There was also no broad trough in the response, but rather a more even series of peaks and dips. After ‘correction’ of the averaged response, the equalised response at position D shows an unacceptable peak around 900 Hz, which almost no calibration technician would wish to see (or hear). However, if only the averaged responses of the measuring microphones were being used as a reference, such alarming responses at individual positions may well (and definitely do) go unnoticed – although perhaps not by the unfortunate member of the audience sitting in that seat.

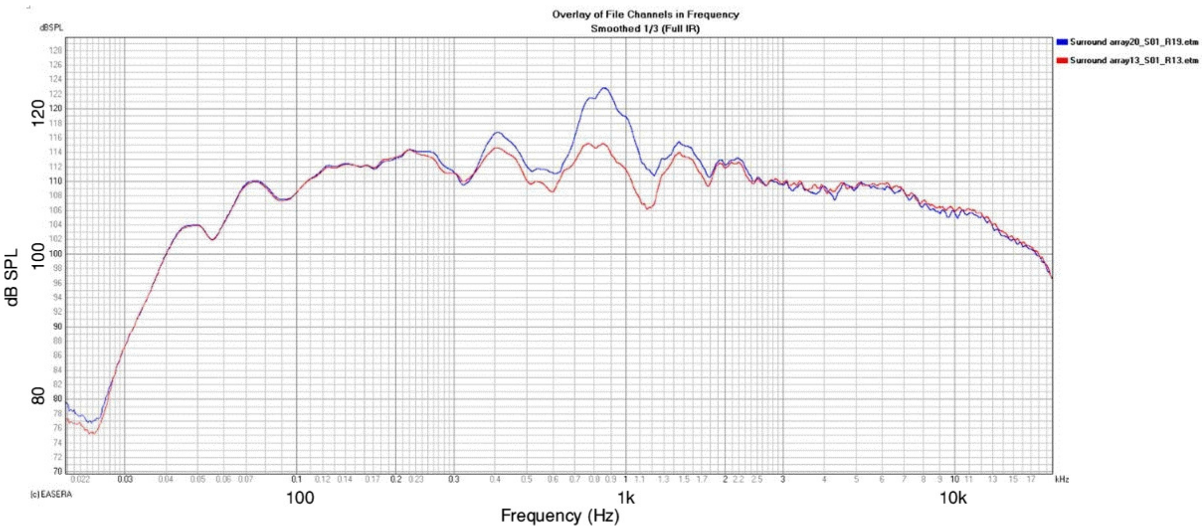

Figure 19. *Result of corrective equalisation from Figure 17 at position D. Blue = equalised response*

Figure 20 shows the response at position C, both before and after equalisation. Between about 200 Hz and 2 kHz, it shows yet another response pattern which bears little similarity to those at positions B or D, nor even to the averaged response. Once again, the correction which seemed to be beneficial to the average response has introduced an unwanted peak around 900 Hz. In general, peaks like this are potentially far more colouring that the dips which they may be intended to correct, so, as a consequence, subjectively, this may be no correction at all.

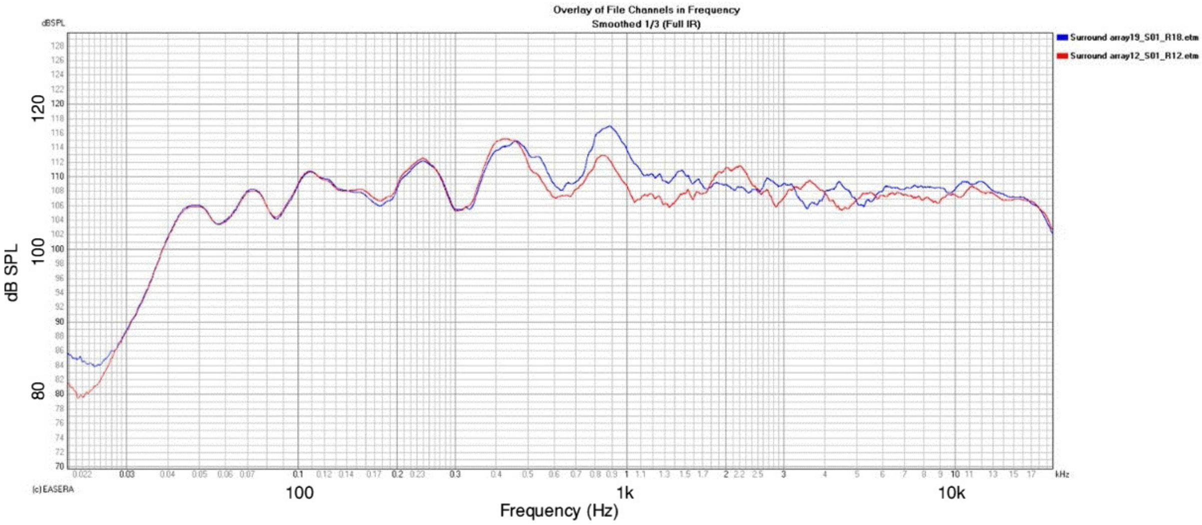

Figure 20. *Result of corrective equalisation from Figure 17 at position C. Blue = equalised response*

At position A, as shown in Figure 21, the response after equalisation shows a pronounced double peak, either side of 1 kHz, in a region which had been within generally-acceptable limits prior to equalisation. Once again, the apparent correction of the averaged response has changed the response at a particular position which would have almost certainly increased the perceived colouration. The absurdity is all of this is that the broad, measured dip in the response shown in Figure 17 would probably not have sounded coloured, so what would be the point of the applied 'correction'? Compliance with the industry norms that require it to be done is perhaps the only valid reason, but these probably need to be changed.

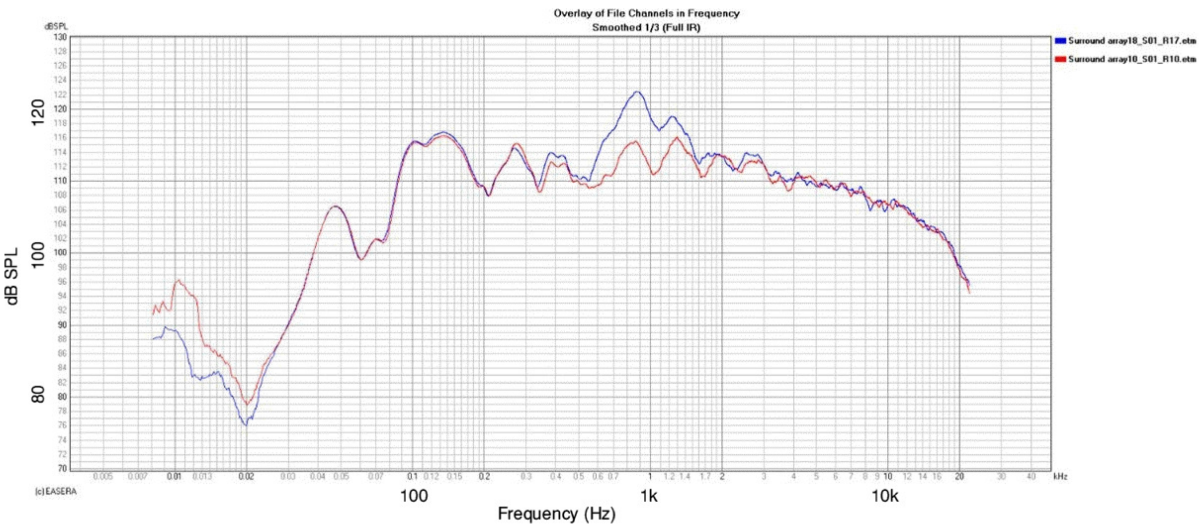

Figure 21. *Result of corrective EQ from Figure 17 at position A. Blue = equalised response*

Most loudspeaker arrays that are not made up of individually-processed elements, such as in immersive sound systems, have only one 'correct' set of equalisation, delay, and gain setting for one position. As such, it is only possible to correct them for one position in space, yet doing this contradicts the motive for using a distributed source in the first place. As can be seen from Figure 22, each measurement position is receiving five distinct signal-arrivals at five separate times, all of which sum to produce the measured frequency response to a time-blind spectrum analyser. Figure 15 clearly shows that at different microphone positions there are different arrival times, and different densities of arrival times, all of which will contribute to the vastly differing amplitude responses which can be seen in all the measurements presented here.

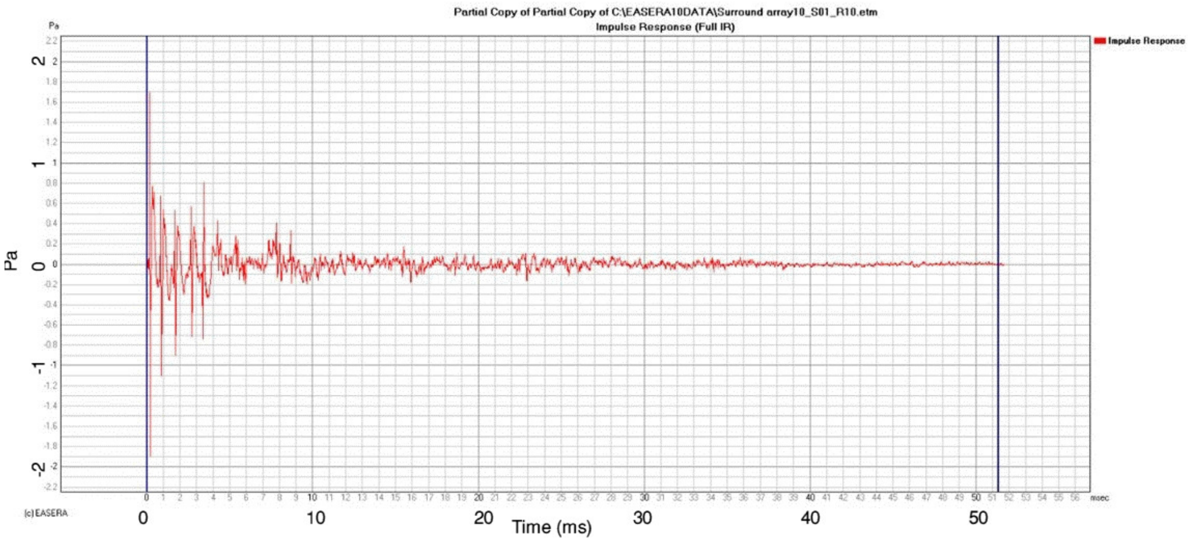

***Figure 22**. Impulse response showing arrivals of a five-loudspeaker array to a single microphone position.*

6 DISCUSSION

The whole point of using 'surround arrays' of multiple, similar loudspeakers is to try to present a diffuse-sounding ambience to the audience in a way in which the general sounds will not be localised, but treating the averaged response of five microphones does not make it possible to correct the array for most listeners. Indeed, it can make it worse for some positions, so using a multi-mic average is not a valid solution to equalising an array.[4] In nature, no such thing exists, as no identical sound would emanate from multiple sources simultaneously. The use of this technique in cinema systems generally began with the introduction of Dolby Stereo, in the 1970s. This was a four-channel system which fed the left, centre and right channels to loudspeakers behind the screen, and where the remaining, single, mono, ambience channel was fed to multiple loudspeakers, distributed around the sides and rear of the room. The concept of using an averaged response to measure the whole array which was something developed empirically as a practical way of assessing the situation in each theatre under some very dubious acoustic conditions. The method was never based on any scientific study.

Subsequently, the mono array was divided in two halves for Dolby Digital and their formats such as Sony SDDS, the divided further for. 6.1, 7.1 and so forth. The original calibration technique was transferred to the expanded systems without much further analysis, and the concept that the measured, average response would somehow relate to the perceived response was somehow based on perceived 'common sense' rather on any, further scientific study. Given the background, it was somewhat inevitable that the consistency and the sonic neutrality of the surround channels would be inconsistent from cinema to cinema, yet the technique continues to be applied. Also given that nothing in nature emits identical sounds from multiple sources, the whole concept of producing ambience in this way is effectively only trying to trick the ears; but ears are not so easily fooled.

It is now many years since Floyd Toole published the drawing shown in Figure 23. What a measuring microphone receives is 'time blind'. Over a period of time, an analyser sums all the pressures from all the sources and their reflections. Summing the outputs of multiple microphones only increases the chaos. By contrast, the ears make the brain aware of individual source locations, and also recognises their different arrival times and different spectral balances. The brain needs no 'help' from equalisers because it perceives the reality of a situation – an ability that has been crucial to our survival since prehistoric times.

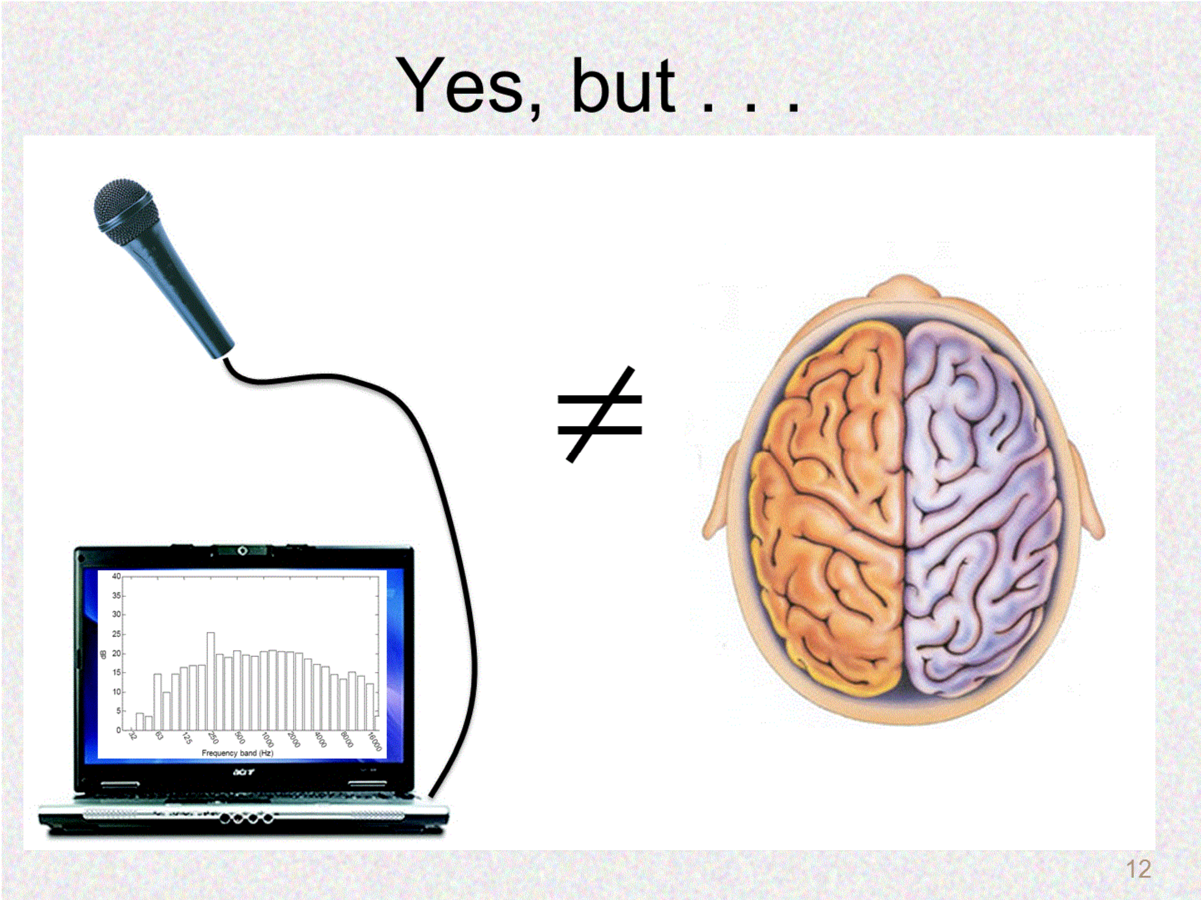

Figure 23. *What is shown on a time-blind analyser and what is perceived by the ears and brain are not equal*

7 CONCLUSIONS

The measurements presented in this paper have shown that attempts to globally equalise an array of multiple, identical, spacially-separated loudspeakers, in order to attempt to correct deficiencies seen in the measurements, will not, by any means, lead to a general improvement in the response at individual seating positions. In fact, severe colouration can result at individual seating positions after attempt to use global equalisation to 'flatten' a whole array of loudspeakers.

It has been noted by experienced listeners that arrays of similarly responding loudspeakers are more likely to timbrally match a single source with the same measured response when each individual loudspeaker, at least in the mid range, measures similarly to the single source. That is to say, the best timbral match between a surround array and a screen loudspeaker is achieved when the close-field response of each loudspeaker in the array closely matches that of the screen loudspeaker. Attempts to match the measured response of a whole array to that of a screen loudspeaker, measured at the same spacially-averaged points in a room, will tend only to introduce unwanted colouration into the response of the array. What is more, at individual positions, the colouration can be very significant.

It is also worth remembering that no cinema alignment engineer ever calibrates the LCR channels with more than one of them playing at the same time, even though signals may be panned between them. They are not even *checked* in pairs, yet they play together as one sound field without any complaints of timbral imbalance. In some ways, it is puzzling that it should be *required* to align cinema surround arrays in entire sections, but it is most likely that as the mono surround arrays were, historically, series-parallel connected arrangements, powered from one amplifier, the process of calibrating them all at once became common practice through a combination of necessity and legacy, rather than being instigated as a process of 'good practice'. In reality, it is a total anomaly that has zero acceptance elsewhere in the world of professional audio.

The authors now intend, along with others, to carry out intensive subjective testing to determine in a more scientific way how people perceive the spectral balance of multiple sources when compared to a single source in similar acoustic conditions.

8 REFERENCES

1. SMPTE Standards document ST202:2010 – 'Dubbing Stages (Mixing Rooms), Screening Rooms and -- For Motion-Pictures B-Chain Electroacoustic Response -- Indoor Theaters.'
2. SMPTE Standards document RP 2096 – 1 / 1.12.6.b. 'Measuring and Adjusting Surround Channels' Electroacoustic Response.'
3. Newell, P., Holland, K., Torres-Guijarro, S., Newell, J., Santos-Dominguez, D., 'Human Factors Affecting the Measurement of Rooms', Proc. of the Institute of Acoustics. Reproduced Sound Conference, Vols. 34, Part 4, Brighton, UK (Nov. 2012)
4. Newell, J., Newell, P., Holland, K., 'Consideration for the Generation and Measurement of Low Frequency Effects in Cinema Rooms', AES 57th International Conference, Hollywood, CA, USA, (March 6–8 2015)

APPENDIX

The measurement microphone positions relative to the five surround loudspeakers