# LOW FREQUENCY PHASE CALIBRATION FOR A CIRCULAR MICROPHONE ARRAY

*Paul D. Teal*

Victoria University of Wellington
P.O. Box 600, Wellington, New Zealand
Email: paul.teal@vuw.ac.nz

*Mark A. Poletti*

Industrial Research Limited
P.O. Box 31-310, Lower Hutt, New Zealand
Email: m.poletti@irl.cri.nz

Previous work has indicated that a limitation on the low frequency performance of a circular microphone array for holographic sound field recording is phase mismatch between the microphones in the array. At low frequencies these variations become more significant than at mid-range and high frequencies because the high order phase mode responses are lower in amplitude. This paper investigates the possibility of performing a "self" calibration of a microphone array. The basis of the calibration is to estimate the location of one or more sources using mid-range frequencies and to use this source location information to perform correction to the array at low frequencies. This of course implies that the calibration is performed in a relatively anechoic environment, since multipath effects at widely differing frequencies are uncorrelated. Initial results confirm that a significant improvement in the array response is possible using this approach.

***Index Terms—*** Microphones, Acoustic Arrays, Calibration, Phase Estimation

## 1. INTRODUCTION

Microphone arrays offer the possibility of recording the sound at not just a single point, but an entire sound field [1, 2]. Most methods of storage and subsequent reproduction of the sound field information involve transforming the microphone data into spherical or circular harmonics. Circular microphones are particularly suited to recording of two-dimensional sound fields, as they allow recording of the horizontal spherical harmonics with fewer microphones than required for a three-dimensional array [3].

Previous work [4] has examined construction of polar responses from a sound field by the use of a regular circular microphone array. It was shown that for an $N$ element array, the $m$th term of the sampled phase mode response to a plane wave from azimuth angle $\phi_i$ is given by

$$r_m(k,\phi_i) = \frac{1}{N}\sum_{n=0}^{N-1} z_n(k,\phi_i-\phi_n)e^{-jm\phi_n} \tag{1}$$

where $z_n(k,\phi_i-\phi_n)$ is the response of microphone $n$ located at angle $\phi_n$ to a plane wave of wave number $k$ with angle of arrival $\phi_i$.

This response can be expressed in terms of its Fourier series decomposition, as

$$z_n(k,\phi_i) = \sum_{q=-\infty}^{\infty} a_n(k,q)e^{-jq\phi_i}$$

$$\Updownarrow$$

$$a_n(k,q) = \frac{1}{2\pi}\int_0^{2\pi} z_n(k,\phi_i)e^{jq\phi_i}\,d\phi_i \tag{2}$$

If all the microphone elements are identical, then $a_n(k,q) = a(k,q)$, and so the sample response becomes

$$\begin{aligned} r_m(k,\phi_i) &= \frac{1}{N}\sum_{n=0}^{N-1}\sum_{q=-\infty}^{\infty} a(k,q)e^{-jq(\phi_i-\phi_n)}e^{-jm\phi_n} \\ &= \sum_{q=-\infty}^{\infty} a(k,q)e^{-jq\phi_i}\left[\frac{1}{N}\sum_{n=0}^{N-1} e^{j(q-m)\phi_n}\right] \\ &= \sum_{l=-\infty}^{\infty} a(k,m-lN)e^{-j(m-lN)\phi_i}. \end{aligned} \tag{3}$$

The last line follows from the fact that the microphone elements are equally spaced ($\phi_n = 2\pi n/N$) and so the term in square brackets is equal to 1 for $q = m - lN$ for any integer $l$, and is zero otherwise. If $a(k,m-lN)$ is small for any values of $l$ other than zero, we then obtain the important result that *the low order polar responses of the array are independent of the polar responses of the individual microphones.* For most microphone patterns this will be true up to some frequency. Beyond this frequency, we say that aliasing has occurred. The complex value $a(k,0)$ will in general vary depending on the frequency, and so equalisation across frequency will be required.

The number of modes we are able to accurately estimate from the measured data is an important measure of the effectiveness of the array for holographic sound field recording. This in effect means that we wish the approximation $a(k,m-lN) \ll a(k,0)$ to be true for as large a value of $m$ as possible. However, we have the conflicting requirement that $a(k,m)$ must be large for as large a value of $m$ as possible.

A prototype array was constructed using 16 cardioid microphones, and the complex polar response of each measured as the array was rotated through 360°.

The conflicting requirements are clearly illustrated in Fig. 1. At high frequencies (e.g., 8000 kHz), the response is not small for $m - lN = \pm 16$. This results in aliasing for all modes. At low frequencies (e.g., 125 Hz) the response is small for $m - lN = \pm 16$, and so the zeroth order mode ($m = 0$) can be reliably constructed. However, for first order terms ($m - lN = \pm 1$), the response is quite small; -7 dB compared to the zeroth order. The second order term is smaller still at -30 dB, and so higher order modes are considerably attenuated compared to those at mid-range frequencies.

Example omnidirectional, first order and second order responses are shown in Fig. 2. The aliasing at high frequencies is very apparent. Unfortunately there is little that can be done about this short of increasing the number of microphones (at additional cost) or decreasing the radius of the array (at the cost of low frequency SNR).

Fig. 2 also shows the low amplitude and the distortion of the second order response at low frequencies. Here some improvements

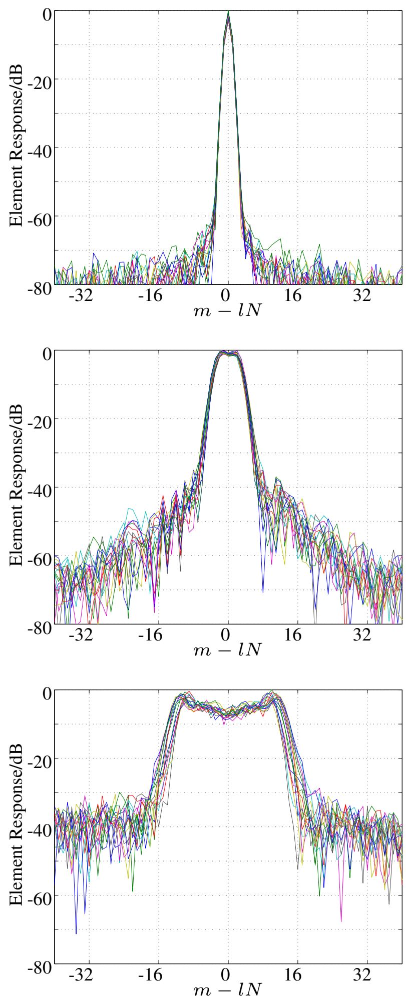


**Fig. 1**. Normalised Fourier transforms of the complex responses at 125 Hz (top), 2 kHz (middle), and 8 kHz (bottom).

may be possible. The analysis of [4] suggests that this distortion is caused by differences in the phase response between the elements of the array. These differences become more significant at low frequencies as the high order modes decrease in amplitude.

The conjecture motivating the research in this paper is that the difference between the low frequency responses may be estimated from certain signals. In this way the microphone array can "self-calibrate". In particular, the idea is to estimate the direction of arrival of a source or sources at mid-range frequencies, where the phase difference is not so significant. If the sources are also generating signals at low frequencies, the location information can be used to estimate the phase difference to each of the microphones, and hence the phase error can be calculated.

## 2. MODE ERROR CRITERION

The polar plots shown in Fig. 2 illustrate the problem in generating second order modes at low frequencies. However, for purposes of comparison it is desirable to have some sort of error measure. The measure used here is a simple squared error comparison with the ideal mode response. The ideal phase mode response is of the form

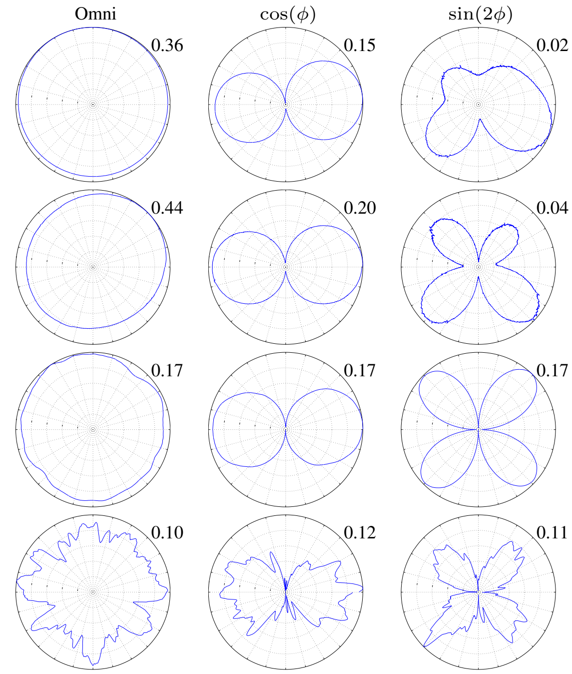


**Fig. 2**. Example raw polar amplitude responses from measured microphone data. 125 Hz (top row), 250 Hz (second row), 2 kHz (third row), and 8 kHz (bottom row). The maximum amplitude is the number at the top right of each response plot.

[4]

$$r_{m,s}(k, \phi_i) = \beta_m(k) e^{-jm\phi_i} \tag{4}$$

If we have measurements $r_m(\phi_i)$ of this response at a discrete set of angles $\phi_i$, then the total normalised error is given by

$$\eta = \sum_{i=1}^{N} \left| e^{jm\phi_i} - \frac{r_m(\phi_i)}{\beta_m} \right|^2 \tag{5}$$

For the purposes of providing an error measure, we are free to choose a value of $\beta_m$ that minimises this error. (This ignores frequency continuity, which will become important when equalisation of the responses across frequency is performed, but at present we are only concerned with phase correction at individual frequencies.) We therefore equate to zero the derivative of this expression with respect to the complex conjugate $\beta_m^*$ of $\beta_m$, and choose

$$\beta_m = \frac{\sum_{i=1}^{N} r_m(\phi_i) r_m^*(\phi_i)}{\sum_{i=1}^{N} e^{-jm\phi_i} r_m^*(\phi_i)} \tag{6}$$

This value of $\beta_m$ can then be used in (5) to obtain the error measure. This can then be used to evaluate the effectiveness of any proposed calibration scheme.

## 3. MICROPHONE MODEL

Implicit in the concept of calibration of the microphone array in this paper, is the notion that there is some small number of parameters (ideally one parameter) which characterise each element of the array. To derive a calibration method, we thus need a model for the elements. The model should be capable of being decomposed into two components

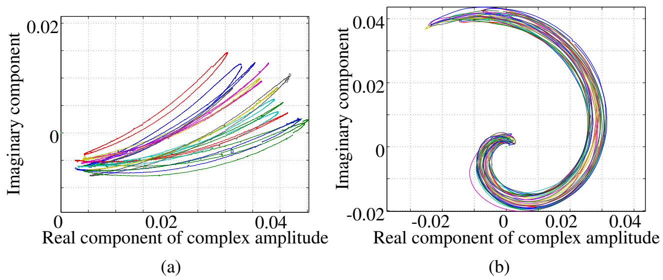


**Fig. 3**. Overlaid raw complex responses for 16 microphone elements at 125 Hz (a) and at 2 kHz (b).

1. the parameter or parameters required for calibration, which encapsulate the differences between the elements, and
2. the complex amplitude of the element, after the simple parametrisation has been removed. Ideally this should be approximately the same for the manufacture of an entire batch of microphone elements.

Some effort was put into producing an electrical equivalent model of each element based on physical considerations. The idea behind this approach is that the differences between each element must ultimately arise from some physical difference in their construction or arrangement. Some useful references for development of these models are [5] and [6]. However, since the elements are active, having a J-FET amplifier stage, element differences may also be due to the electronics. For this paper it was decided to use an empirical model for the elements.

### 3.1. Element Differences for Calibration

The complex polar responses for each of the 16 elements are overlaid in Fig 3 for a low and a mid-range frequency. The responses at low frequencies, where we most which to correct for differences, suggests that a simple rotation is all that is necessary to compensate for microphone differences. A complex rotation, which includes an amplitude scaling as well as a rotation may achieve better alignment, but at the cost of an additional real parameter to estimate.

We first investigate whether this model can be used to obtain better first and second order harmonics at low frequencies, in the situation where all of the response information is available.

Suppose then that we have 2 complex vectors $\mathbf{x}$ and $\mathbf{y}$, representing respectively the measured response and an ideal response. We wish to find a complex scaling $a$ so that $a\mathbf{x}$ and $\mathbf{y}$ are closest in some sense. For simplicity we again choose for our alignment metric the sum of the squared error between the two vectors, or

$$\eta_n = \|a_n \mathbf{x}_n - \mathbf{y}\|^2 \tag{7}$$

Equating the derivative with respect to $a_n^*$ to zero, we obtain

$$a_n = \frac{\mathbf{x}_n^H \mathbf{y}}{\mathbf{x}_n^H \mathbf{x}_n} \tag{8}$$

If we instead use a strict rotation, $e^{j\theta_n}$ with no scale adjustment, and equate the derivative with respect to $\theta_n$ to zero, we obtain

$$e^{j\theta_n} = \frac{\mathbf{x}_n^H \mathbf{y}}{\|\mathbf{x}_n^H \mathbf{y}\|} \tag{9}$$

Continuing with a squared loss metric, we may wish to find a "central" vector $\mathbf{y}$ which produces the lowest total loss summed over all of the $N$ elements of the array, or

$$\eta = \sum_{n=1}^{N} \eta_m = \sum_{n=1}^{N} \|\mathbf{y} - \mathbf{x}_n a_n\|^2 \tag{10}$$

Taking the derivative with respect to $\mathbf{y}$ we obtain

$$\mathbf{y} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{x}_n a_n \tag{11}$$

$$\text{or} \qquad \mathbf{y} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{x}_n e^{j\theta_n} \tag{12}$$

Iteration between (8) and (11), or between (9) and (12) rapidly converges to a self consistent solution. If scaling is allowed, then the vector of $a_n$ must be scaled at each iteration

$$\mathbf{a} \leftarrow \mathbf{a} \frac{\sqrt{N}}{\|\mathbf{a}\|} \tag{13}$$

This is to prevent the iterative process minimising the squared error by simply making the scaling factor smaller at every stage.

### 3.2. Single Element Model

Fig. 4 shows the amplitude and phase response of the microphones presented as functions of $\cos(\phi_i - \phi_n)$. (These are actually the averaged responses for several elements after alignment as discussed in Section 3.1.) Both the magnitude and phase appear to be approximately linear functions of $\cos(\phi_i - \phi_n)$. A least squares fit is shown as the line in each case. The error minimisation for this line was weighted so that values for which the amplitude is small contributed smaller error. The linearity suggests that the first order microphone model presented in [4] is at least approximately correct, and this is the model we use. The complex response for this model (for waves of elevation $\theta = \frac{\pi}{2}$) is given by

$$\begin{aligned} z_m(k, \phi_i - \phi_n) = {} & [\alpha + (1-\alpha)\cos(\phi_i - \phi_n)] \\ & \times \exp\left[jkR\cos(\phi_i - \phi_n)\right] \end{aligned} \tag{14}$$

where $\alpha$ is $\frac{1}{2}$ for an ideal cardioid, and $R$ is the radius of the array. At low frequencies a linear approximation for the phase is evidently not such a good approximation. There is also a loop in the response, indicating that angle of maximum amplitude of the response does not align with the maximum amplitude at other frequencies. Correcting this is beyond the scope of this paper. What is perhaps more significant, and is relatively easy to incorporate into the model, is the fact that at low frequencies, the phase difference between sources separated by 180° is greater than $kR$. At 2 KHz, the array radius can be estimated from the phase gradient as 77.5 mm, which is very close to the physical radius of 77.8 mm. At 125 Hz, the phase gradient corresponding to the least squares line shown in Fig. 4 leads to an estimated radius of 116 mm. This is because each of the elements has a phase response which varies with incidence angle, and the phase is not only due to the time difference of arrival at the element. It is simple to incorporate this into our single element model by using a "virtual" radius which is a function of frequency, rather than the true radius.

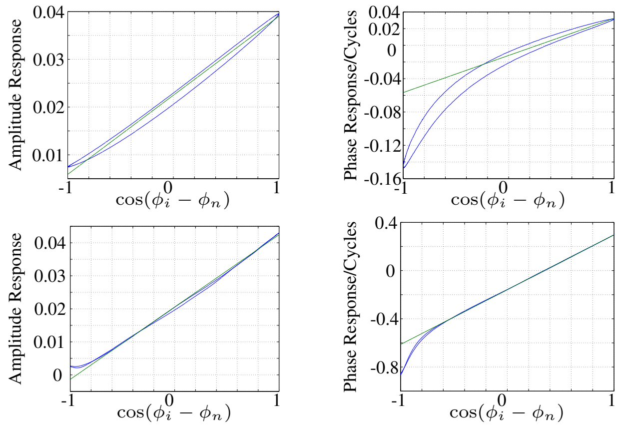


**Fig. 4**. Mean amplitude and phase responses for 125 Hz (top row) and 2 kHz (bottom row), each presented as a function of $\cos(\phi_i - \phi_n)$ along with a weighted linear fit.

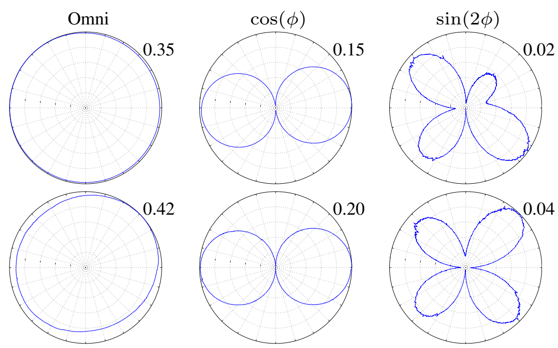


**Fig. 5**. Example raw polar amplitude responses from measured microphone data. 125 Hz (top row), 250 Hz (bottom row). The maximum amplitude is the number in the top right of each response plot. This should be compared to the upper half of Fig. 2.

## 4. ELEMENT DIFFERENCE MODEL VERIFICATION

The approach of applying a rotation to each of the measured complex responses of the array elements, as described in Section 3.1 was performed. Even though the algorithm for applying the rotation was not specifically optimised for minimising the error for any particular mode, a reduction in the error is apparent. This is shown by comparing Fig. 5 with the upper half of Fig. 2.

The errors for the modes, using the measure discussed in Section 2 for both rotation and complex rotation are shown in Table 1. It is apparent that the error, particularly for the second order modes, can be considerably reduced using this simple phase calibration approach.

## 5. SELF CALIBRATION

A maximum likelihood method of beam forming was used at 2 kHz to estimate the location of various single sources:

$$\widehat{\phi}_i = \max_{\phi_i} \left| \mathbf{b}\mathbf{z}^H(\phi_i) \right|^2 \tag{15}$$

where $\mathbf{b}$ is the measured response to a single source, and each element of $\mathbf{z}$ is given by (14). Incidentally, it was found that using the true radius of the array did not necessarily minimise the error variance. The measured response of the source *at a lower frequency* was then divided by the ideal model response to obtain a complex calibration factor for each of the $N$ microphone elements. The error obtained with these calibrations was found to depend on the actual location of the source, but one example is shown in the last two lines of Table 1. The error can be seen to be comparable to that obtained for the "perfect" rotation angle. A more sophisticated approach than that outlined here combines the information from several calibration angles to obtain superior calibration parameter estimates.

| | f/Hz | Omni | $\cos(\phi)$ | $\sin(\phi)$ | $\cos(2\phi)$ | $\sin(2\phi)$ |
|---|---|---|---|---|---|---|
| Raw | 125 | 0.028 | 0.081 | 0.106 | 0.447 | 0.593 |
| | 250 | 0.096 | 0.065 | 0.059 | 0.159 | 0.259 |
| Rotate | 125 | 0.029 | 0.083 | 0.076 | 0.303 | 0.485 |
| | 250 | 0.072 | 0.054 | 0.042 | 0.141 | 0.157 |
| Rotate & scale | 125 | 0.018 | 0.057 | 0.049 | 0.236 | 0.276 |
| | 250 | 0.065 | 0.027 | 0.045 | 0.141 | 0.112 |
| Calibration | 125 | 0.025 | 0.055 | 0.063 | 0.376 | 0.250 |
| | 250 | 0.070 | 0.045 | 0.057 | 0.127 | 0.090 |

**Table 1**. Error for each of several modes for raw data, and data with an optimal simple rotation or an optimal complex rotation applied to each element. The last two rows are the result for a complex rotation estimated from a calibration signal.

## 6. CONCLUSION

It has been shown that estimation of the second order modes of a sound field using a microphone array is difficult at low frequencies. This is primarily because the magnitude of the modes is very small. This results in any variations in the element properties becoming very significant. It has been further shown that these variations can be simply accounted for by a simple rotation or complex rotation, and in this way the usable low frequency range of the array can be extended. The rotation can be estimated from real signals, provided that the location information derived at mid-range frequencies can be applied to low frequencies. Further work must be performed to establish reliable methods of estimation of these parameters in real environments.

## 7. REFERENCES

[1] G. Weinreich and E.B. Arnold, "Method for measuring acoustic radiation fields," *J. Acoustic. Soc. Am.*, vol. 68, pp. 404–411, Aug 1980.

[2] T.D. Abhayapala and D.B. Ward, "Theory and design of high order sound field microphones using spherical microphne arrays," in *Proc. IEEE ICASSP*, Orlando, 2002, pp. 1949–1952.

[3] M.A. Poletti, "A unified theory of horizontal holographic sound systems," *J. Audio. Eng. Soc.*, vol. 48, pp. 1155–1182, Dec 2000.

[4] M.A. Poletti, "Effect of noise and transducer variability on the performance of circular microphone arrays," *J. Audio Eng. Soc.*, vol. 53, no. 5, pp. 371–384, May 2005.

[5] H.F. Olson, *Elements of Acoustical Engineering*, D. Van Nostrand, New York, 1940.

[6] L. L. Beranek, *Acoustics*, Acoustical Society of America, 1993.