

An overview of multichannel level alignment

Nick Zacharov

Nokia Research Center, Speech and Audio Systems Laboratory, Tampere, Finland

`nick.zacharov@research.nokia.com`

As multichannel sound systems become more and more widespread, the issues to how to obtain optimal sound reproduction become more apparent. Although the matter of level calibration is rather trivial for stereo sound systems, it has been found that it does critically affect perceived sound quality. The more complex and often sub-optimal multichannel set-up introduces a whole new range of problems in this respect. This article aims to provide the reader with some background into the questions surrounding multichannel level alignment and discusses some of the topical issues and research presently in progress.

1 INTRODUCTION

The purpose of this paper is to provide the reader with a background into the issues of multichannel level alignment and review some of the work performed to date.

Multichannel systems consist of numerous loudspeakers, often acoustically dissimilar, set-up in a sub-optimal manner, due to the practical constraints of the domestic environment. These types of tendencies away from the standardised and idealised multichannel set-ups of ITU-R BS 775 [15] can lead to significant variations in the amplitude response of each channel, leading to differences in the perceived channel level. This is an important factor, as level alignment has often been shown to be critical to the perceived quality of reproduction [1, 2, 4, 21]. Whilst the matter of multichannel level alignment has been addressed, this issue is still far from being understood and resolved.

To study these matters in greater depth, this paper is divided into two main sections. The background section will provide a review of the level alignment and will then proceed to consider issues associated with the reproduction system, the acoustic environment, and briefly consider aspects of perception. Section 3 will consider some new perspectives in level alignment research, discussing the matters of directional loudness, current studies in level alignment and also the influence of source directivity. A summary is presented in section 4.

2 BACKGROUND

With the now widespread availability of multichannel audio, commonly in the form of so called 5.1 channel systems [11], the issue of how to obtain optimal quality of reproduction is once again apparent.

Subjectively, the level or loudness of a reproduced sound has been considered to influence the perceived quality from an early stage [24]. It has been demonstrated in the literature that the perceived quality of reproduction of any sound systems is partially related to level.

Aarts considered level alignment a critical issue in the subjective testing of loudspeaker, to avoid biasing tests [1] and this view has been supported by other researchers in the field. If the levels of compared systems are not equal, a masking of factors or biasing of results can occur which cannot easily be dealt with statistically in listening tests. This type of 'intensity'* masking of other factors is a well-known phenomenon in other psychometric testing and is extensively discussed in the fields of sensory evaluation techniques, such as flavour and smell. It is often the case that the intensity of the product under test must be normalised, such that products of near equal intensity are compared.

Aarts considered different methods of evaluating the loudness of loudspeaker reproduction employing different methods and compared this data against subjective

* In this context we refer to intensity in the non-acoustic sense

alignments [1, 2]. He concluded that the best manner for the alignment of frontally placed loudspeaker levels was by use of the involved Zwicker loudness method outlined in [19] and ISO 532B [14]. His finding also concluded that the A-weighted sound pressure level (SPL) measurement was less suitable for alignment purposes of this nature.

In a later study Aarts [2] continued this work by considering the suitability of other linear SPL measures including A, B, C, and D-weightings in comparison with the Zwicker loudness metric. Once again compared with subjective alignments, the Zwicker loudness metric proved most favourable, whilst the B-weighted SPL measure was also found to be satisfactory. The A-weighted SPL measure again failed to find favour for this task.

Bech [4] has also found that within the limits tested, that perceived quality does increase as a function of level for multichannel sound systems. This result was concluded in an audio-visual subjective experiment on the influence of stereo base width on perceived quality. Three different base widths were compared at two different SPL: 70 and 80 dB (linear), measured with a pink noise signal at the listening position.

However, under different circumstances, when the level difference between channels is excessive, this may lead to degradation in the quality, as found by Rumsey [21]. In this study Rumsey considered the perceived quality of so called 'up conversion algorithms', which generate a 5 channel signal from stereo source material. His findings were that listeners often found difficulty is differentiating quantity from quality. Also, imbalance in the front/back level that may occur in multichannel systems can also degrade the perceived spatial sound quality.

The matter of correct level alignment of signals is considered of such importance in subjective testing that certain standards strictly define methods of alignment both in the field of multichannel audio reproduction. ITU-R BS 1116 [16] is one such standard that defines the reproduction level with pink noise for each channel under test by

$$L_{ref} = 85 - 10\log(m) \pm 0.25 \quad (\text{dBA}) \quad (1)$$

where m is the number of reproduction channels in the total set-up.

Although this is a well-defined method, it has been found that it is perhaps not suitable under all circumstances as will be discussed later.

In this study we will only consider how to align multichannel systems for domestic reproduction spaces. The issues of cinema calibration have developed and evolved over many years and are now quite well understood and controlled. The reader is referred to

Holman for a review of aspects of cinema sound [7, 9]. In practice films sound is recorded and produced with the well-controlled and defined cinema acoustics in mind. Thus, the cinema reproduction must be considered as the reference situation. With the advent of the home theatres the aim is to transform the cinema sound experience into the home, though the acoustics are far from constrained by comparison to the cinema.

So what does incorrect multichannel alignment lead to? Well the answer to this question is not so simple, but here are a few possible results

- a lack of surround information leading to missing spatial information,
- excessive surround information leading to an unnatural or undesirable effect,
- demasking of multichannel coding artefacts.

It is clear that the apparently trivial matter of level alignment is quite far from that.

The issue of level alignment of the subwoofer, low frequency (LFE) or .1 channel is a complex one [25] and is not discussed in this paper. It is also assumed that matters of time of flight alignment do not influence the level calibration, in itself, and thus will not be considered here. However, time of flight corrections are essential for the correct reproduction of spatial information.

2.1 Reproduction system

The starting point to the level alignment issues is the reproduction system. For the benefit of clarity a brief review of some of the simple sound systems will be considered with respect level alignment.

The simplest mode of sound reproduction is that of a monophonic sound system. This of course has a very simple level alignment strategy which consists of our personal preference of reproduction level and is controlled by the well known volume knob.

The stereo system is the next level of complexity, which nowadays is also considered trivial to set-up. In this set-up, two speakers are employed, which are typically of the same type, i.e. having similar sensitivity, directivity, and amplitude response characteristics.

In practice, when people care somewhat about the quality of the stereo reproduction, the loudspeakers are set-up in as symmetrical a fashion as is feasible.

Lastly, the interested listener tends to be aware that to achieve good reproduction, he should sit on the axis of symmetry of the speaker set-up at an equal distance from each speaker.

Based upon these simple, but now quite well accepted

considerations, the overall level alignment of the system should be quite reasonable without any further user adjustment. The volume control can then serve to control overall level of reproduction

If, however, the loudspeakers are not equidistant to the listening position due to the constraints of the listening environment, this may lead to an imbalance in the left/right alignment. To correct for such anomalies, many reproduction systems contain the well known balance control which provides the user with means of correcting the level alignment of the two channels. This is sufficient, assuming that the set-up has been created as described above. However, when the speakers are set-up in a very asymmetric environment or the loudspeaker are of very different types (and this is not advised) the balance control may not suffice to level align the two channels, we begin to hear problems.

Whilst the stereo system is still quite manageable in terms of level alignment, the multichannel system is not so simple.

Considering the 5-channel set-up, based upon the ITU-R BS 775 standard [15], illustrated in figure 1, speakers should be of a similar type and symmetrically placed. It is generally considered that the speakers should be of a similar type even in practical domestic set-ups, though this is not always the case as illustrated by Holman [6]. Quite often different loudspeaker types are employed for the surround channels and perhaps also for the center channel. As discussed earlier, even the use of non-similar loudspeakers for a stereo set-up can lead to complications in the level alignment. This is certainly a far greater problem with 5 channels.

We are now at the point where a 5-channel balance control is not very feasible and level alignment need to be performed channel by channel. An accepted means of performing such an alignment is to replay a noise signal through each of the channels and allow the user to align the levels to be equal. Whilst this method is clear, the ideal definition of the signal to achieve a level alignment of non-similar loudspeakers is less clear and should also consider the implications of the reproduction environment.

2.2 Reproduction environment

One of the primary complications in multichannel reproductions is due to the complexity of the speaker/room interaction. Of course this is nothing new, but now we must consider numerous loudspeaker, perhaps of different characteristics, which maybe non-ideally located in a non-ideal environment. Each loudspeaker in the room will have to interact with the boundaries of the room to create the amplitude response characteristic at the listening position. This is a complicated issue and it is not the aim of this paper to discuss the basics of room acoustics, which are quite involved and complex to model, but to highlight

some of the sources of variation which include:

- first order wall reflections,
- standing waves,
- acoustic radiation space,
- material acoustic impedance,
- room geometry.

The influence of all of these factors should be accounted for during the level alignment, which further complicates the task in hand.

Under free field conditions with identical loudspeakers, set-up symmetrically, no alignment is required assuming the electrical gains are the same. In this case the levels at the central listening position should be identical. However, this is a very unrealistic situation that only occurs under laboratory conditions. In practice the domestic environment is far from a so-called 'free field' and we must consider the effects of the room interaction and the practical constraints of the set-up.

The issue of symmetry is a critical issue to the speaker room interaction and can be divided into two groups, namely, that of the reproduction set-up and that of the room.

The theoretically ideal symmetrical set-ups for multichannel loudspeaker has been extensively studied and the currently accepted configuration in accordance with ITU-R BS 775 [15] is illustrated in figure 1. In this situation speakers are positioned on a radius of 2-3m at angles of $0, \pm 30^\circ, \pm 110^\circ$. The speaker are here both time and level aligned in terms of the direct sound energy, and only the room interaction should influence the steady state amplitude response. This set-up was created in an ITU-R BS 1116 [16] standard listening room which is highly damped with a reverberation time (RT60) < 0.35 seconds[†]. Steady state amplitude response measurements were made at the listening position and are illustrated in figure 2. As expected it can be seen that there are only very small differences between each channel, dominantly around 200Hz in this case. It was found that to align each reproduction channel with pink noise to a loudness level of 20 Sones (64 dBA), the gains required were identical for all channels. This implies that the room interaction provides little complication in terms of level alignment. In this case level alignment is a rather trivial matter.

In practice such ideal conditions are rarely encountered in the domestic environment. To study the effects of asymmetry a set-up was created to break symmetries, as illustrated in figure 3. In this case the room in itself was acoustically symmetrical, but the central listening position was offset by 1m from the axis of symmetry. To further aggravate the situation speakers were placed with an

[†] within the range 250 - 8k Hz

exaggerated non-symmetry. Once again steady state amplitude responses have been measured at the listening position and are presented in figure 4. Clearly the situation this time is far more grave than previously. At higher frequencies, above 2 kHz it can be seen that the spectra differ mainly in terms of level, which is strongly related to the speaker distance from the listening position. However, below this frequency we can see some quite major differences that are associated with the room coupling. A full examination of these effects is not intended in this text, though these measurements are illustrative of the complexity and extent of the differences in amplitude response that also lead to differences in level alignments.

Another means of studying the differences between these two configurations is to look at the direct-to-reverberant energy ratio.

To do this the clarity index (40ms), as defined in equation 2, was calculated from measured impulse responses, for all of the difference channels and for both set-ups, the results of which can be found in table 1. The C_{40} measurement is presented as opposed to the more traditional C_{50} or C_{80} measurement, as it is more illustrative of differences in this highly damped environment.

$$C_{40} = 10 \log \frac{\int_0^{0.04} p^2(t) dt}{\int_{0.04}^{\infty} p^2(t) dt} \quad (\text{dB}) \quad (2)$$

where p is the acoustic pressure.

Table 1 Clarity indices for the symmetrical and asymmetrical loudspeaker set-ups illustrated in figures 1 & 3.

Channel	C_{40} (dB)	
	Symmetrical set-up	Asymmetrical set-up
Left	12.5	10.6
Center	12.8	12.6
Right	12.8	16.9
Left Surround	12.6	13.7
Right Surround	12.4	11.5

As we can see from this table, the C_{40} figures are quite constant for the symmetrical set-up. The asymmetrical set-up shows greater variance, as can be expected, with greater distance leading to lower values.

In more realistic domestic set-up, the acoustic symmetry of the listening environment may be more complex, with differing acoustic absorption properties associated with each wall. This will further aggravate the level alignment.

Both set-ups illustrated in figures 1 & 3 were employed in

a study of subjective alignment with different types of noise signals, which were reported in [27].

2.3 Perception

So far we have discussed purely the reproduction aspects of the multichannel reproduction scenario, whilst ignoring the presence of a listener in the final set-up. In this section we will briefly discuss the relationship between level and loudness and how these are evaluated.

Different models and metrics have been defined over the years to specify level and loudness [17, 18, 19]. Loudness can be defined a perceptual measure of level. Whilst linear SPL measures of level have been used to describe human perception for many years, this is perhaps not the most ideal method to describe loudness. Weighting curves as described in [12, 13] provide coarse approximation to the auditory systems response and have been found useful in certain applications, typically associated with noise emission. However, it should be noted that each of these weighting functions are designed to be correct only at specific loudness levels and in practice should not be applied beyond this scope.

Loudness models have existed for over three decades but are still little employed due to their relative complexity compared to linear SPL measures. In practice the complexity of loudness models today is quite trivial to implement in real time. The findings of various researchers [1, 2, 27], presented throughout this text suggests that the loudness metrics are very suited to the task for which they were intended and are superior to linear SPL measures.

The loudness models suggested by Paulus and Zwicker [19] and Moore et al [17] follow the basic function illustrated in Figure 5. The Zwicker and Moore models differ in a number of areas of which the most important are

- The characteristics of the transmission through the outer and middle ear
- Calculation of the excitation patterns
- Transformation of the excitation to a specific loudness scale

A detailed discussion of these models is not intended in this text and the interested reader is referred to the original papers on these topics for further information.

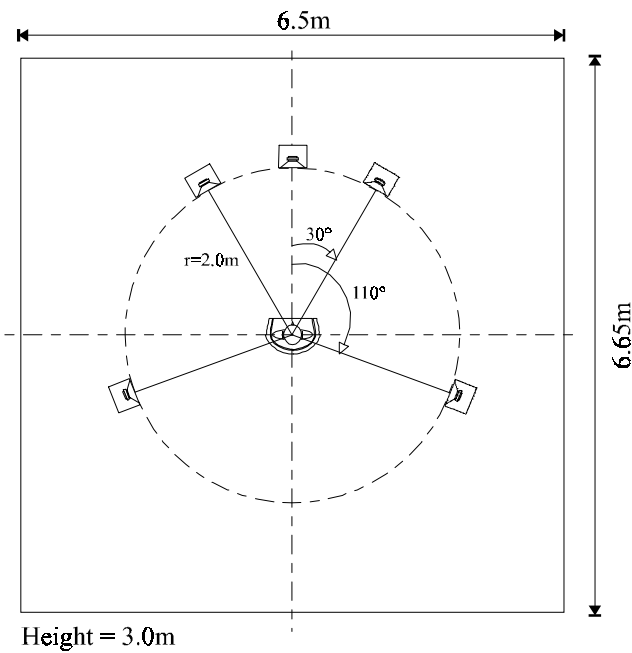


Figure 1 An idealised loudspeaker set-up [27] in accordance with ITU-R BS 775 [15]

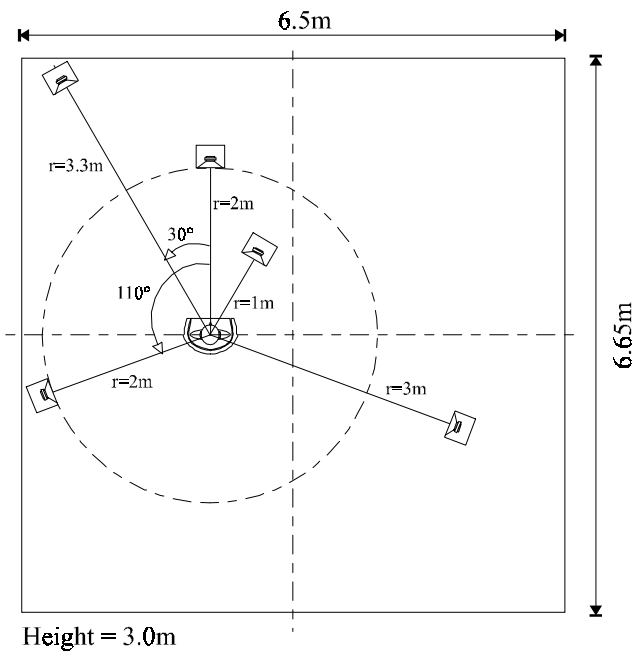


Figure 3 An asymmetrical loudspeaker set-up [27]

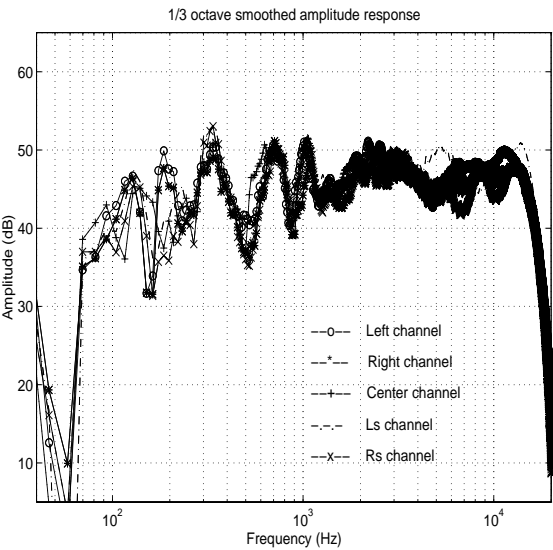


Figure 2 1/3 octave smoothed amplitude responses of the symmetrical set-up (figure 1) measured with a pressure microphone at the listening position

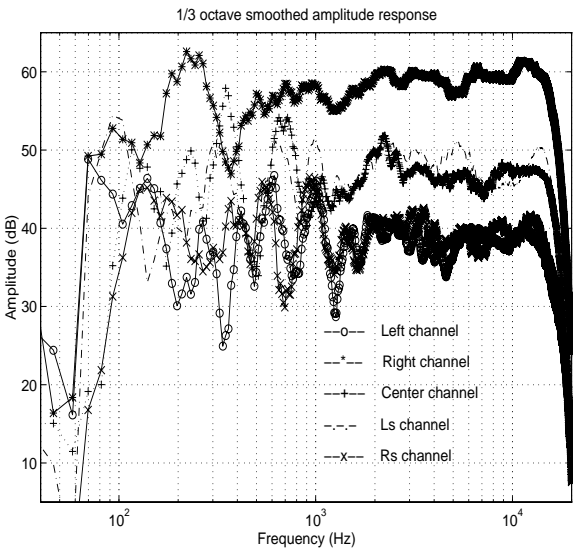


Figure 4 1/3 octave smoothed amplitude responses of the asymmetrical set-up (figure 3) measured with a pressure microphone at the listening position

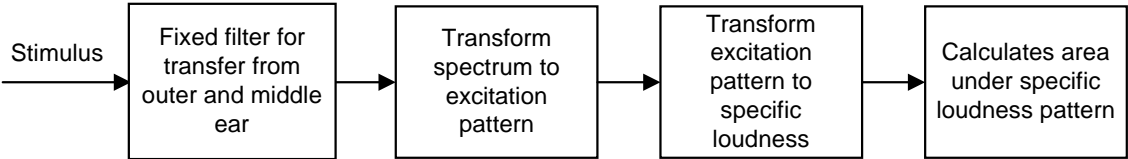


Figure 5 Block diagram of the Moore loudness model [17]

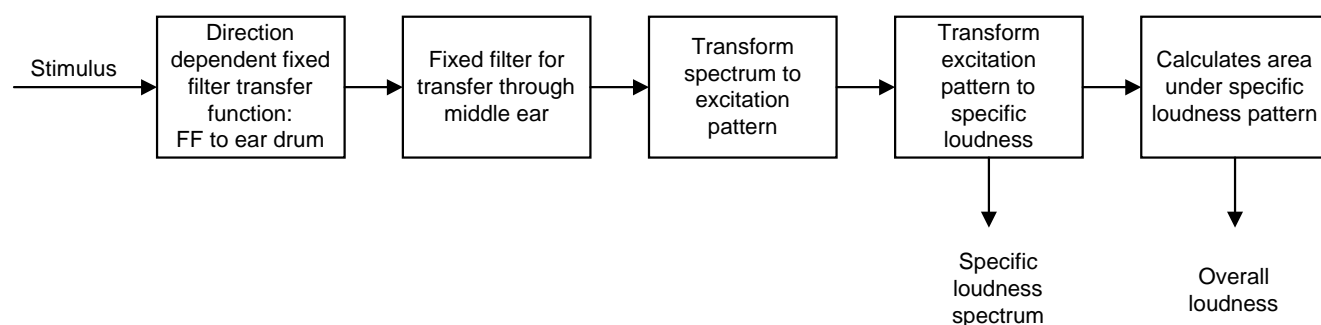


Figure 6 Block diagram adapted from the Moore loudness model [18]

3 NEW PERSPECTIVES

In this section we will discuss certain aspects which further influence level alignment in the multichannel scenario.

3.1 Directional loudness

In all listening situations, a person is present. Whilst for a stereo system, the sources of the sound are placed at equal angles on either side of a listeners, nearly symmetrical, head, this is not the case for the multichannel system. The question arises that what are the directional effects associated with the head and torso and how does this affect level alignment?

The issues of directional loudness have been studied to an extent by Robinson and Whittle [20] and more recently by Sørensen et al [22]. The findings of Robinson and Whittle were that there are significant directional interaural level differences (ILD) which show themselves mostly within the range 1.6-10kHz. These are principally caused by the physical nature of the head and pinna, providing directivity, causing shadowing, diffraction and other phenomena. However, at that time loudness models were not yet as well developed as today and so this data was not transformed to the loudness domain.

Sørensen et al, have also presented data on the directional level difference, which shows similarity to those presented in this paper.

In this study we choose to consider what the loudness is as a function source azimuth in free field conditions. To perform this study, head related transfer functions (HRTF) from a Brüel and Kjær head and torso simulator (type 4128) were employed in conjunction with the Moore loudness model [18]. This model consists of five stages to calculate the perceived loudness of steady state signals, the first of which is a free field to ear drum transfer function. In this model, the assumed direction of arrival of the source is 0° azimuth. For the purposes of this task, this block was omitted and the HRTF's of the head and torso simulator employed instead, which also includes the meatus (see figure 6). Furthermore, it has been assumed that the

binaural loudness of the 0° source is 20 Sones with a lower cut-off frequency of 50Hz, these being parameters employed in other studies by the author [27, 23]. Based upon these assumptions the specific loudness spectra have been plotted as a function of angle as illustrated in figures 7 and 8, employing a 0.3 ERB (equivalent rectangular bandwidth) grid.

The Moore model [18] states the binaural loudness as simply the summation of the monaural loudness levels for each ear. This method has been employed to estimate figure 9. The overall loudness has been calculated monaurally by summing the specific loudness (per ERB) for the whole ERB scale. This data is presented in table 3, the appendix.

When considering both the monaural and binaural loudness spectra, it is clear that below ERB 10 (~444 Hz), the interaural loudness difference (ILoD) is quite angle independent. The largest differences can be found in the midrange frequencies around ERB 25 (~3000 Hz), where monaural values range from 0.25-0.61 Sones and 0.5-1.21 Sones binaurally.

Overall monaural loudness differences vary in the range 5.2-10 Sones as a function of angle with a minimum at 110° for the left ear. This is quite a significant loudness difference and clearly perceptible. Naturally, this difference decreases when the overall binaural loudness is considered, varying in the range 16.6-20 Sones. Considering the typical set-up angles for 5 channel reproduction we can see from table 3, that the binaural levels vary quite considerably from 17-19-20 Sones for azimuth angles of 110°, 30° and 0° respectively.

In practice, multichannel systems are rarely set-up in free field conditions, in which case there will always be a diffuse field component as already illustrated. Furthermore, loudspeaker directivity affects the direct-to-reverberant ratio, as will be illustrated, which will also have an affect on the directional loudness characteristics. Under these circumstances it is suspected that the angular ILoD will be only marginal for broad band signals. For narrow band signals this may be another matter as binaurally loudness

differences are still quite significant.

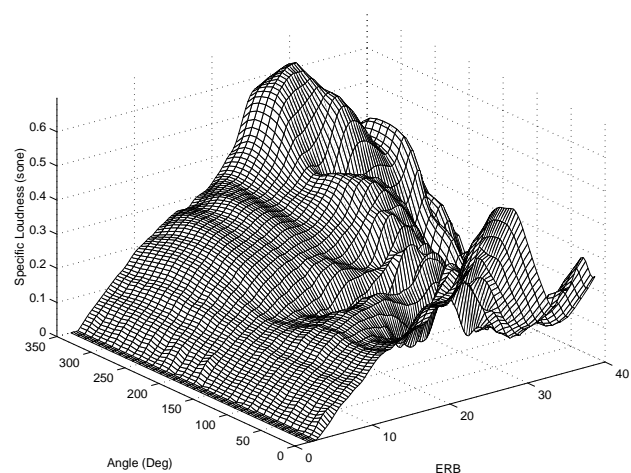


Figure 7 Monaural specific loudness spectrum (left ear) as a function of azimuth for a head and torso simulator

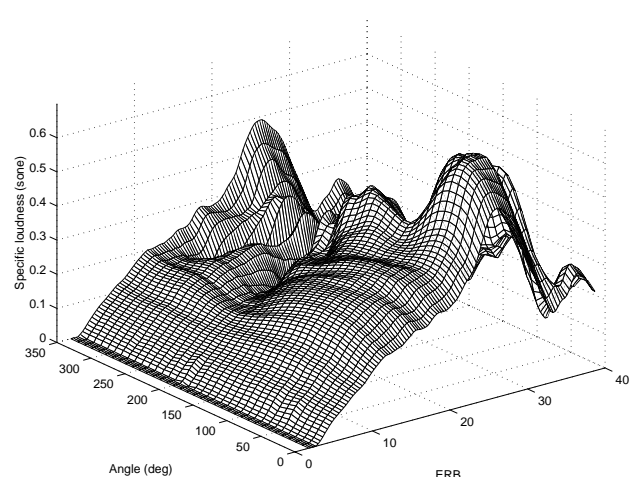


Figure 8 Monaural specific loudness spectrum (right ear) as a function of azimuth for a head and torso simulator

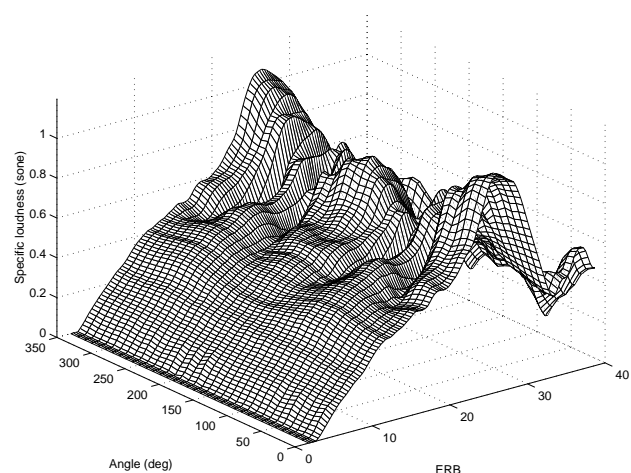


Figure 9 Binaural specific loudness spectrum as a function of azimuth for a head and torso simulator

3.2 Level alignment methods

A wide range of signals have been developed and employed over the years for calibration. In theory and under ideal conditions (i.e. a free field with identical channels and loudspeakers), alignment would be possible with a pure tone sine wave. In practice this is very unwise as matters are not so idealistic.

Broad band noise signals have been employed traditionally as a means of completely exciting the whole systems. Noise of various shapes have been developed for a wide variety of purposes and have often been named by colours. Whilst a discussion of the entire colour of noises is beyond the scope of this text, pink noises shall be discussed due to its relevance to this field.

Pink noise has been defined as a random noise signal having a spectral level, which decrease by 3 dB per doubling in frequency. This signal has been widely used in auditory research over the years. The motivation for this signal lies in the fact that when considered in terms of the 1/3 octave filtering, sound pressure level and logarithmic frequency, the spectrum is flat as illustrated in figure 10. Each of these metrics can be considered as simple approximation to those of the auditory system.

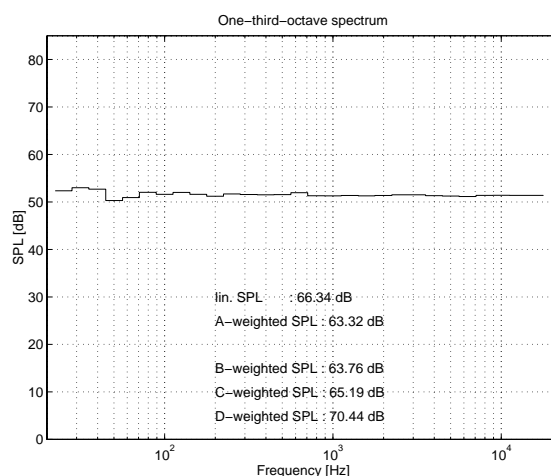


Figure 10 1/3 Octave spectrum of Pink Noise

Whilst pink noise has been the basis of many level alignment tasks, it has often been found non-ideal. One of the problems with the pure pink noise signal is that it places too much emphasis on the low frequency energy in comparison to the auditory system. Aarts [2] considered this matter and tested various weighting filters which are employed to approximate more closely to the auditory systems response. At that time A, B, C and D weighting filters were considered in measurement terms. Although A-weighted measurements have been considered elsewhere as a correct means of alignment [16], Aarts [1] did not found this to be the case for loudness alignment. In a further study of objective measured for loudness alignment, Aarts further concluded that the simple B-weighted SPL measure to be

more closely aligned with the Zwicker loudness measure [2].

Based upon this information Bech [3] employed a B-weighted pink noise signal to subjectively align systems with satisfactory results. As earlier concluded in loudspeaker directivity studies [27], the A-weighted pink noise measure was also found non-ideal and once again the B-weighted signal was found to provide a superior solution.

Whilst broad band signals have been employed by researchers in the field, commercial multichannel systems have tended towards more narrow band solutions. Bech [3], Suokuisma et al [23] and Zacharov et al [27] have reported and studied the use of certain commercially available narrow band test noise signals:

- Test signal A:
Filtered pink noise
Highpass: second order Butterworth with corner frequency of 700 Hz
Lowpass: first order Butterworth with corner frequency of 700 Hz
- Test signal B:
Filtered pink noise
Highpass: first order Butterworth with corner frequency of 500 Hz
Lowpass: first order Butterworth with corner frequency of 500 Hz
- Test signal C:
Filtered pink noise
Highpass: third order Butterworth with corner frequency of 2000 Hz
Lowpass: third order Butterworth with corner frequency of 500 Hz

Clearly these signals only excite a narrow portion of the frequency. The motivation behind all of these signals is not clear, but one of the signals has been developed for domestic multichannel calibration with the following aims in mind [8]:

- to avoid the low frequency variations between rooms occurring below the Schoeder frequency (approximately 500 Hz in domestic rooms),
- to minimise the position dependant effects in the sound field at higher frequencies (approximately 2000 Hz in domestic rooms),
- to provide a sufficiently broad frequency range signal to be representative of the loudspeakers output.

In practice these signals are well suited to in-situ midrange loudspeaker sensitivity alignment, in the frequency range where there are only small variations between rooms. However, these signals do not provide the user with a

broad band excitation that would be required to compensate for effects of the room interaction and source directivity.

What is the 'ideal calibration signal' is thus still an open question.

In an effort to establish how people perform multichannel level alignment, certain members of the Eureka 16553 Medusa (Multichannel Enhancement of Domestic User Stereo Applications) project group (Bech, Suokuisma and Zacharov), have commenced extensive studies into subjective and objective level multichannel alignment [23, 27].

Nine test signals have been considered in this work, as described in table 4. Several signals have been designed to take into account the characteristics of the auditory system and the source/room interaction in a detailed fashion.

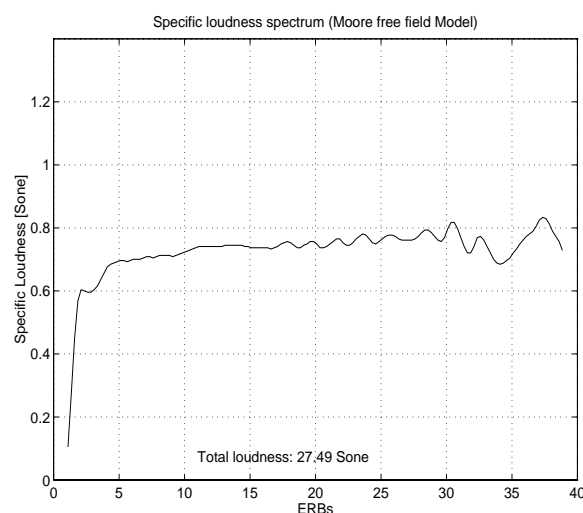


Figure 11 Example constant specific loudness signal [23] in accordance with the Moore model [17]. Note that total loudness here is according to the Moore model, which is equivalent 20 Sones (Zwicker diffuse field).

The constant specific loudness signal [23], in accordance with the Moore model[‡] [17] has been developed with the aim of performing level alignment over a broad frequency range. This is achieved by a level dependent spectral shaping of the signal that places equal perceptual weight on each frequency band (ERB). This type of strategy has been applied and tested for both the Zwicker and Moore models.

The initial experiment was performed in two sites based upon the set-up illustrated in figure 1. The task was to subjectively align the level of the test channel, of a 5 channel system, to that of the centre channel, employing the method of adjustment [5]. Loudspeakers were selected that were very closely matched. To ensure that the reference centre channel were equally loud for all signals,

[‡] The earlier Moore model was employed for this study, which differs from that presented in [18]

this channel was aligned for an equal loudness of 20 Sones with each signal, employing a Zwicker diffuse field model. A loudness alignment was essential in this case, due to the widely differing bandwidth and spectral characteristics of the signals under test. Other methods of center channel alignment were informally tested including linear, A, B, C and D weighted SPL measures and were found to provide very poor subjective level alignment in this case.

The experiment was performed with six trained subjects at each site in standardised listening rooms and analysed with a covariate analysis of variance model (ANCOVA). The findings of this work were that there are only marginal differences between the calibrations resulting with each signal. Though this was initially a surprise, when considering the similarity in the amplitude responses of the systems as shown in figure 2, the result is easier to understand. Small differences were found between the calibrations for different channels, which might be associated with the directional loudness characteristics HRTF's.

The experiment was repeated at one site with the asymmetrical set-up illustrated in figure 3 and analysed in a similar fashion. Once again the signal type was found to be only of marginal significance. Channel was found to be the dominating factor, which was both related to the distance of the source and the room interaction. A closer study indicated that in this case, listeners appeared to be compensating principally for level as a function of the source distance. However, it is unclear how the perception integration of the direct and diffuse level information occurs.

The conclusions of this work so far are that,

- there is a strong indication that for identical loudspeakers and idealised room acoustics, that the calibration signal characteristics are not significant,
- for the asymmetrical case, listeners are performing a level calibration based principally upon distance from the loudspeaker, which is a function of the room acoustics,
- in all cases the calibration signal was only found to be marginally significant.

3.3 Source directivity

Loudspeaker directivity is another complicating factor in the reproduction chain. Whilst the standards propose identical loudspeakers, domestic set-up often consider different directivity types [10, 6], particularly in the surround channels. The issues of spatial impression as a function of directivity has previously been studied [26] and found to have a profound influence on the perceived quality of spatial perception. During the pilot study of this experiment a wide range of domestic loudspeakers were studied and aligned employing the ITU-R BS 1116 alignment signal in accordance with equation 1. The study

compared different directivity loudspeakers for different groups of channels, such that all channels did not have identical directivity loudspeakers (bandwidths were similar). Initial informal subjective comparisons of these systems concluded that there were large differences in the front/back balance for different configurations employing this calibration procedure. Whilst set-ups employing loudspeakers of similar or identical directivity were well aligned with this methods, systems with dipole surrounds and more directivity frontal loudspeakers had an inferior alignment. Further informal listening tests were carried out with different alignment methods and the method proposed by Bech [3] was found to provide a superior calibration. This method, evolved from the finding of Aarts [2] proposed the use of a B-weighted pink noise signal feed to each channel and aligned for equal linear SPL (slow meter). For this study a calibration level 76 ± 0.2 dB (linear weighting, slow meter), was employed.

It became clear from this study that the loudspeaker directivity has a strong influence on the level alignment and the method of alignment must take this matter into account.

We have already seen that the distance from the source to loudspeaker has an influence on the direct-to-reverberant ratio. It is natural to assume that the directivity will also affect matters.

The clarity index (40ms) was measured for six different loudspeaker types in a BS1116 listening room. A detailed specification of these loudspeakers performance can be found in [26], which are best described as commercially available types. Speakers were placed at 2m from the centre of the listening room at 110° , as shown in figure 1. Speakers were calibrated with B-weighted pink noise to a level of 76dB (linear, slow meter) in accordance with [3], with a bandlimited frequency range of 110-18k Hz to ensure equal measurement bandwidth. Impulse response measurements were made with a calibrated pressure microphone for each speaker type from which the clarity indices were estimated. Results are presented in table 2 in rank order of clarity.

Table 2 Clarity indices for different directivity loudspeakers.

Speaker type	C_{40} (dB)
Dipole (null towards listener)	5.3
Dipole (lobe towards listener)	10.9
Pseudo omni-directional source	12.6
Cardioid	12.8
Horizontal line source	13.1
Vertical line source	13.3

Most of the speakers show small differences in the clarity index, with the exception of the dipole (null towards listener). In this configuration the clarity index falls to 5.3 dB, which is less than half the value for any of the other loudspeakers. This is of some concern, as this speaker

configuration often considered desirable in the surround channels.

The differences in clarity observed here are far larger than those associated with distance and placement as illustrated in table 1, in the same reproduction space. The question of how to perceptually compensate for the different room excitations associated with the source directivity is quite interesting. Once again it is clear that the perceptual integration of direct and diffuse level information is important in this respect.

Loudspeaker bandwidth is another factor that can affect matters. Whilst channels in 5.1 systems are capable of full bandwidth reproduction, this is not certain of the reproduction loudspeakers. This may become an issue in set-ups where different loudspeakers types are employed. Overall loudness is significantly affected by bandwidth, particularly at low frequencies. Thus if loudspeakers with limited low frequency performance are to be level aligned to wider band loudspeakers, problems may occur.

4 SUMMARY

This paper has considered some of the work and issues associated with the level alignment of multichannel systems. It is apparent that although many different strategies for calibration exist, few addresses the real problems associated with the widespread non-ideal multichannel set-up.

Based upon the studies presented in this paper it can be concluded that,

- level alignment is a critical factor in terms of perceptual quality,
- the ideal characteristics of a noise signals for level calibration are not yet known,
- source distance is a significant factor that influence level calibration,
- source directivity is a more significant factor that influence level calibration,
- directional loudness, though significant in the free field, may be less significant in the more reverberant domestic listening environment.

Clearly, at this time there are still some open questions as how to best align multichannel sound systems. Other researchers have showed the benefits of loudness alignment and B-weighted pink noise signals. However, research is still needed to consider the alignment signal requirement for non-ideal set-ups.

With the advent of virtual sound source technology for reproducing multichannel sound, an interesting question poses itself. How loud should virtual loudspeakers be and how should it be assessed and aligned?

Further work in this area should consider the effects of

loudspeaker directivity, bandwidth and absolute reproduction level as a function of different calibration signal.

5 ACKNOWLEDGEMENTS

The author would like to thank the funding body, Tekes (Technology Development Centre of Finland), for supporting the Eureka 1653 Medusa project. Pekka Suokuisma (Nokia Research Center) is thanked for assisting in the preparation of the directional loudness data. Matti Hämäläinen (Nokia Research Center) is thanked for his comments to drafts of this paper. All members of the Eureka Medusa project are thanked for their comments and discussion throughout the project so far.

6 REFERENCES

- [1] Aarts R. M., Calculation of the Loudness of Loudspeakers during Listening Tests, *Journal of the Audio Engineering Society*, Vol. 39, pp.27-38, January/February 1991.
- [2] Aarts R. M., Comparison of Some Loudness Measures for Loudspeaker Listening Tests, *Journal of the Audio Engineering Society*, Vol. 40, pp.142-146, March 1992.
- [3] Bech S., Calibration of relative level differences of a domestic multichannel sound reproduction system, *J. Audio Eng. Soc.*, vol. 46, pp 304 – 313, April 1998.
- [4] Bech S., The influence of stereophonic width on the perceived quality of and audio-visual presentation using a multichannel sound system, *J. Audio Eng. Soc.*, vol. 46, pp 314-322, April 1998.
- [5] Cardozo B. L., Adjusting the method of adjustment: SD vs DL, *J. Acoustical Society of America*, 37(5), May 1965.
- [6] Holman T., Audio for digital television, *Audio Media*, pp 114-118, April 1998.
- [7] Holman T., New factors in sound for cinema and television, *J. Audio Eng. Soc.*, vol. 39, no. 7/8, July/August 1991.
- [8] Holman T., personal communication, May 1998.
- [9] Holman T., *Sound for Film and Television*. Focal Press, Oxford and Boston, 1997.
- [10] http://www.thx.com/theatre_home/hth_dolby.html.
- [11] http://www.thx.com/theatre_home/hthx51.html.
- [12] IEC 537, Frequency weighting for measurement of aircraft noise (D-weighting), International Electrotechnical Commission, Geneva, Switzerland, 1976.

- [13] IEC 651, Sound level meters, International Electrotechnical Commission, Geneva, Switzerland, 1979.
- [14] ISO Rec. R. 532, Method for calculating Loudness Level, Method B, International Organization for Standardization, Geneva, Switzerland, 1966.
- [15] ITU-R Recommendation BS 775-1, Multichannel stereophonic sound system with and without accompanying picture, Geneva, 1994.
- [16] ITU-R Recommendation BS.1116, Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems, Geneva, 1994.
- [17] Moore B. C. J., Glasberg B. R., A revision of Zwicker's Loudness Model, *Acustica*, Vol.82, pp. 335-345, 1996.
- [18] Moore B. C. J., Glasberg B. R., and Baer T., A model for the prediction of thresholds, loudness, and partial loudness, *J. Audio Eng. Soc.*, vol. 45, pp 224 – 239, April 1997.
- [19] Paulus E., Zwicker E., Programme zur automatischen Bestimmung der Lautheit aus Terzpegeln oder Frequenzgruppenpegeln, *Acustica*, Vol. 27. pp. 253-266, 1972.
- [20] Robinson D. W., Whittle L. S., The loudness of directional sound fields, *Acustica*, vol. 10, pp. 74-80, 1960.
- [21] Rumsey F., Controlled subjective assessments of 2-to-5 channel surround sound processing algorithms, presented at the 104th Convention of the Audio Engineering Society, May 1998.
- [22] Sørensen M. F., Lydolf M., Frandsen P. C., Møller H., Directional dependence of loudness cues and binaural summation, proceedings of the 15th international congress on acoustics, Trondheim, Norway, pp. 293-296, June 1995.
- [23] Suokuisma P., Zacharov N., Bech S., Multichannel level alignment, part I: Signals and methods, presented at the 105th Convention of the Audio Engineering Society, September 1998.
- [24] Toole F. E., Subjective measurements of loudspeaker sound quality and listener performance, *J. Audio Eng. Soc.*, vol. 33, no. 1/2, January/February 1985.
- [25] Zacharov N., Bech S., Meares D., The use of subwoofers in the context of surround sound program reproduction, *J. Audio Eng. Soc.*, vol. 46, pp 276 – 287, April 1998.
- [26] Zacharov N., Subjective appraisal of loudspeaker directivity for multichannel reproduction, *J. Audio Eng. Soc.*, vol. 46, pp 288 – 303, April 1998.

- [27] Zacharov, N., Bech, S., and Suokuisma, P., Multichannel level alignment, part II: The influence of signals and loudspeaker placement, presented at the 105th Convention of the Audio Engineering Society, September 1998.

APPENDIX 1

Table 3 Overall loudness levels as a function of azimuth to the listener based upon the Moore model [18]

Azimuth	Monaural loudness, left ear (Sones)	Monaural loudness, Right ear (Sones)	Binaural loudness, two ears (Sones)
0°	10.0	10.0	20.0
10°	9.2	10.7	19.9
20°	8.3	11.3	19.6
30°	7.4	11.6	19.0
40°	6.6	11.9	18.5
50°	6.0	12.2	18.1
60°	5.5	12.3	17.8
70°	5.3	12.3	17.5
80°	5.4	12.3	17.8
90°	5.9	12.4	18.3
100°	5.4	12.2	17.5
110°	5.2	11.8	17.0
120°	5.4	11.3	16.7
130°	5.9	10.8	16.7
140°	6.5	10.3	16.8
150°	7.2	9.7	16.9
160°	7.8	9.3	17.0
170°	8.2	8.8	17.0
180°	8.5	8.5	17.0
190°	8.8	8.2	17.0
200°	9.2	7.8	17.0
210°	9.7	7.2	16.9
220°	10.2	6.6	16.7
230°	10.6	6.0	16.6
240°	11.1	5.5	16.6
250°	11.5	5.2	16.7
260°	11.8	5.2	17.0
270°	11.9	5.7	17.7
280°	11.9	5.4	17.3
290°	11.9	5.2	17.0
300°	12.0	5.5	17.5
310°	12.0	5.8	17.8
320°	11.8	6.4	18.3
330°	11.7	7.2	18.9
340°	11.4	8.1	19.5
350°	10.8	9.0	19.8

APPENDIX 2

Table 4 Description of test signals employed is [23, 27]

Signal name	High pass filter characteristics	Low pass filter characteristics	Comments
	Hz , dB/Oct.	Hz , dB/Oct.	
1.	700, 12	700, 6	Commercially available signal
2.	250, 6	500, 6	A signal
3.	500, 18	2k, 18	Commercially available signal
4.			Zwicker constant specific loudness according to ISO 532 (diffuse field)
5.			Zwicker constant specific loudness according to ISO 532 (free field)
6.			Constant specific loudness according to Moore
7.			Uniform excitation noise according to Zwicker
8.			Pink noise
9.			B-weighted pink noise