# 3D impulse response measurements of spaces using an inexpensive microphone array

Daniel Protheroe (daniel.protheroe@marshallday.co.nz)
Marshall Day Acoustics
PO Box 5811, Wellesley St
Auckland 1141, New Zealand

Bernard Guillemin (bj.guillemin@auckland.ac.nz)
Dept. of Electrical and Computer Engineering, Faculty of Engineering
The University of Auckland
Private Bag 92019, Auckland Mail Centre
Auckland 1142, New Zealand

## ABSTRACT

The acoustical characteristics of a room are traditionally determined using omnidirectional impulse response measurements, yielding information about sound reflections in terms of magnitude and time, but not direction. However, the direction of reflections is often important, and thus the need for a practical, low cost measurement system for determining this. In this paper we present the performance of a low cost measurement system utilising an inexpensive microphone array, namely the Core Sound TetraMic, for the determination of 3D room impulse responses. These can then be visualised, for example, as a "hedgehog pattern". Experiments undertaken in an anechoic chamber indicate that the accuracy of directional estimation of this system is in the region of ±7.5°.

## 1   INTRODUCTION

In the field of room acoustics the measurement and visualisation of room impulse responses is of great interest. Room impulse responses are traditionally measured using an omnidirectional receiver. Sound reflections can therefore be analysed in terms of magnitude and time, but not direction. A 3D impulse response measurement system is required for the complete analysis of a sound field in terms of magnitude, time and direction of reflections. It then becomes possible, for example, to relate sound reflections to the physical features of the room, observe the directional distribution of early and late sound energy, as well as identify surfaces causing problematic reflections.

Several researchers have made 3D impulse response measurements, for example Abdou and Guy,[1] Merimaa et al.,[2] Fukushima et al.,[3] Ohta et al.,[4] Tervo et al.,[5] and Bassuet.[6] However, their measurement systems often use proprietary, expensive or relatively impractical hardware. Further, details of the necessary calibration routines, as well as their practical limitations, are often not readily available. The consequence of this is that their use is essentially limited to the researchers who developed them.

In this work we propose a 3D impulse response measurement system which uses low cost, commercially available hardware, namely the Core Sound TetraMic microphone array,[7] and

investigate its performance. The necessary calibration routines for the TetraMic are software based and come supplied by the manufacturer. The subsequent processing routines leading to a fully directional analysis and visualisation of a sound field are relatively straightforward to implement and are presented in this paper. The system should be highly accessible to practitioners in the acoustics arena, as well as being relatively simple to operate in a real-world environment.

This paper is structured as follows. In Section 2 sound intensity theory is presented. Our measurement system is discussed in detail in Section 3, followed by some experimental results of the system in an anechoic chamber and a real room in Section 4. Our conclusions follow in Section 5.

## 2  SOUND INTENSITY THEORY

The system resolves magnitude and direction of reflections using a sound intensity technique. Sound intensity, $\mathbf{I}$, at a position in a sound field is defined as the product of instantaneous sound pressure $p(t)$ and the instantaneous 3D particle velocity vector $\mathbf{u}(t)$ at that location:

$$\mathbf{I} = \overline{p(t)\mathbf{u}(t)} \tag{1}$$

where the average is made over a suitable length of time.

One approach for measuring sound intensity is with a B-format microphone which comprises a coincident array of transducers which can measure pressure and particle velocity in three dimensions.

## 3  MEASUREMENT SYSTEM OVERVIEW

The major functional blocks of the measurement system are shown in Fig. 1. First, the 3D impulse response is recorded at a position in the room. The resulting signals are segmented into short time windows of interest, and for each the time averaged sound intensity is estimated, producing a series of 3D vectors which describe arriving sound energy in terms of magnitude, direction and time. Finally, these vectors are visualised as 2D or 3D "hedgehog patterns".
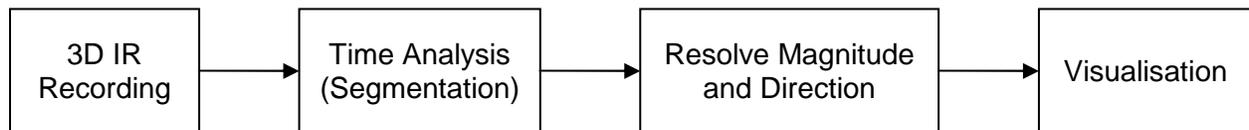
| 3D IR Recording | → | Time Analysis (Segmentation) | → | Resolve Magnitude and Direction | → | Visualisation |
|---|---|---|---|---|---|---|

**Figure 1:** Major functional blocks of the measurement system.

### 3.1  3D Impulse Response Recording

Capturing a room's intensity impulse response in 3D in B-format presents us with the challenge of arranging a pressure sensor and three orthogonal velocity sensors at the same point in space. One practical method to capture B-format, developed by Gerzon[8] in the 1970s, uses a tetrahedral microphone array of cardioid capsules. The outputs of this A-format array are summed and filtered to obtain the equivalent B-format signals. Our measurement system utilises a Core Sound TetraMic which is a relatively inexpensive and commercially available A-format microphone array. The manufacturer calibrates each device and supplies this data, as well as A-format to B-format conversion routines, in the form of software.

### 3.2 Time Analysis

The pressure signal from the B-format room impulse responses is analysed in time to identify important events including the direct sound, early reflections and any problematic echoes. The transition time between the early reflection period and late reverberation is also determined. This information guides a segmentation process whereby the B-format impulse response is divided into segments, with each segment ideally encapsulating an event. This is necessary because the direct sound and reflections are not delta functions but have dispersed energy. This is as a result of band-limited and non-flat frequency responses inherent in the measurement system and the room under test, as discussed in Section 3.4.

#### 3.2.1 *Direct Sound and Early Reflections*

The magnitude, direction and timing of early reflections are important in concert hall acoustics and should ideally be analysed individually. Fig. 2 shows the results of the segmentation process applied to the early part of a pressure impulse response signal. The shaded areas correspond to rectangular windows which encapsulate the direct sound (blue) and three prominent early reflections (red). This process could be conducted manually based on visual inspection, or automated through a peak-picking or arrival detection algorithm,[9,10] an example of which is discussed in Section 3.5.
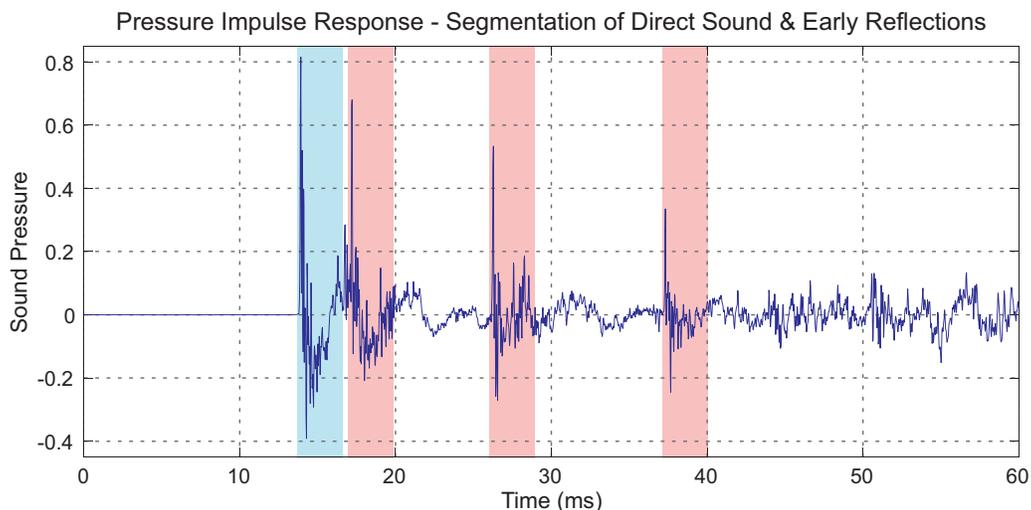


**Figure 2:** A pressure impulse response showing the segmentation of the direct sound (shaded blue) and early reflections (shaded red).

#### 3.2.2 *Late Reverberation*

After the so-called transition time,[11] from the end of the early reflection period to the onset of the late reverberation period, it is no longer possible to consider reflections individually, and there is no value – from the point of view of the subjective sound – in doing so. In this region of the room response the signal is segmented into a series of equal length, non-overlapping rectangular windows. The magnitudes and directions found from these windows will indicate the diffuse nature of the field.

#### 3.2.3 *Window Duration*

For both the early reflection and later periods, the window durations used should be equal in order to accurately compare magnitudes within a non-steady-state signal, but it is not

necessarily obvious what window length (nor the precise window function) will be optimal. Our experience so far suggests that a window length based on the duration of the direct sound is suitable. This ensures good timing resolution, but will be at the expense of some low frequency information in the resulting analysis. A longer window is more likely to encapsulate more than one reflection in the early period.

### 3.3   Determination of Magnitude and Direction

Magnitude and direction is estimated for each of the time segments according to the following process. For each segment sound intensity is determined in the X, Y and Z directions. The time averaged intensity for the $i$th segment $S_i$ in the X direction is given by:

$$I_{S_i,X} = \frac{1}{N}\sum_{n=0}^{N-1} I_X[S_{i,TOA} + n] = \frac{1}{N}\sum_{n=0}^{N-1} p[S_{i,TOA} + n]u_X[S_{i,TOA} + n] \qquad (2)$$

where $S_{i,TOA}$ is its time of arrival and $N$ its length (both in samples), $p[n]$ is the pressure signal, and $u_X[n]$ is the velocity signal in the X direction. Similar calculations are performed in the Y and Z directions resulting in the vector:

$$\mathbf{I}_{S_i} = \begin{bmatrix} I_{S_i,X} \\ I_{S_i,Y} \\ I_{S_i,Z} \end{bmatrix} \qquad (3)$$

The length and direction of $\mathbf{I}_{S_i}$ correspond to the estimates of magnitude and direction, respectively, of the sound in segment $S_i$.

The pressure and velocity signals may also be filtered if particular frequency bands are of interest. It is recommended that signals are windowed first before filtering to avoid filter delay problems. Each window should be sufficiently zero padded to include energy delayed by the filters.[12]

As will be discussed in Section 4.1, the individual transducers of a TetraMic don't appear to act as coincident in space for frequencies above about 5 kHz. This introduces error in the direction estimation process, but this can be avoided by filtering. In a broadband analysis, direction should be estimated from B-format segments which have been processed through a low pass filter with a cut-off frequency of 5 kHz. However, the sound intensity magnitudes should be estimated from the unfiltered signals.

### 3.4   Impact of Non-Ideal Frequency Responses

In the example shown in Fig. 2, the direct sound and early reflections are well separated in time. However, in many situations this is not the case, and identifying individual events in the early reflection period can be challenging. The imperfect (i.e. non-flat) frequency responses inherent in the measurement system and the reflecting surfaces in the room will cause energy to smear out in time. For example, the impulse response of a dodecahedron loudspeaker, a source often used in room acoustic measurements, is shown in Fig. 3. Clearly this deviates significantly from a delta function, as therefore will be the case for any resulting reflections that arise. This smearing of energy makes it difficult to identify individual early reflections, especially when these are overlapping. One approach of potential help in identifying reflections in this case has been suggested by Lee.[13]
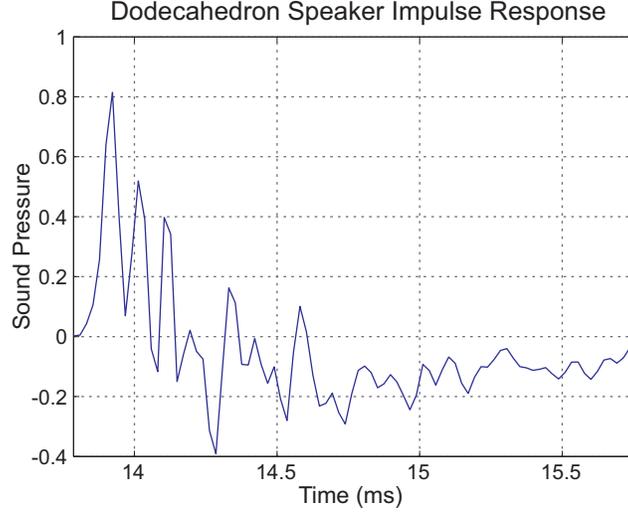
**Figure 3:** Impulse response of a dodecahedron speaker.

### 3.5 Lee's Correlation Technique

This assumes that each early reflection is a relatively undistorted version of the direct sound. The idea is to cross-correlate the entire impulse response signal with the windowed direct sound.

Let $p[n]$ be the sampled pressure waveform and $p_{Direct}[n]$ be the segment identified as the direct sound, zero padded to equal the length of $p[n]$. The cross correlation between $p_{Direct}[n]$ and $p[n]$ is given by:

$$R_{pp_{Direct}}[m] = \sum_{n=0}^{N-1} p[n+m]p_{Direct}[n] \tag{4}$$

where $N$ is the length of $p[n]$. Provided that the early reflections exhibit a high degree of similarity with the direct sound, $R_{pp_{Direct}}[m]$ should exhibit large, positively valued peaks at each arrival. Fig. 4(a) shows a hypothetical impulse response consisting of a direct sound and one reflection. Fig. 4(b) shows the resulting cross correlation function (note that the correlation delay has been shifted in time) with single, positive peaks at the position of the direct sound and reflection.

The correlation technique may be applied in 3D, also suggested by Lee.[13] The three particle velocity impulse responses are correlated with $p_{Direct}[n]$:

$$\mathbf{R}_{up_{Direct}}[m] = \sum_{n=0}^{N-1} \mathbf{u}[n+m]p_{Direct}[n] \tag{5}$$

where $\mathbf{R}_{up_{Direct}}[m]$ is a 3D vector quantity and $\mathbf{u}[n]$ is the instantaneous 3D particle velocity vector. This correlation function is effectively the sliding time-averaged sound intensity. In an analysis of the early reflection period, arrival times may be estimated by locating the peaks in $R_{pp_{Direct}}[m]$. The value of each peak, in each of the X, Y and Z signals of $\mathbf{R}_{up_{Direct}}[m]$, forms a 3D vector which points in the direction of the incoming arrival. Magnitude information will be incorrect, as the direct sound part of the pressure signal is used instead of the actual pressure. This alternative method of resolving direction may be particularly valuable in situations where

reflections are overlapping in time, an example of which is shown in Section 4.2. The B-format impulse responses should be filtered to remove high frequency non-coincidence effects before calculating $\mathbf{R}_{up_{Direct}}[m]$. A linear phase filter is recommended so that $\mathbf{R}_{up_{Direct}}[m]$ can be time aligned with $R_{pp_{Direct}}[m]$ after filtering.
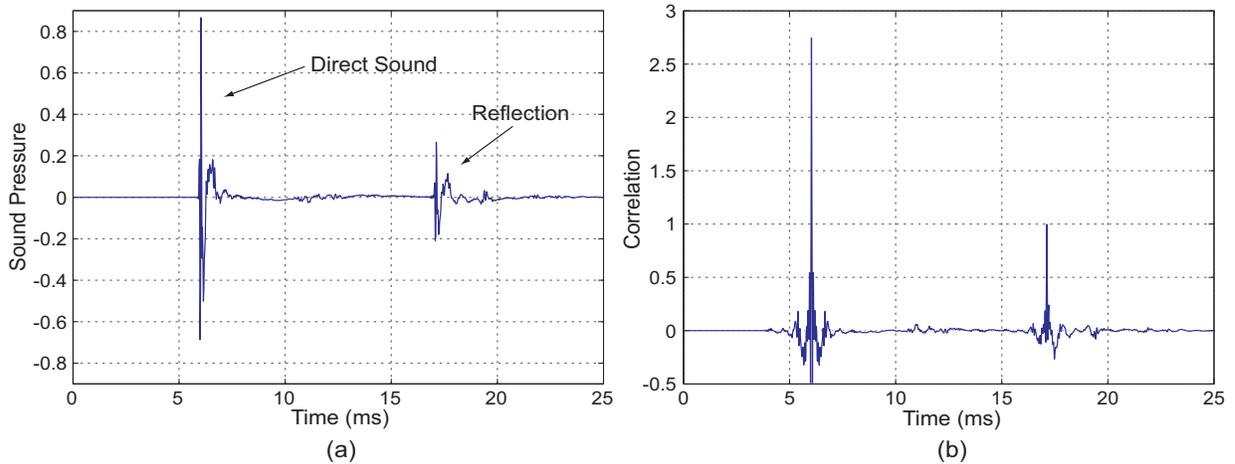


**Figure 4:** (a) Pressure impulse response showing the direct sound and a reflection, (b) Cross correlation of the impulse response with the direct sound segment.

## 3.6   Visualisation of 3D Impulse Response

The method presented here for viewing the sound intensity vectors generates a 2D or 3D "hedgehog" image in a similar style to Thiele's "Igel"[14] and Bassuet's[6] 3D impulse response visualisation. An example is shown in Fig. 5. Each sound intensity vector or segment is represented by a single line, plotted on a 3D Cartesian diagram with respect to the measurement position (i.e. the origin). The length of each line corresponds to the relative sound intensity level (dB) normalised to the direct sound magnitude, and its direction is the estimated direction of incoming energy flow. The time information is represented by colour, as described by the key in Fig. 5. The direct sound is coloured light green and thicker than the other vectors. The dynamic range is restricted to 40 dB. The clarity of the presentation can be enhanced by rotating the view. A 2D hedgehog pattern is useful for relating the sound intensity data to plans of the room, as shown in Section 4.3.

## 4   EXPERIMENTAL RESULTS AND DISCUSSION

In this section we present the results from two experiments to demonstrate the performance of the system in an anechoic chamber, and then the results of measurements conducted in a real room (the Auckland Town Hall auditorium). Farina's exponential sine sweep technique[15] was used to excite the rooms, and recordings were conducted at 24 bit resolution and a sampling rate of 44.1 kHz.

## 4.1   Investigation of Directional Accuracy

Some of our preliminary experiments indicated that the angular estimation error present in the system was a function of direction. Thus this error was investigated for arrivals in the horizontal plane. In the anechoic chamber the source, a Tapco S5 studio monitor, was setup 1.6 m from the TetraMic mounted on a turntable, as shown in Fig. 6. A total of 72 impulse response

measurements were taken in 5° azimuth increments of the turntable from 0° to 360° ($\theta_D$), and for each, the azimuth of the direct sound was estimated ($\hat{\theta}_D$).
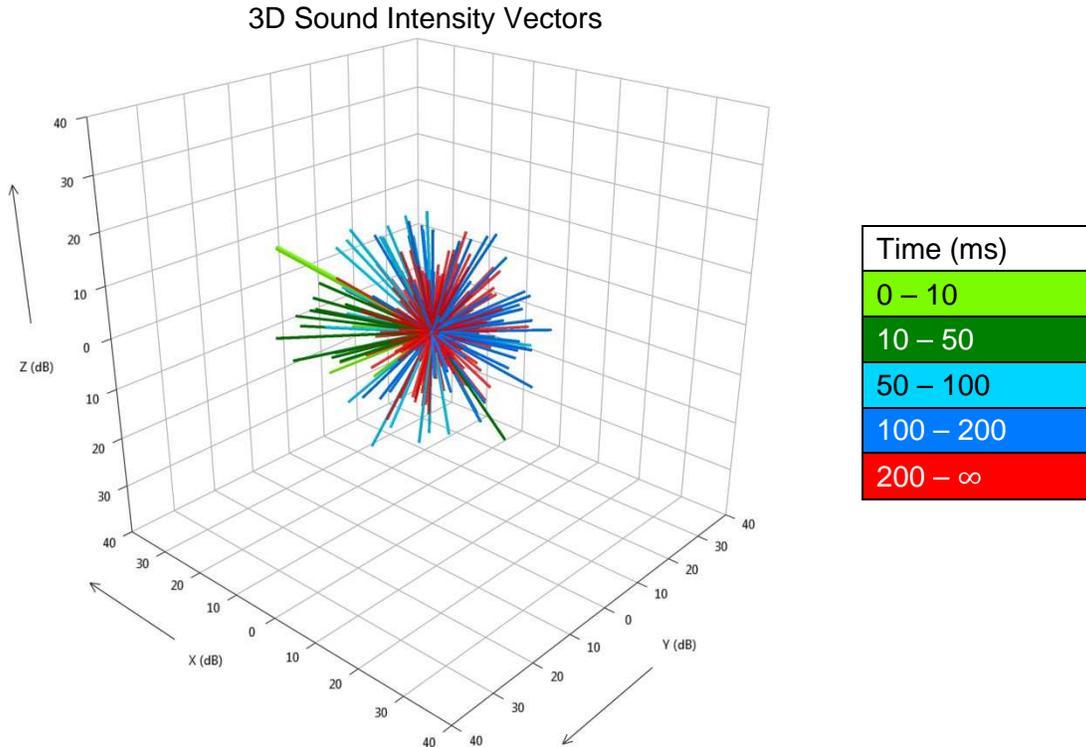
### 3D Sound Intensity Vectors



**Figure 5:** 3D sound intensity vectors visualised as a "hedgehog pattern".

The resulting broadband azimuth estimation error $\theta_E = \hat{\theta}_D - \theta_D$ is shown plotted in blue in Fig. 7. An average error of 5° was subtracted from $\theta_E$ before plotting the diagram. This offset implies a slight rotational misalignment of the TetraMic in the physical setup. The maximum error (without filtering) was $\pm 18.8°$. The smoothness and repeating pattern of this error function suggests it can be corrected for. Two causes for this error are likely to be the non-coincidence for high frequencies, and physical alignment errors in the experimental setup.

At high frequencies above about 5 kHz the TetraMic begins to no longer appear coincident, distorting the B-format polar patterns and thus introducing directional error. This error should be predictable and correctable, but further investigation is required in order to determine how this might be done. A simpler approach is to filter out this high frequency data. The angular error, after applying such a non-coincidence filter, is shown by the red trace in Fig. 7. In this case, a Butterworth low pass filter with a cut-off frequency of 5 kHz was used. The maximum error after filtering is $\pm 7.5°$, which represents a substantial improvement over the unfiltered results.

Errors in the physical setup of this experiment would have also contributed to this error. For example, the TetraMic was unintentionally slightly lower in height than the centre of the source and azimuth estimation is dependent on the elevation angle. This could be partly responsible for the cyclic shape of the error. Further, the source was not a true point source but a loudspeaker enclosure consisting of two separated drivers. The TetraMic was also displaced slightly in the horizontal plane as it was rotated. Careful alignment would be expected to reduce the errors for horizontally arriving sound.
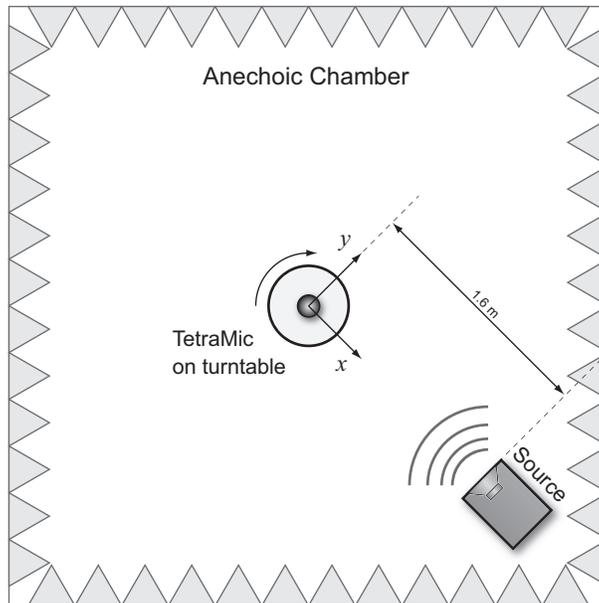
**Figure 6:** Experimental setup in the anechoic chamber for investigating directional accuracy in the horizontal plane.
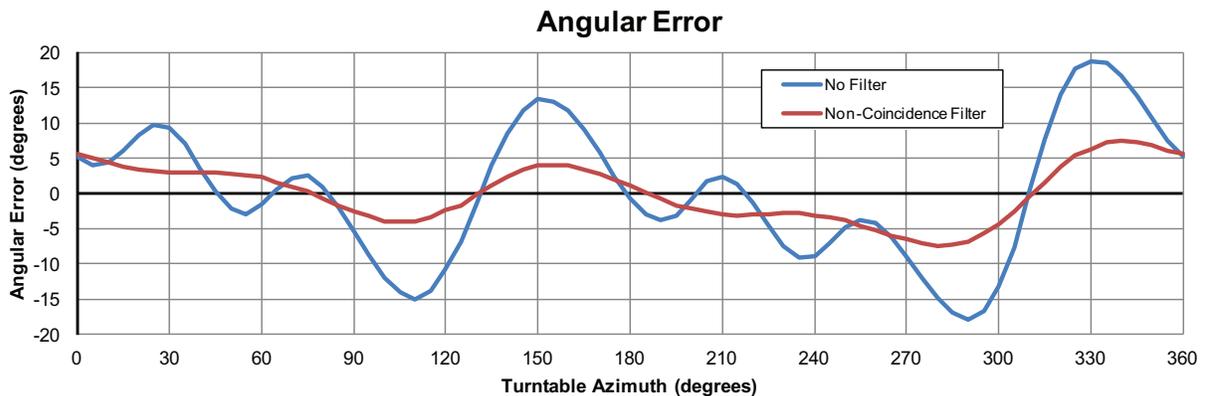


**Figure 7:** Azimuth error for unfiltered signals (blue) and signals filtered with a non-coincidence filter (red).

This experiment presented the error for horizontally arriving sound, but in reality this should be examined in all directions. Additional investigation and modeling may lead to further improvements in the overall accuracy of the measurement system.

## 4.2   Overlapping Reflections

An experiment was carried out to investigate the effect of two overlapping reflections arriving from different directions, and whether or not their original directions could be estimated acceptably accurately. Again this was undertaken in the anechoic chamber (see Fig. 8). The positions of the source (Tapco S5) and TetraMic were the same as in the previous experiment. The positive X direction of the TetraMic was directed towards the source. Two reflectors of identical dimensions were arranged symmetrically about the TetraMic to generate two equal and opposite sound reflections.
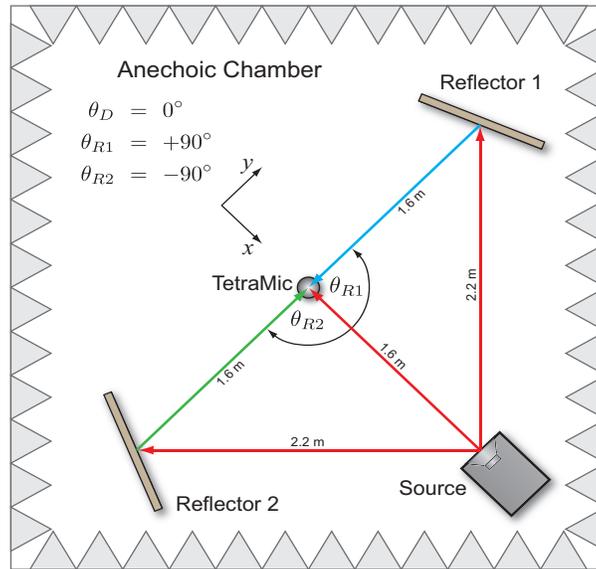
**Figure 8:** Setup for the overlapping reflections experiment.

The recorded impulse response for this arrangement is shown in Fig. 9(a). As expected, it appears that only one reflection is present and thus it is concluded that the two reflections are highly overlapped. The estimated azimuth for these combined reflections is 136°.

When analysed using Lee's correlation technique, the part of the correlation function $R_{pp_{Direct}}[m]$ relating to the reflections yields two distinct peaks as shown in Fig. 9(b). From analysing the corresponding peaks in $\mathbf{R}_{up_{Direct}}[m]$, the estimated azimuth of the first peak was -93° and the second peak was 93°, from which it can be concluded that the reflection resulting from Reflector 2 arrived slightly before Reflector 1. So in this case, even though these reflections were highly overlapped, their azimuths were resolved relatively accurately. There is, however, no clear way of estimating the magnitude of each reflection in this case.

It would thus seem appropriate to use this correlation technique to identify the time and direction of arrivals in the early reflection period where reflections might be overlapping. However, it has been our experience that the usefulness of this technique is limited in a real world environment. Often the cross correlation signal would look equally as confusing as the original impulse response, and the significance of each peak would not be clear. One issue is that the impulse response of a dodecahedron speaker varies with direction. Correlating signals with a template radiated from only one direction of the source will therefore be erroneous. Another important issue is the band-limited responses of reflecting surfaces in the room which can cause significant time smearing of energy in the early reflection period.

Further work to improve the robustness of the correlation technique is required. Other methods to assist with identification of arrivals in an impulse response should also be investigated. Sturm and Defrance,[9] and Defrance et al.[11] developed an arrival detection technique based on matching pursuit. This is similar to the correlation technique presented here, but with some additional steps. Usher[10] presented a technique based on a kurtosis analysis. Tervo et al.[5] suggested detecting reflections based on spherical variance.
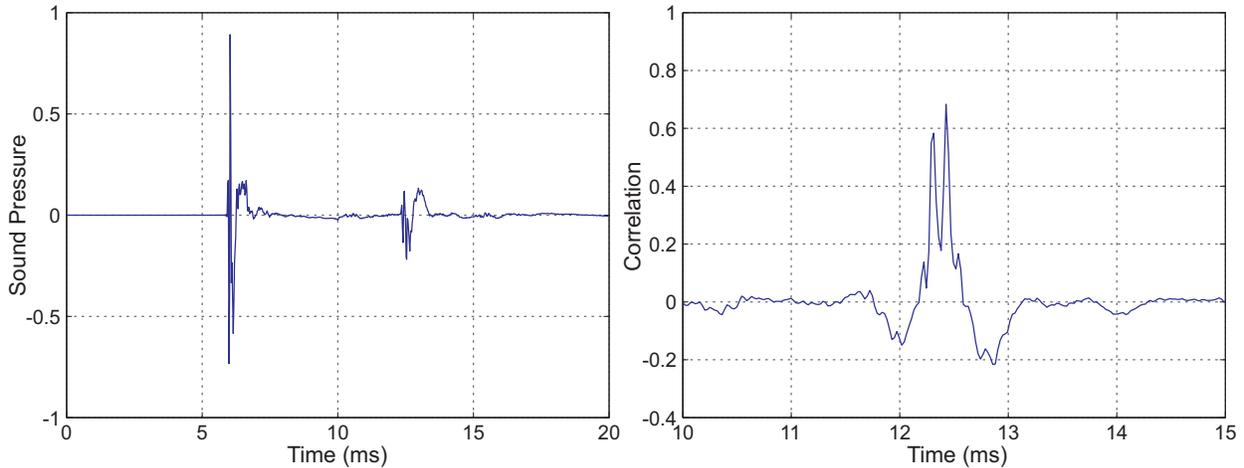
**Figure 9:** (a) Pressure impulse response of the overlapping reflections experiment, (b) Part of the cross-correlation signal showing two reflection peaks.

When the impulse response in the early reflection period is not sufficiently clear to allow aligning of windows with individual reflections, this period can be segmented into a series of equal length rectangular windows, similar to the analysis of the late reverberation period (see Section 3.2.2). Segments will not correspond to individual reflections, but a series of segments in the early period should reliably indicate how energy is directionally distributed. This technique is also very simple to automate. The impulse response measurements presented in the next section were analysed in this way.

### 4.3   Measurements in a Concert Hall

The measurement system was used to obtain 3D impulse responses in the Great Hall of the Auckland Town Hall, a traditional shoebox concert hall which seats about 1500 people across three levels. The source, a dodecahedron loudspeaker, was placed at the centre of the stage, approximately 2 metres back from the front and at a height of 1.5 m above the stage floor. The receiver positions, shown in Fig. 10, have the resulting hedgehog diagrams overlaid. The TetraMic was placed at listener head height, i.e. 1.2 m above the floor.

At each position, a 30 second sweep signal was used to excite the room. The impulse responses were segmented into a series of 2 ms rectangular windows following the onset of the direct sound. In each of the resulting broadband hedgehog diagrams, the direct sound segment is assigned a magnitude of 40 dB, and the magnitudes of the subsequent sound intensity vectors are referenced to this level. As a result, magnitude information cannot easily be compared between different hedgehog diagrams. The detected reflections have been grouped into several bands of time as shown by the different colours in Fig. 10.

Considering the hedgehog diagram at receiver position R1, in the stalls, the thicker light green line, pointing to the source, is the direct sound intensity vector. The other light green lines may be related to direct sound arriving from other drivers in the dodecahedron loudspeaker, or floor reflections. The next interval of time, dark green, corresponds to the early reflection period. The directions of most vectors in this interval imply laterally reflected energy from the side walls. As time progresses (light blue to red) the sound intensity vectors gradually become more uniformly distributed about the measurement position, indicating the sound field is becoming diffuse.
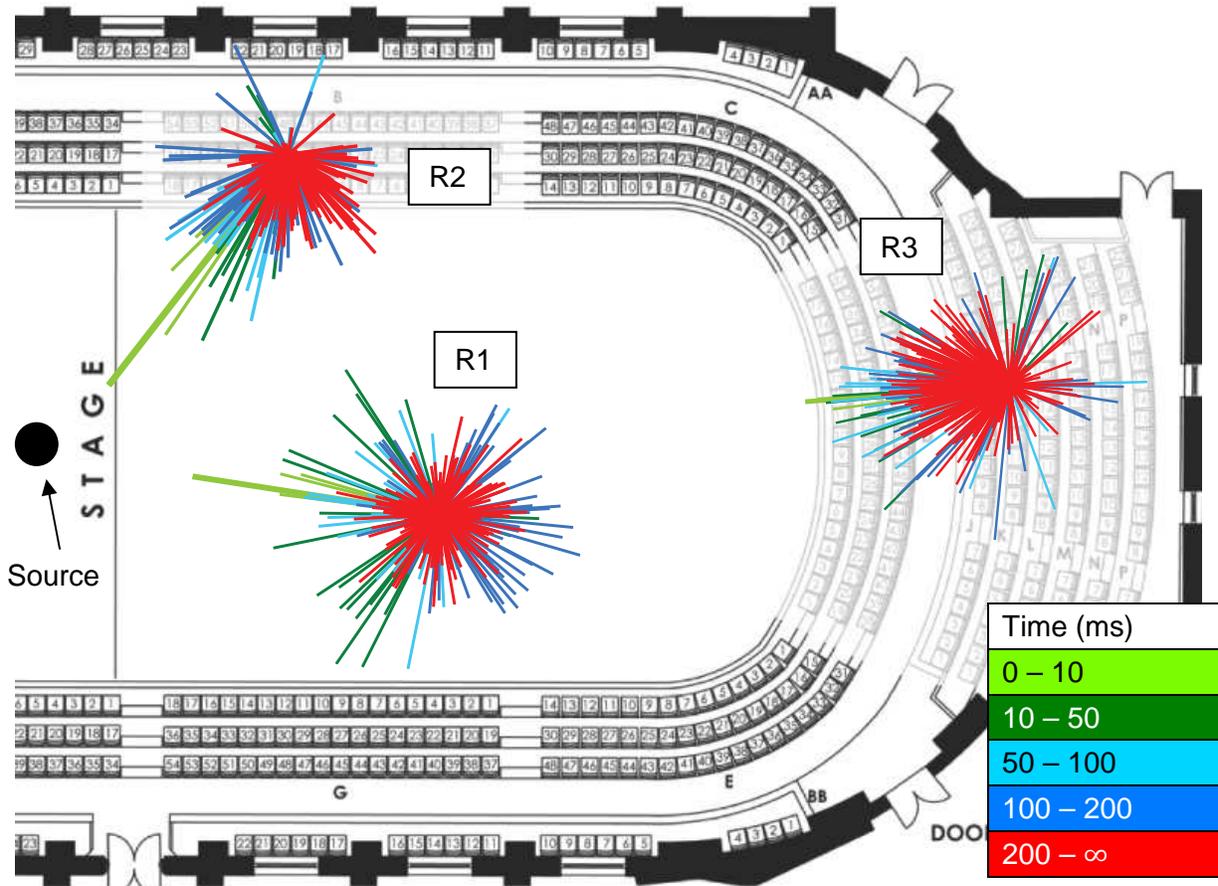
**Figure 10:** 2D hedgehog patterns for three receiver positions in the stalls and balcony levels of the Great Hall.

At position R2, in the balcony level, the direct sound (light green) is quite prominent, followed by some early energy (dark green) in a similar direction to the direct sound. The sound field doesn't become as diffuse as in position R1, indicated by the lack of red sound intensity vectors in the top half of the hedgehog. This is consistent with having rear shielding and absorption from the raked padded seats located in this seating area, and the large area of curtains along the balcony walls.

Position R3, also in the balcony, is relatively far from the source and is overshadowed by the ceiling of the second balcony. As a result we see the length of the direct sound vector is similar to those of the reflection vectors. This indicates a smaller direct to reverberant ratio compared to the other positions. As with position R2, the steep raked seating behind R3 absorbs a significant amount of energy in these directions.

These results are well in agreement with what is expected from the form and contents of the room, and the hedgehog style of visualisation is, we feel, an effective way of comparing the acoustical properties at the different seating positions.

## 5    CONCLUSIONS

This paper has overviewed a technique for measuring and visualising 3D impulse responses of rooms using inexpensive and commercially available hardware, specifically the Core Sound TetraMic. The measurement system was tested in an anechoic chamber to assess its practical

limitations, and it was found that directional accuracies in the horizontal plane are in the range of $\pm7.5°$. A technique based on cross-correlation was also described to assist with discriminating arrivals in the early reflection period, but improvements to the robustness of this technique are required. Finally, the measurement system was demonstrated in a real space, the Auckland Town Hall auditorium, and the resulting visualisations show detail which is well matched to the physical environment.

## ACKNOWLEDGEMENTS

## REFERENCES

[1] A. Abdou and R. W. Guy, "Spatial information of sound fields for room-acoustics evaluation and diagnosis," J. Acoust. Soc. Am. **100** (5), 3215-3226 (1996).

[2] J. Merimaa, T.Lokki, T. Peltonen, and M. Karjalainen, "Measurement, analysis and visualization of directional room responses," in *AES 111th Convention*, New York (2001).

[3] Y. Fukushima, H. Suzuki, and A. Omoto, "Visualization of reflected sound in enclosed space by sound intensity measurement," Acoust. Sci. Tech. **27** (3), 187-189 (2006).

[4] T. Ohta, H. Yano, S. Yokoyama, and H. Tachibana, "Sound source localization by 3-d sound intensity measurement using a 6 channel microphone system – part 2: application in room acoustics," in *Proceedings of Inter-Noise 2008*, Shanghai, China (2008).

[5] S. Tervo, T. Korhonen, and T. Lokki, "Estimation of reflections from impulse responses," in *Proceedings of the International Symposium on Room Acoustics,* (2010).

[6] A. Bassuet, "New acoustical parameters and visualization techniques to analyze the spatial distribution of sound in music spaces." J. Build. Acoust. **18** (3-4), 329-347 (2011).

[7] L. Moskowitz, *Core Sound – TetraMic*, http://www.core-sound.com/TetraMic/1.php

[8] M. A. Gerzon, "The design of precisely coincident microphone arrays for stereo and surround sound," in *AES 50th Convention*, London, UK (1975).

[9] B. L. Sturm, and G. Defrance, "Detection and estimation of arrivals in room impulse responses by greedy sparse approximation," in *18th European Signal Processing Conference (EUSIPCO)*, Aalborg, Denmark (2010).

[10] J. Usher, "An improved method to determine the onset timings of reflections in an acoustic impulse response," J. Acoust. Soc. Am. **127** (4), EL172-EL177 (2010).

[11] G. Defrance, L. Daudet, and J.-D. Polack, "Using matching pursuit for estimating mixing time within room impulse responses," Acta Acust. United Ac. **95** (2009), 1071-1081 (2009).

[12] ISO 3382-1:2009; *Acoustics – measurement of room acoustic parameters – Part 1: Performance spaces* (International Organization for Standardization).

[13] R. Lee (private communication).

[14] E. Meyer and R. Thiele, "Raumakustische untersuchungen in zahlreichen konzertsälen und rundfunkstudios unter anwendung neuerer messverfahren," Acustica **6**, 425-444 (1956).

[15] A. Farina, "Simultaneous measurement of impulse response and distortion with a swept-sine technique," in *AES 108th Convention,* (2000).