The effects of non-cardioid directivity on incidence angle estimation using the polar energy time curve

J. James Esplin, Brian E. Anderson, and Timothy W. Leishman
Acoustics Research Group, Department of Physics and Astronomy, Brigham Young University, N283 Eyring Science Center, Provo, Utah 84602
jamesesplin@byu.edu, bea@byu.edu, tim_leishman@byu.edu

Brian T. Thornock
SRC Inc., 7502 Running Ridge Road, North Syracuse, New York 13212
brianthornock@gmail.com

Abstract: Assessment of desirable reflections and control of undesirable reflections in rooms are best accomplished if the reflecting surfaces are properly localized. Several measurement techniques exist to identify the incident direction of reflected sound, including the useful polar energy time curve (Polar ETC), which requires six cardioid impulse response measurements along the Cartesian axes. The purpose of this investigation is to quantify the incidence angle estimation error introduced into the Polar ETC by non-cardioid microphone directivities. The results demonstrate that errors may be minimized with a cardioid-family microphone possessing a certain range of directivities and by maximizing the measurement signal-to-noise ratio.

PACS numbers: 43.55.Mc, 43.58.Fm [NX]
Date Received: June 21, 2011 **Date Accepted:** August 17, 2011

1. Introduction

Over the years, many methods have been derived to determine the direction of arrival of distinct sound reflections in rooms. One of the first commercialized approaches was the polar energy time curve (Polar ETC) introduced by Becker.[1,2] It involves six sequential directional impulse response measurements using a cardioid microphone oriented along the positive and negative Cartesian axes. The difference in the so-called "energy time curve" (ETC) for each axis is used to identify the incidence angles (azimuth and elevation) for a given reflection.

Since Becker's work, several additional reflection localization measurements have been introduced. Yamasaki *et al.*[3] and Sekiguchi *et al.*[4] developed cross-correlation methods for four-microphone arrays. Choi *et al.*[5] employed a similar approach but included a fifth microphone at the geometric center of a tetrahedral array. Noël *et al.* added additional complexity to the microphone arrangement.[6] Gover *et al.* introduced an alternative scheme involving 32-microphone spherical arrays with a beamforming algorithm to determine directions of arrivals.[7–9] Abdou *et al.* used intensity measurements to characterize directional room response characteristics,[10] while Essert[11] and Farina *et al.*[12,13] employed ambisonic microphone arrays with intensity-based schemes. Each of these methods has its benefits and drawbacks as does the Polar ETC. However, because the Polar ETC involves only a single microphone and is available commercially, it continues to be a practical cost-effective tool for the problem of reflection localization.

The Polar ETC is based on an important assumption—namely, that the microphone maintains an ideal cardioid directivity pattern over the entire bandwidth of interest. In reality, typical cardioid microphones may only maintain a pattern approaching true cardioid directivity over a very limited bandwidth. To our knowledge, no study

quantifying errors introduced by violation of this assumption has been published. The purpose of this paper is to quantify angular estimation errors introduced by all members of the cardioid family of microphones, particularly to determine how important it is to use a true cardioid microphone rather than, for example, a subcardioid microphone or a hypercardioid microphone. The errors provide the confidence level one may expect from experimental results obtained with the method.

The paper will show that a certain range of cardioid family microphones can actually provide acceptable results (as long as the microphone does not have a pattern approaching omnidirectional or bidirectional directivity). The following sections explore theoretical error estimations when any type of cardioid family microphone is used and when the measurements possess various signal-to-noise ratios (SNR). Experimental results validate the theoretical findings.

2. Theoretical predictions

To predict the effects of microphone directivity on the Polar ETC, one must first understand the Polar ETC itself. The ETC is the envelope or magnitude of the analytic signal describing the impulse response. It is computed by creating a complex time signal with the impulse response as the real part and the Hilbert transform of the impulse response as the imaginary part.[14–17] Originally, it was interpreted as a measure proportional to either instantaneous energy density or instantaneous sound intensity. Becker and others[2,14] used it under these assumptions. However, it has since been shown that the ETC is based upon an acausal operation and does not accurately represent energy flow.[18,19] Additionally, the *ETC* components, used later in Eqs. (1) and (2), are not quadratic in the linear field variables, therefore they cannot be considered energy quantities. Other names have subsequently been used to describe it, including the "time response"[20] and the "envelope time curve".[21] Regardless, the Polar ETC does allow the localization of points of reflection and has been widely used for that purpose.

The essential forms of the equations developed by Becker[1] and D'Antonio *et al.*[2] are based on standard Cartesian-to-spherical coordinate system transformations. The process of incidence angle estimation using the Polar ETC is governed by the following two equations [corresponding to Eqs. (5) and (10c) in Ref. 2]:

$$\phi_M = \tan^{-1}\frac{ETC_{+y} - ETC_{-y}}{ETC_{+x} - ETC_{-x}} \tag{1}$$

and

$$\theta_M = \sin^{-1}\frac{ETC_{+z} - ETC_{-z}}{2E_0}, \tag{2}$$

where ϕ_M represents the measured (estimated) azimuthal angle for an incoming wave from the $+x$ axis, θ_M represents the measured (estimated) elevation angle for an incoming wave from the x-y plane, *ETC* is for the directional measurement taken in the Cartesian orientation indicated, and E_0 represents the instantaneous *ETC* value measured with an omnidirectional microphone obtained from the six measurements (to be specified later).

The impulse response h_d, measured by a directional microphone, is $h_d = H(\theta,\phi)h_{omni}$, where $H(\theta,\phi)$ is the microphone directivity function and h_{omni} is the impulse response that would be measured at the same location with an omnidirectional microphone. The corresponding ETC is defined by

$$ETC = \sqrt{(h_d)^2 + \left(\hat{h}_d\right)^2}, \tag{3}$$

where $\hat{}$ represents a Hilbert transform operator.[19] Accordingly, we may then express Eq. (3) as

$$ETC = |H(\theta,\phi)|\sqrt{(h_{omni})^2+\left(\hat{h}_{omni}\right)^2}, \tag{4}$$

and one may consider each of the directional impulse response quantities given in Eqs. (1) and (2) to be proportional to the absolute value of the directivity function for the microphone.

Here we consider only cardioid family directivity functions[22,23]

$$|H(\theta,\phi)| = |A + B\cos\phi\cos\theta|, \tag{5}$$

where $A \geq 0$, $B \geq 0$, $A + B = 1$, and ϕ and θ represent general azimuth and elevation angle dependencies respectively. If $A = 1$ and $B = 0$, Eq. (5) represents an omnidirectional directivity. If $A = 0.75$ and $B = 0.25$, Eq. (5) represents a subcardioid directivity. If $A = B = 0.5$, Eq. (5) represents a true cardioid directivity. If $A = 0.25$ and $B = 0.75$, Eq. (5) represents a hypercardioid directivity. If $B = 1$ and $A = 0$, Eq. (5) represents a bidirectional or "figure-8" directivity. D'Antonio *et al.* assumed that the directional ETC components in Eqs. (1) and (2) were proportional to the directivity function instead of its absolute value.[2] Using an approach similar to theirs, but with the absolute value in place, the Cartesian ETC components become

$$ETC_{\pm x} = E_0|A \pm B\cos\phi\cos\theta|, \tag{6}$$

$$ETC_{\pm y} = E_0|A \pm B\sin\phi\cos\theta|, \tag{7}$$

$$ETC_{\pm z} = E_0|A \pm B\sin\theta|, \tag{8}$$

$$E_0 = \sqrt{(ETC_{+x} - ETC_{-x})^2+(ETC_{+y} - ETC_{-y})^2+(ETC_{+z} - ETC_{-z})^2}\Big/2. \tag{9}$$

As another source of error, there is always a finite amount of noise in measurements that is inseparable from the desired arrival of sound energy. To quantify the angular estimation error introduced by a finite amount of noise, we use a theoretical approach, based upon statistical averaging. We modify the incident E_0 by adding a random number, $-1 \geq \varepsilon \leq 1$, the maximum amplitude of which is determined by a SNR (expressed in dB), so that $E_{0,i} = E_S + |E_S|\varepsilon_i 10^{-SNR/20}$. Equations (6) through (9) are then substituted into Eqs. (1) and (2) with each of Eqs. (6) through (8) having newly generated ε values to simulate random noise in subsequent measurements. A set of "actual" angles (ϕ_A, θ_A) are then substituted in place of (ϕ, θ) in the resulting equations:

$$\phi_M = \tan^{-1}\left[\frac{E_{0,3}|A + B\sin\phi_A\cos\theta_A| - E_{0,4}|A - B\sin\phi_A\cos\theta_A|}{E_{0,1}|A + B\cos\phi_A\cos\theta_A| - E_{0,2}|A - B\cos\phi_A\cos\theta_A|}\right], \tag{10}$$

$$\theta_M = \sin^{-1}\left[\frac{E_{0,5}|A + B\sin\theta_A| - E_{0,6}|A - B\sin\theta_A|}{E_0'}\right], \tag{11}$$

$$E_0' = \sqrt{\begin{array}{l}\left[E_{0,1}|A + B\cos\phi_A\cos\theta_A| - E_{0,2}|A - B\cos\phi_A\cos\theta_A|\right]^2\ldots \\ +\left[E_{0,3}|A + B\sin\phi_A\cos\theta_A| - E_{0,4}|A - B\sin\phi_A\cos\theta_A|\right]^2\ldots \\ +\left[E_{0,5}|A + B\sin\theta_A| - E_{0,6}|A + B\sin\theta_A|\right]^2\end{array}}. \tag{12}$$

The simulated random noise causes the "measured" angles to differ from the actual angles. The angular estimation error σ, which represents the central angle difference between the unit vectors associated with (ϕ_A, θ_A) and (ϕ_M, θ_M), is then determined through application of the standard trigonometric spherical law of cosines

$$\sigma = \cos^{-1}[\sin\theta_M \sin\theta_A + \cos\theta_M \cos\theta_A \cos(\theta_M - \theta_A)]. \quad (13)$$

To obtain a statistically significant result, we conduct the theoretical analysis, as described in Eqs. (10) through (12), for each value of $\phi_A = 0°...90°$ and $\theta_A = 0°...90°$ in 1° steps, 5 000 times with newly generated random numbers for each case, and average σ over the iterations. A weighted average $(\cos\theta_A)$ is applied to adjust the weighting of angles near the poles of a sphere where the angular area on the surface of the sphere is smaller than angular areas near the equator,

$$\bar{\sigma}(A) = \frac{1}{5000}\sum_{\varepsilon^n=1}^{5000}\left(\frac{1}{91}\sum_{\theta_A=0°,1°,2°...}^{90°}\left\{\frac{1}{91}\sum_{\phi_A=0°,1°,2°...}^{90°}(\cos\theta_A)\cos^{-1}\begin{bmatrix}\sin\theta_M\sin\theta_A + ... \\ ... + \cos\theta_M\cos\theta_A\cos(\phi_M - \phi_A)\end{bmatrix}\right\}\right), \quad (14)$$

where ε^n represents a newly generated set of random numbers for each Cartesian direction ($\varepsilon_{1..6}$). We then carry out this statistical process for various values of *SNR*. Figure 1 displays the results of this theoretical work averaged over all possible combinations of angles (we only utilize the ranges $\phi_A = 0°...90°$ and $\theta_A = 0°...90°$ due to symmetry). The minima of these curves generally occurs at $A = B = 0.5$, but the errors don't grow substantially until approximately $0.25 \geq A \geq 0.75$. Note that the errors rise significantly as $A \to 1$ (omnidirectional) or as $A \to 0$ (bidirectional). Even for the case of very little noise ($SNR = 100$) there exists some finite amount of error for values of $A < 0.5$. (Empirical studies confirm this result.) The information displayed in Fig. 1 can then be used to estimate the angular estimation error or confidence level one would expect when conducting Polar-ETC measurements, given a *SNR* and a specified cardioid-family directivity.

3. Experimental methods

To experimentally explore the angular estimation error of Polar ETC measurements, a setup involving a hemi-anechoic chamber, a Tannoy dual-concentric loudspeaker (System 800), and an AKG 414 microphone, with variable polar pattern settings [omnidirectional (OD), subcardioid (SC), cardioid (C), hypercardioid (HC), and figure-8 (F8)] was used. While the directivity of this microphone varies over frequency, it should be noted here that we assumed the selectable AKG 414 directivity patterns remained consistent over the entire audio measurement bandwidth. This approach allowed us to simplify the analysis while exploring the impact of distinct pattern

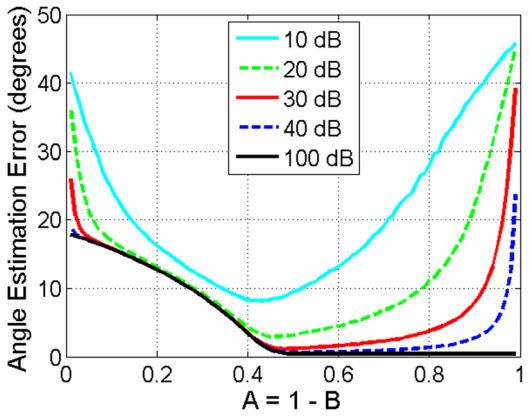

Fig. 1. (Color online) Simulated angular error in the Polar ETC as a function of the cardioid factor A, which varies from 0 (bidirectional directivity) to 1 (omnidirectional directivity). The five curves represent expected errors for the indicated signal-to-noise ratios.

variations. The effect on angular estimation error due to the variability of cardioid microphone directivities over frequency may be the subject of future work.

A set of Polar ETC experiments with reflecting surfaces at known angles of incidence and reflection was conducted to quantify the localization accuracy of the directions of arrival at the microphone. To take sequential measurements about a point in space with a single microphone, a microphone positioner was used that was able to rotate over a full 4π sr in desired increments, while maintaining consistency of the microphone diaphragm location throughout the rotations. The Polar ETC experiments then yielded the estimated angles. An altazimuth-mounted laser pointer, centered at the same point in space as the microphone diaphragm, was then used in conjunction with a set of planar mirrors on the reflecting surfaces to determine actual angles of arrival for the direct sound and the first reflection. The angular estimation error was then calculated using Eq. (13).

With this approach, two measurement configurations were employed. The first involved the microphone and loudspeaker placed in a hemi-anechoic chamber (hemi-anechoic above approximately 150 Hz). The direct sound arrived from an angle $(\phi = 0^\circ, \theta = 0^\circ)$ (measured in degrees) while the first reflection was incident from $(\phi = 0^\circ, \theta = -43.5^\circ)$. The second configuration employed the same microphone and speaker placement as the first configuration, but the x axis of the microphone was rotated -45° in ϕ relative to the loudspeaker, thus shifting the microphone coordinate system. In this case, the direct sound arrived at an angle of $(\phi = 45^\circ, \theta = 0^\circ)$ while the first reflection was incident from $(\phi = 45^\circ, \theta = -43.5^\circ)$. Ten tests were conducted for each configuration and the measured angles (ϕ_M, θ_M) were averaged for both the direct sound and the first reflection. The measurements employed a bandwidth up to approximately 30 kHz for the radiated sound and a 192 kHz sampling frequency. A MOTU 896mk3 digital audio interface was used in conjunction with EASERA software to acquire the data with a swept-sine excitation signal. Each set of 10 Polar ETC measurements for each configuration was then repeated for all five AKG microphone settings.

4. Results

Experimental error was determined from the average of the error results for both the direct sound and the first reflection for each of the different experimental setups. To compare experimental results to the theoretical results presented previously, an appropriate SNR of the experimental data had to be determined. Because the source settings were maintained throughout the experiments, the sound arrivals should have possessed equal incident sound pressure levels irrespective of the microphone configuration employed. We determined the SNR from the impulse responses obtained when the microphone was used in its omnidirectional setting. The noise level was taken as the peak level of the noise preceding the direct sound. The ETC peak levels of both the direct sound and the first reflection were averaged to represent the signal levels. The average SNR for the experiments was 35.1 dB. This ratio is admittedly low but results from an average of the direct sound and the first reflection, each having different signal-to-noise ratios. It is important to point out here that, for a typical impulse response in a given room, the SNR will decrease for successive reflection arrivals.

Figure 2 shows the comparison of the theoretical analysis averaged over the four incident angles employed in the experiments and the corresponding experimental results. Note that the theory curve here does not appear to fall in line with the theory in Fig. 1, but this is because the averaging here only takes place over four angles rather than averaging over a full octant as done in Fig. 1. Apparently theoretical error estimation for these four angles results in a different shaped curve, particularly between $0 \leq A \leq 0.5$. Further, note that if one wishes to compute an expected error for a particular experimental situation, one should compute the theoretical error estimation for that specific case because the theory depends on the incidence angle(s) and the SNR. We see that the experimental Polar ETC results follow the predicted theoretical trend.

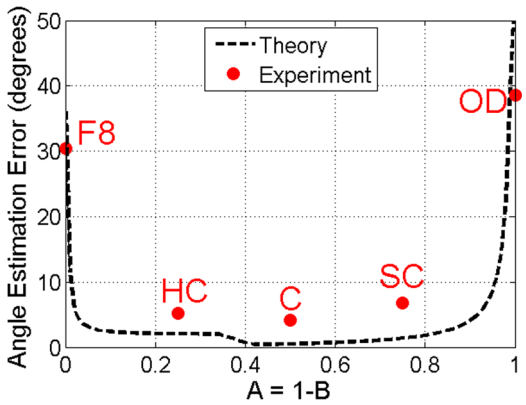

Fig. 2. (Color online) Comparison of theory (for a signal-to-noise ratio of 35.1 dB and averaged over the 4 angles of incidence) to the angular errors from the experiments.

Departure from theory, in terms of absolute levels of error estimation, may be due to imprecise positioning of the laser pointer setup in place of the microphone positioner, imprecise rotation of the microphone positioner about the microphone diaphragm, and microphone directivities that depart from the theoretical formulas over the measurement bandwidth. It is interesting to note that the errors for the SC, C, and HC microphone settings are within 2.8° of each other, suggesting that the strict use of a true cardioid microphone is not necessary. As long as the microphone is not nearly OD or F8, the angular errors remain relatively low. This also implies that one should consider band limiting Polar ETC measurements to eliminate the OD behavior typical at low frequencies and the erratic behavior typical at high frequencies for cardioid microphones.

5. Conclusions

This paper has quantified the angular estimation error introduced when non-cardioid microphones are used for Polar ETC measurements. Theoretical developments and experimental results have shown that angular estimation errors in the use of the Polar ETC significantly increase when omnidirectional or figure-8 microphones are used. Further, the errors introduced by the use of subcardioid or hypercardioid microphones are not significantly higher than those when using a cardioid microphone. Theoretical results presented here also provide insight into how angular estimation errors increase with decreasing signal-to-noise ratio.

Acknowledgments

The authors thank Wes Lifferth for his machining of the altazimuth-mounted laser pointer and Ken Forster for his machining of the microphone positioner. Funding for author J. J. Esplin was provided by the Brigham Young University Department of Physics and Astronomy.

References and links

[1]F. M. Becker, "The polar energy time curve," in *6th International Conference on Sound Reinforcement* Audio Eng. Soc. (1988), pp. 54–57.

[2]P. D'Antonio, J. Konnert, F. Becker, and C. Bilello, "Sound intensity and interaural cross-correlation measurements using time-delay spectrometry," J. Audio Eng. Soc. **37**(9), 659–673 (1989).

[3]Y. Yamasaki and T. Itow, "Measurement of spatial information in sound fields by closely located four point microphone method," J. Acoust. Soc. Jpn. **10**(2), 101–110 (1989).

[4]K. Sekiguchi, S. Kimura, and T. Hanyuu, "Analysis of sound fields on spatial information using a four-channel microphone system based on a regular tetrahedron peak point method," Appl. Acoust. **37**(4), 305–323 (1992).
[5]C. Choi, L-H. Kim, Y. Oh, S. Doo, and K-M. Sung, "Measurement of early reflections in a room with five microphone system," IEICE Trans. Fund. Elect. Comm. Comp. Sci. **E86-A**, 3283–3287 (2003).
[6]C. Noël, V. Planeaua, and D. Habault, "A new temporal method for the identification of source directions in a reverberant hall," J. Sound Vib. **296**(3), 518–538 (2006).
[7]B. N. Gover, J. G. Ryan, and M. R. Stinson, "Microphone array measurement system for analysis of directional and spatial variations of sound fields," J. Acoust. Soc. Am. **112**(5), 1980–1991 (2002).
[8]B. N. Gover, J. G. Ryan, and M. R. Stinson, "Measurements of directional properties of reverberant sound fields in rooms using a spherical microphone array," J. Acoust. Soc. Am. **116**(4), 2138–2148 (2004).
[9]B. N. Gover, "Directional measurement of airborne sound transmission paths using a spherical microphone array," J. Audio Eng. Soc. **53**(9), 787–795 (2005).
[10]A. Abdou and R. W. Guy, "Directional accuracy of transient sound employing an intensity probe," Appl. Acoust. **50**(1), 65–77 (1997).
[11]R. Essert, "Progress in concert hall design—developing an awareness of spatial sound and learning how to control it," EBU Tech. Rev. Winter, 31–39 (1997).
[12]A. Farina, A. Capra, L. Conti, P. Martignon, and F. Fazi, "Measuring spatial impulse responses in concert halls and opera houses employing a spherical microphone array," in *19th International Congress on Acoustics* (Madrid, Spain, September, 2007), pp. 1–6.
[13]A. Farina, A. Capra, P. Martignon, and S. Fontana, "Measuring impulse responses containing complete spatial information," in *the 22nd Audio Engineering Society of the United Kingdom Conference* (Cambridge, UK, April, 2007), pp. 1–5.
[14]R. C. Heyser, "Determination of loudspeaker signal arrival times: Part 1," J. Audio Eng. Soc. **19**(9), 734–743 (1971).
[15]R. C. Heyser, "Determination of loudspeaker signal arrival times: Part 2," J. Audio Eng. Soc. **19**(10), 829–834 (1971).
[16]R. C. Heyser, "Determination of loudspeaker signal arrival times: Part III," J. Audio Eng. Soc. **19**(11), 902–905 (1971).
[17]R. C. Heyser, "Instantaneous intensity," in *the 81st Convention of the Audio Engineering Society*, preprint 2399, 1–11 (1986).
[18]D. B. Keele, "The analytic impulse and the energy-time curve: The debate continues," in *the 93rd Convention of the Audio Engineering Society*, preprint 3399, 1–19 (1992).
[19]J. Vanderkooy and S. P. Lipschitz, "Uses and abuses of the energy-time curve," J. Audio Eng. Soc. **38**(11), 819–836 (1990).
[20]H. Biering and O. Z. Pedersen,"System analysis and time delay spectrometry (Part I)," Brüel and Kjær Tech. Rev., No. **1**, 3–51 (1983).
[21]D. Davis and E. Patronis, Jr., *Sound System Engineering*, 3rd ed., (Focal Press, London, 2006), pp. 135–142.
[22]G. Martin, "The significance of interchannel correlation phase and amplitude differences on multichannel microphone techniques," in *the 113th Convention of the Audio Engineering Society*, preprint 5671, 1–18 (2002).
[23]J. Braasch, "A loudspeaker-based 3D sound projection using virtual microphone control (ViMiC)," in *the 118th Convention of the Audio Engineering Society*, preprint 6430, 1–11 (2005).