Using a spherical microphone array for stage acoustics: A preliminary case for a new spatial parameter

Lilyan Panton\(^{(a)}\), Densil Cabrera\(^{(b)}\), Damien Holloway\(^{(c)}\)

\(^{(a)}\) University of Tasmania, Australia, Lilyan.Panton@utas.edu.au
\(^{(b)}\) University of Sydney, Australia, Densil.Cabrera@sydney.edu.au
\(^{(c)}\) University of Tasmania, Australia, Damien.Holloway@utas.edu.au

Abstract

The acoustic conditions on stage for musicians are traditionally assessed with an omnidirectional receiver; however, with the use of a spherical microphone array the directionality of on-stage sound fields can be examined. This paper explores the issues around using such a microphone for stage acoustic measurements. As part of this study the 32-channel spherical microphone array Eigenmike has been used for acoustic measurements on-stage in six Australian auditoria; additionally, in four of these venues a traditional omnidirectional receiver was also used. This paper compares the results of standard omnidirectional parameters with the Eigenmike and omnidirectional receiver to assess the validity of the omnidirectional parameters derived from measurements with the Eigenmike. For example, the stage acoustic parameters the ‘support measures’ deviate between the microphones by no more than 0.5 dB. Additionally, the paper explores redefining the standard acoustic parameters to consider directionality, and presents these results in comparison to subjective musician assessments. A new parameter \(TS_{20–50}\) is proposed that corresponds well with the preferences of musician playing in ensemble. This work is being completed as part of larger study examining stage acoustics for chamber orchestras, which has also included subjective musician surveys with the Australian Chamber Orchestra regarding the venues included in the objective acoustic study.

Keywords: stage acoustics, higher order ambisonics,
Using a spherical microphone array for stage acoustics: A preliminary case for a new spatial parameter

1 Introduction

A growing field of research relates to the use of spherical microphone arrays for stage measurements in auditoria [1] [2]. These microphone arrays allow the directionality of the sound fields to be examined through a set of techniques that are often referred to as higher order ambisonics, or HOA. In this paper the spatial responses on stage are compared to subjective musician's assessment from the Australian Chamber Orchestra (ACO); details on the subjective musician surveying have been previously reported in [3].

In addition to using spherical microphone arrays to investigate the spatial response on stage, this work examines whether these microphone arrays may also be used for the measurements of the standard omnidirectional parameters in auditoria. In this paper measurements have been undertaken on stage in four Australian auditoria with a 32 channel spherical microphone array known as an Eigenmike 32, capable of resolving up to a 4th order ambisonic representation of the sound field, and are compared with a standard omnidirectional receiver (B&K 4190). The best known omnidirectional on-stage acoustic parameters are the 'support measures' (ST$_{\text{early}}$ and ST$_{\text{late}}$), which are measured with a 1 m source-receiver distance at several locations on stage and then arithmetically averaged. ST$_{\text{early}}$ is defined to assess the energy of reflections between 20–100 ms, whereas ST$_{\text{late}}$ is defined to assess the energy of reflections between 100–1000 ms. These measures were proposed by Gade [4], and are included in ISO 3382.1 [5]. ST$_{\text{early}}$ and ST$_{\text{late}}$ derived from the two types of microphone are presented in this paper. Additionally, sound strength $G$ and reverberation time $T_{30}$ on stage are derived from both microphones; $G$ and time $T_{30}$ are also defined in [5].

As well as comparing omnidirectional parameters derived from the two microphones, this paper proposes a procedure to consider spatial parameters defined to assess directionality of on-stage sound fields. This is motivated by a major study of on-stage acoustics for symphony orchestra musicians by Dammerud, which found that musicians had a preference for high and narrow stage enclosures, suggesting sound energy arriving from the sides before the top was preferred [6]. In a recent study using spherical microphone arrays on stage and musicians playing in the laboratory, Guthrie defined a spatial parameter to assess the ratio of sounds from ‘above’ to from the ‘sides’ and concluded that lower values of this parameter were preferred, again indicating sound from the sides, rather than above, was beneficial [7]. This paper discusses the issues around defining such a spatial acoustic parameter using 4th order ambisonics, and finally compares the results of a proposed spatial parameter $T_{S_{20-50}}$ with the subjective musician preferences.
2 Measurement procedure

In this study 1 m source-receiver distance measurements were made around four locations on stage, and across-stage measurements were also made with source-receiver distances between 2.7 m and 6 m. An omnidirectional source (B&K omnidirectional loudspeaker type 4295) was used; for the receiver a 32-channel spherical microphone array (Eigenmike 32) was used and in four auditoria an omnidirectional receiver (B&K omnidirectional receiver type 2669) was also used for comparison. The measurements were undertaken in unoccupied auditoria, and on stage with no musicians or stage furniture present. The auditoria in which the on-stage measurements were taken were all major Australian purpose built concert halls: City Recital Hall Angel Place, Sydney, (AP), Llewellyn Hall, Canberra, (LH), Sydney Opera House (SO) and Wollongong Town Hall (WH). Two additional venues were measured with only the Eigenmike: Perth Concert Hall (PH) and Adelaide Town Hall (AH).

The on-stage positions are shown in Figure 1; around each source position four 1 m measurements were taken (in positions front, back, left and right in relation to the stage orientation). Additionally, 12 across-stage measurements were undertaken between the four source positions. These measurements were processed using the AARAE software. AARAE is a MATLAB-hosted environment used to analyze room acoustics, for further details see [8].

3 Effect of microphone on omnidirectional parameters

The use of the Eigenmike was validated by comparing $ST$, $T_{30}$ and $G$. Results presented below are typical, though all other locations were considered, producing similar observations. In Table 1 the support measures ($ST$) are presented for source positions S1 and S2 defined in Figure 1; these results are the arithmetic average of the values in decibels from four measurements taken at a 1 m radius around each source position, and again averaged arithmetically over the 250–2000 Hz octave bands. The difference between the $ST$ measures determined by the Eigenmike and the B&K omnidirectional microphone is at worst 0.5 dB and generally within ±0.3 dB.
T_{30} derived from a 6 m source-receiver distance across stage (from S1 to S4) was investigated for octave bands 125–4000 Hz. The maximum deviation between the two microphones was 0.08 s. Such a small deviation is expected as the microphone type should not impact a reverberation parameter. This observed difference is also significantly less than differences that can be caused by the underlying calculation method used to calculate T_{30} [9].

Finally, deviations between microphones for the sound strength parameter G on stage were investigated. This is a more stringent test since it involves absolute rather than relative levels. For this analysis the B&K omnidirectional loudspeaker was calibrated in an anechoic chamber using the B&K omnidirectional microphone, according to [4]. Additionally, measurements in the anechoic chamber with the Eigenmike and omnidirectional microphone allowed a ‘transfer function’ between the Eigenmike and omnidirectional microphone to be produced, used to adjust the Eigenmike G values before comparing. The across stage measurements (source-receiver distance between 2.7 m and 6 m) and octave bands 125–4000 Hz were investigated. The deviations for G from each microphone were mostly less than 1 dB, with the absolute worst deviation for any source-receiver distance being 1.7 dB.

Table 1: Difference between ST measures with Eigenmike and omnidirectional microphone

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Microphone</th>
<th>250–2000 Hz Average (arithmetic)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AP</td>
</tr>
<tr>
<td>Pos S1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST_{early} (dB)</td>
<td>Omni</td>
<td>-11.64</td>
</tr>
<tr>
<td></td>
<td>Eigen</td>
<td>-11.33</td>
</tr>
<tr>
<td></td>
<td>Diff (Omni – Eigen)</td>
<td><strong>-0.31</strong></td>
</tr>
<tr>
<td>ST_{late} (dB)</td>
<td>Omni</td>
<td>-12.05</td>
</tr>
<tr>
<td></td>
<td>Eigen</td>
<td>-11.74</td>
</tr>
<tr>
<td></td>
<td>Diff (Omni – Eigen)</td>
<td><strong>-0.32</strong></td>
</tr>
<tr>
<td>Pos S2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ST_{early} (dB)</td>
<td>Omni</td>
<td>-13.57</td>
</tr>
<tr>
<td></td>
<td>Diff (Omni – Eigen)</td>
<td><strong>0.21</strong></td>
</tr>
<tr>
<td>ST_{late} (dB)</td>
<td>Omni</td>
<td>-13.78</td>
</tr>
<tr>
<td></td>
<td>Eigen</td>
<td>-13.91</td>
</tr>
<tr>
<td></td>
<td>Diff (Omni – Eigen)</td>
<td><strong>0.13</strong></td>
</tr>
</tbody>
</table>

4 Spatially defined stage parameters

In this section we introduce higher order ambisonics (HOA) and discuss the issues around defining spatial acoustic parameters on stage. Then a spatially defined acoustic parameter is presented to assess sound energy from ‘above’ relative to the ‘sides’ on stage, and discussed compared to musicians’ preferences regarding six auditoria.

4.1 Background to Higher Order Ambisonics

Any arbitrary spatial directionality function s(\theta, \phi) can be represented as a weighted sum of any infinite set of orthogonal basis functions
where \( c_i(\theta, \phi) \) are the basis functions (in the higher order ambisonic form these are the spherical harmonics \( Y_n^m(\theta, \phi) \) [10]) and \( w_j \) are the weights. We multiply by \( c_i(\theta, \phi) \), integrate over the unit sphere and use the orthogonality condition \( \sum_{j=1}^{\infty} \int \int c_i(\theta, \phi) c_j(\theta, \phi) \, dA = 0 \) for \( i \neq j \) to obtain the spherical Fourier transform

\[
\int \int c_i(\theta, \phi) s(\theta, \phi) \, dA = \sum_{j=1}^{\infty} \int \int w_j c_i(\theta, \phi) c_j(\theta, \phi) \, dA = w_i \int \int c_i^2(\theta, \phi) \, dA,
\]

which easily yields the weights \( w_i \) of harmonic component \( i \) by rearrangement. In this implementation the functions \( Y_n^m(\theta, \phi) \) are normalised such that the area averaged value is 1, hence the last integral is \( 4\pi \), the area of the unit sphere. We have two tasks: to convert the 32 microphone signals into spherical harmonic channels; and to find the optimum combination of these channels to capture and reject sound respectively originating from a desired solid angle its complement. The two are closely related problems solved very similarly, though the former must account for the scattering of the incident waves by the rigid sphere that approximates the microphone, and is already implemented in AARAE [8]. The current work implements the latter problem, and the following provides background.

In discrete form, if we define the surface of the sphere by a large number \( N \) points each representing as nearly as possible equal area, the integrals above may be approximated as sums, hence equation (2) becomes

\[
w_i = \frac{\sum_{k=1}^{N} c_{k,i} s_k}{\sum_{k=1}^{N} c_{k,i}^2} = \frac{1}{N} \sum_{k=1}^{N} c_{k,i} s_k,
\]

where \( c_{k,i} \) and \( s_k \) are respectively \( c_i(\theta, \phi) \) and \( s(\theta, \phi) \) evaluated at the \( k^{th} \) point with coordinates \( (\theta_k, \phi_k) \), and again the denominator has been simplified to \( N \) by the normalization of \( c_i(\theta, \phi) \). If we truncate the infinite series to represent the function \( s_k \) with finite order ambisonic components (hence a finite number of basis functions and weights) then this can be expressed in matrix form as

\[
\{ w \} = \frac{1}{N} [c]^T \{ s \},
\]
where columns of \([c]\) represent each basis function evaluated at all points. We note that an alternative form based on the second integral of (2) is \(\{w\} = ([c]^T[c])^{-1} [c]^T \{s\}\), (where \([c]^T[c] = N[I]\) due to the orthogonality and normalization of \(c_l(\theta, \phi)\)). If \(\{s\}\) is longer than \(\{w\}\) this is simply the least squares solution to the overspecified problem \(\{s\} = [c]\{w\}\). Thus in intuitive terms the weights are those that provide the best fit of the spherical harmonics of a given order to the desired spatial directionality function. Finally, in order to capture or reject sound in a given region we define \(s_k = 1\) or \(0\) respectively. The definition of the desired regions, corresponding functions \(\{s\}\), and corresponding weights \(\{w\}\), will be discussed below.

\[\]

\[\]

**Figure 2:** a) Cartesian and spherical coordinates used in the analysis, courtesy of [11]. The x axis points towards the audience (front); \(\theta\) is defined as the angle towards the y axis from the x axis, and \(\phi\) is defined as the angle downwards from the z axis towards the xy plane. b) The spatial regions are defined as the faces of a cube centered at the origin with edges parallel to the x, y and z axes projected onto the \((\theta, \phi)\) space; left, right, front, back, top and bottom are defined as shown.

### 4.2 Definitions of spatial regions and spatial filtering procedure

For the purposes of defining spatial acoustic parameters spatial regions have been proposed. The regions are designated as ‘top’, ‘bottom’, ‘left’, ‘right’, ‘front’ and ‘back’. Sound arriving from either ‘left’ or ‘right’ directions is defined as ‘sides’. These regions are defined as the solid angles subtended by the faces of a cube centred at the origin and aligned with the x, y and z axes shown in Figure 2a. For example, if sound arrives through the front of the cube it is assigned to the ‘front region’. The cube is oriented on stage such that the front faces towards the audience and the top faces towards the stage ceiling. Using spherical coordinates (defined in Figure 2a), the regions can be projected onto the \((\theta, \phi)\) space, as shown in Figure 2b.

Spatial filtering is performed to capture the sound energy arriving from each region. To achieve the desired spatial filtering first the 32 physical channels corresponding to the individual microphones must be decoded into the ambisonic channels. For \(N^{th}\) order there will be \((N+1)^2\) ambisonic channels and this number must be less than the number of microphones in order to obtain a unique ambisonic decoding. Thus with the Eigenmike we are limited to a 4\(^{th}\) order HOA...
representation with 25 channels. By combining these 25 channels with appropriate weights as determined by the method outlined in Section 4.1 energy from any desired region (a single direction or a solid angle) can be captured. With infinite order this can be achieved perfectly, but if the series is truncated to a finite order the representation will of course be an approximation. Appropriate weighing must be applied to the HOA channels to give as closely as possible a unit signal strength within the desired region and zero signal strength outside of the desired region.

![Image of spatial filtering](image)

**Figure 3:** a) The ideal spatial filtering for the right direction where in desired region weighting is one and outside desired region weighting is zero. b) The actual spatial filter for the right direction, based on 4th order ambisonics.

To design the spatial filter first the ideal case was specified by the signal weight vector \( \{s\} \) to represent complete signal rejection from all the undesired regions \((s_k = 0)\), but complete signal acceptance within the desired region \((s_k = 1)\); this was done at 3002 points very nearly equally spaced over the sphere, and is demonstrated in Figure 3a for the ‘right’ region. The weights \( \{w\} \) were determined according to the procedure in section 4.1 to create the best spatial filter achievable using 4th order ambisonics (25 HOA channels), as shown in Figure 3b. By the reciprocity principle a hypothetical sound field of the form of Figure 3a will be detected in 4th order ambisonic form as in Figure 3b, i.e. it will not be detected uniformly over the desired region and there will be ‘leakage’ as sound appears to come from adjacent regions. We can quantify the latter artefact by comparing the energy from each undesired region relative to energy in the desired region. This was found by integrating over the solid angle corresponding to each cube face (left, right, front, back, top and bottom) for the surfaces shown in Figure 3b, taking the ratios, and converting to decibels. As an example if we are examining the ‘left’ spatial filter and we wish to assess how much energy will be captured originating from the front region (relative to from the left region) we would use

\[
E_{\text{front-left}} = 10 \log(s_{\text{front}}^2) - 10 \log(s_{\text{left}}^2),
\]  

(5)
where $s^2_{\text{front}}$ is the integral of the squared $s$ values in region ‘front’, and $s^2_{\text{left}}$ is the integral of the squared $s$ values in region ‘left’. In a more generic sense we could call this $E_{a-b}$ where $a = \text{any region being assessed}$, and $b = \text{desired region}$. We find that $E_{a-b} = -17$ dB if $a$ and $b$ are adjacent regions, or $-30$ dB if $a$ and $b$ are opposite.

4.3 Defining a Top/Sides spatial parameter

The use of spherical microphone arrays, and higher order ambisonics, to investigate stage acoustics is a relatively new field of study. However, some previous work has been undertaken to adapt existing stage parameters to ‘directional’ versions. Guthrie redefined many of the common stage acoustic parameters (e.g. the $ST$ measures, $LQ_{7-40}$, $G_p$, $G_l$) based on directional regions, or alternatively as ratios of directional regions [6]. However, Guthrie used different definitions for spatial regions to those used in this paper, and the spherical microphone array used in that study was limited to 2nd order ambisonics. Guthrie found a spatial ratio of $LQ_{7-40}$ from Top/Sides was relevant for musician playing in ensemble conditions (decreasing values were preferred). A similar parameter has been investigated in this work, but with some changes to the integration limits.

The lower integration time limit of 20 ms was selected; this was based on an investigation of the on-stage impulse responses which found this was an appropriate time to remove the direct sound and floor reflection, but was before the occurrence of any of the stage enclosure reflections. It was crucial to remove the direct sound as it is significantly higher in magnitude than the stage reflections and thus if it was included any differences in enclosure reflections between stages could not easily be observed. The upper integration time limit of 50 ms was selected to capture ‘early’ reflections. Traditionally 100 ms has been used as the upper limit for the arrival of early reflections (such as in the support parameters proposed by [3]). However, based on work by Guthrie and Dammerud it appears very early arriving reflections (i.e. before 40 or 50 ms) are highly relevant for ensemble playing conditions [1] [5]; this also reflects that we are interested in whether first order reflections from ‘sides’ arrive before or after those from the ‘top’ region. We thus propose a spatial ratio of ‘top’ relative to ‘sides’ $TS_{20-50}$ defined as

$$TS_{20-50} = 10\log\frac{\int_{20ms}^{50ms} p^2(t)_{\text{TOP}} \cdot dt}{\int_{20ms}^{50ms} p^2(t)_{\text{SIDES}} \cdot dt}.$$  (6)

4.4 Measurement procedure and results for Top/Sides on stage

The $TS_{20-50}$ parameter was derived from the on-stage 1 m source-receiver measurements. The parameter was found as an average value of the four measurements around the source. In Figure 4a the results are presented for source position S1; the venues are plotted from left to right in terms of the musicians’ most preferred to least preferred venue (i.e. PH is the most preferred venue and WH is the least preferred venue). We see that the trend of decreasing
TS_{20–50} being preferred is clearly observed in our dataset. The same results for source position S3 are shown in Figure 4b, and similar trends were observed for the average of the 1 m measurements at S2 and S4. The method of averaging 1 m measurements on stage is in line with procedure used for the ST parameters; however, it may also be of interest to investigate the ‘across stage’ measurements (with larger source-receiver distances) when considering ensemble playing. This will be presented in future work, as well as full statistical significance testing, but preliminary results confirm the above findings.

![Graphs showing TS_{20–50} for different venues](image)

Figure 4: Average TS_{20–50} (dB) for the four 1 m source-receiver distances for 250 – 2000 Hz octave bands a) around source position S1 b) around source position S3; venues are listed in Section 2.

5 Discussion

This paper has provided a comparison of results for standard omnidirectional stage parameters with a spherical microphone array and with a standard omnidirectional receiver. This has shown generally good agreement between the two datasets, particularly for the ST parameters, suggesting the use of an omnidirectional receiver in addition to a spherical microphone array is not necessary. The advantages of using a spherical microphone array became clear in Section 4, when spatial acoustic parameters could be investigated. A parameter TS_{20–50} is proposed representing a Top/Sides ratio and could be derived from on-stage measurements. This is an extension of the work by Dammerud based on physical dimensions of stages to assess energy from sides relative to energy from above [5]. From these initial results it is evident that lower values of Top/Sides are preferred; indicating early reflections arriving from the ‘sides’ rather than from the ‘top’ (before 50 ms) are preferred for musicians playing in ensemble. This project is ongoing and the subjective relevance of such a parameter for musicians will be further investigated with more on-stage acoustic measurements, and musician surveying.
6 Conclusions

In this paper the use of a spherical microphone array for deriving standard omnidirectional parameters has been investigated. Additionally, a potentially useful spatial parameter to assess sound energy arriving from above compared to from the sides has been proposed and investigated. It was found that decreasing values of the proposed parameter $T S_{20-50}$ may be preferred for musicians playing in ensemble. Further musician surveying and on-stage measurements are planned to confirm this trend.

Acknowledgments

Lily Panton is supported by an Australian Postgraduate Award, with additional support from the Dr Joan Woodberry Fellowship in Engineering and the Georgina Sweet Fellowship. Lily Panton received an ICA-ASA Young Scientist Conference Travel Grant. This research is also supported by an Australian Acoustical Society Education Grant. This research is also supported under the Australian Research Council's Discovery Projects funding scheme (project DP120100484).

References


