Analysis and Design of Spherical Microphone Arrays

Boaz Rafaely, *Senior Member, IEEE*

Abstract—**Spherical microphone arrays have been recently studied for sound-field recordings, beamforming, and sound-field analysis which use spherical harmonics in the design. Although the microphone arrays and the associated algorithms were presented, no comprehensive theoretical analysis of performance was provided. This paper presents a spherical-harmonics-based design and analysis framework for spherical microphone arrays. In particular, alternative spatial sampling schemes for the positioning of microphones on a sphere are presented, and the errors introduced by finite number of microphones, spatial aliasing, inaccuracies in microphone positioning, and measurement noise are investigated both theoretically and by using simulations. The analysis framework can also provide a useful guide for the design and analysis of more general spherical microphone arrays which do not use spherical harmonics explicitly.**

Index Terms—**Plane-wave decomposition, sound-field analysis, spatial aliasing, spatial sampling, spherical Fourier transform, spherical microphone arrays.**

I. INTRODUCTION

RECENT studies of spherical microphone arrays embraced a wide range of applications, including sound-field recordings, beamforming for speech enhancement, and sound-field analysis. Meyer and Elko presented spherical microphone arrays for spatial sound recordings [1] and beamforming [2], while Abhayapala and Ward [3] studied the limitations of a similar array in recording a plane-wave sound field. Weinreich and Arnold [4] designed a spherical measurement array with a single rotating microphone to analyze acoustic radiation fields, while recently, Rafaely [5]–[7] showed that such a spherical measurement array can be used for plane-wave decomposition by employing spherical convolution. The advantage of the spherical configuration is the three-dimensional symmetry useful in spatial sound-field analysis. The reader is referred to the papers above for further details and a more comprehensive review. Although these studies presented the array processing employed and in some cases a preliminary analysis of performance, a more comprehensive, theoretical analysis of performance which is valuable in the design and analysis of spherical microphone arrays in practice was not provided.

This paper presents such a theoretical framework of analysis. First, three spatial sampling schemes are presented and compared, which can be used for microphone positioning. Then, a theoretical error analysis is developed where the separate effects

of finite order, or finite number of microphones, inaccuracies in the positioning of microphones, spatial aliasing and measurement noise are evaluated. The theoretical analysis is supported by simulation examples of microphone arrays with various configurations. The analysis in this paper is based on spherical-harmonics decomposition, or the spherical Fourier transform, and could therefore, directly apply to the recently studied spherical-harmonics based microphone arrays [1]–[3], [5]–[7]. Nevertheless, since the analysis involves performance limits due to physical factors, the framework and results presented here could also be useful in the design of spherical microphone arrays in general (which are not explicitly designed using spherical harmonics), such as the array designed by Gover *et al.* for sound-field analysis [8].

II. SPHERICAL FOURIER TRANSFORM

The spherical Fourier transform [9], or spherical harmonics decomposition [10], employed in this paper is briefly revised in this section. The reader is referred to [5], [9], and [10] for more details. The standard spherical coordinate system (r, θ, ϕ) is used (e.g., [13]), where in addition Ω denotes the position (θ, ϕ) on a unit sphere. Consider a function $f(\Omega)$ which is square integrable on the unit sphere, then its spherical Fourier transform, denoted by f_{nm}, and the inverse transform, are given by [9]

$$f_{nm} = \int_{\Omega \in S^2} f(\Omega) Y_n^{m^*}(\Omega) d\Omega \qquad (1)$$

$$f(\Omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} f_{nm} Y_n^m(\Omega). \qquad (2)$$

The spherical harmonics Y_n^m are defined by

$$Y_n^m(\Omega) = Y_n^m(\theta, \phi) = \sqrt{\frac{(2n+1)}{4\pi} \frac{(n-m)!}{(n+m)!}} P_n^m(\cos\theta) e^{im\phi} \qquad (3)$$

where n is the order of the spherical harmonics and $i = \sqrt{-1}$. The spherical harmonics are the solution to the wave equation, or the Helmholtz equation in spherical coordinates [10], [13]. The associated Legendre function P_n^m represents standing spherical waves in θ while the term $e^{im\phi}$ represents traveling spherical waves in ϕ. The integral $\int_{\Omega \in S^2} d\Omega = \int_0^{2\pi} \int_0^{\pi} \sin\theta d\theta d\phi$ covers the entire surface area of the unit sphere, denoted by S^2 [10]. The spherical harmonic are orthonormal and satisfy [13]

$$\int_{\Omega \in S^2} Y_{n'}^{m'}(\Omega) Y_n^{m^*}(\Omega) d\Omega = \delta_{n-n'} \delta_{m-m'}. \qquad (4)$$

Also, the addition theorem [13] provides the following simplification

$$\sum_{m=-n}^{n} Y_n^m(\Omega_1) Y_n^{m^*}(\Omega_2) = \frac{2n+1}{4\pi} P_n(\cos\Theta) \qquad (5)$$

where Θ is the angle between Ω_1 and Ω_2, and P_n is the Legendre polynomial. This further simplifies to $(2n+1)/4\pi$ when $\Omega_1 = \Omega_2$.

III. SPATIAL SAMPLING ON A SPHERE

Microphone arrays spatially sample the sound field, and in particular, spherical arrays perform spatial sampling of functions (e.g., sound pressure) defined on a sphere. Spatial sampling, similar to time-domain sampling, requires limited bandwidth (limited harmonic order) to avoid aliasing. Several sampling schemes are presented, which offer a tradeoff between the required number of microphones and the simplicity of their arrangement. A sampling scheme can be considered exact if the spherical Fourier coefficients can be computed from the spatial samples without error. This requires the following equality:

$$f_{nm} = \sum_{j} \alpha_j f(\Omega_j) Y_n^{m^*}(\Omega_j) \qquad (6)$$

where $\Omega_j = (\theta_j, \phi_j)$ are the samples on the sphere, and j ranges for the entire sample set. The weights α_j are introduced to support the equality. Note that given f_{nm}, $f(\Omega)$ can be computed exactly from (2). Substituting (2) in (6), the condition for exact sampling as in (6) is satisfied if the following holds:

$$\sum_{j} \alpha_j Y_{n'}^{m'}(\Omega_j) Y_n^{m^*}(\Omega_j) = \delta_{n-n'} \delta_{m-m'}. \qquad (7)$$

Equation (7) can be considered as a modified version of the orthogonality property of the spherical harmonics [10], [13]. Parseval's relation can be written for the spherical Fourier transform as defined in (6) and (2), which includes the weighting parameter α_j, as

$$\sum_{j} \alpha_j |f(\Omega_j)|^2 = \sum_{n=0}^{N} \sum_{m=-n}^{n} |f_{nm}|^2 \qquad (8)$$

where N is the highest order for which f_{nm} is nonzero. Several alternatives for the sampling set Ω_j and the corresponding weights α_j are discussed below.

A. Equiangle Sampling

Driscoll and Healy [9] presented a sampling theory for band limited functions on a sphere ($f_{nm} = 0$ for $n > N$), which required $4(N+1)^2$ samples, with the angles θ and ϕ each sampled uniformly at $2(N+1)$ locations, e.g., $\theta_j = \pi j/(2N+2)$, $j = 0 \ldots 2N+1$, and $\phi_k = 2\pi k/(2N+2)$, $k = 0 \ldots 2N+1$. The spherical Fourier transform in this case is generally given by [9]

$$f_{nm} = \sum_{j=0}^{2N+1} \sum_{k=0}^{2N+1} \alpha_j f(\theta_j, \phi_k) Y_n^{m^*}(\theta_j, \phi_k). \qquad (9)$$

The weights α_j are introduced to compensate for the denser grid near the poles, and are given in [9]. Further details can be found in Driscoll and Healy [9], where it is shown that the equiangle sampling scheme satisfies the Nyquist sampling condition along ϕ, while the sampling along θ needs to satisfy $\sum_{j=0}^{2N+1} \alpha_j P_n(\cos\theta_j) = \sqrt{2}\delta_n$. The advantage of the equiangle sampling scheme is the regular angle differences. This could be useful when spatial samples are taken by a rotating microphone, for example, where a simple, equal-step mechanical rotation in both θ and ϕ is an advantage. The disadvantage of this scheme is the large number of samples compared to alternative sampling schemes.

B. Gaussian Sampling

The Gaussian sampling scheme described in this section requires only $2(N+1)^2$ samples, which is half of the equiangle sampling scheme. The azimuth angle ϕ is sampled at $2(N+1)$ equiangle samples as in the equiangle sampling scheme, but the elevation angle θ requires only $N+1$ samples, which are nearly equally spaced. The spherical Fourier transform is given in this case by

$$f_{nm} = \sum_{j=0}^{N} \sum_{k=0}^{2N+1} \alpha_j f(\theta_j, \phi_k) Y_n^{m^*}(\theta_j, \phi_k). \qquad (10)$$

The samples θ_j are computed as the zeros of the $N+1$ degree Legendre polynomial [13], [15] $P_{N+1}(\cos\theta_j) = 0, j = 0 \ldots N$. This choice ensures that the sampling condition in (7) is satisfied, since the polynomials in (7) are of degree not higher then $2N$ [15]. The weights α_j which satisfy the sampling condition can then be calculated or taken from tables [13], [16]. The advantage of the Gaussian sampling scheme is the reduced number of sample points for a given order N. The drawback is the potential inconvenience with the nonequal spacings along θ when, for example, microphones are mechanically rotated and equal-step rotation might be an advantage.

C. Nearly Uniform Sampling

This sections describes sampling schemes where the samples are distributed uniformly and nearly uniformly on the surface of a sphere. A sampling scheme where the distance between neighboring samples is constant, give rise to a limited set of special geometries called platonic solids, e.g., 20 samples on the vertices of a dodecahedron [17]. A sampling scheme with the samples distributed nearly uniformly on the sphere surface offer a much wider range of configurations. Hardin and Sloane [18] prove that the number of samples for these configurations is at least $(N+1)^2$, but show that in practice this number is larger then $1.5(N+1)^2$ for many examples, when the weights α_j are taken to be equal. Their configurations satisfy the sampling condition in (7) with negligible error (less then 4×10^{-13} for the examples employed in this paper). Lebedev (see [19] and references within) also provides samples and weights for various configurations, with a total of about $1.3(N+1)^2$ samples. It is important to note that a polynomial degree of $2N$ should be considered when applying these methods since the sampling condition in (7) involves a product of two Legendre polynomials,

each of degree N or lower. The spherical Fourier transform in this case is represented by (6), with Ω_j the set of samples and α_j the corresponding weights.

The advantage of the nearly uniform sampling scheme is the small number of samples. The disadvantage, as in the Gaussian sampling scheme, is the potential inconvenience with the nonequiangle locations. Furthermore, the equiangle and Gaussian sampling schemes can benefit from FFT-based computations along ϕ, which are not suitable for nearly uniform sampling.

IV. SPHERICAL ARRAY PROCESSING

Spherical microphone arrays typically measure the sound pressure p around a sphere, and the array processing is then performed by multiplying the pressure with the aperture weighting function w^* and integrating over the sphere, to give the array output at each frequency

$$y(kr) = \int_{\Omega \in S^2} p(kr, \Omega) w^*(k, \Omega) d\Omega$$
$$= \sum_{n=0}^{\infty} \sum_{m=-n}^{n} p_{nm}(kr) w_{nm}^*(k) \qquad (11)$$

where p_{nm} and w_{nm} are the spherical Fourier transform coefficients of p and w, k and r are the wave number and sphere radius, and (2) and (4) were employed to derive the summation term. The complex conjugate of w was used to simplify notation. If w was used under the integral, then w_{nm}^* would be replaced by $(-1)^m w_{n(-m)}$ under the summation, while if w_{nm} was used instead of w_{nm}^* then $w(\Omega) = w(\theta, \phi)$ would be replaced by $w(\theta, -\phi)$ under the integral [10]. The summation term in (11) can be viewed as weighting in the spherical Fourier transform domain, also called phase-mode processing [11], [12].

In practice, the pressure on the sphere is spatially sampled at the microphone positions Ω_j, and in addition the weighting function w_{nm}^* is limited to order $n \leq N$, which depends on the number of microphones used as discussed in Section III. Given the frequency-dependent output p of M microphones arranged on a sphere according to the chosen sampling scheme, the array output can be calculated as

$$y(kr) = \sum_{j=1}^{M} \alpha_j p(kr, \Omega_j) w^*(k, \Omega_j)$$
$$= \sum_{n=0}^{N} \sum_{m=-n}^{n} p_{nm}(kr) w_{nm}^*(k). \qquad (12)$$

The coefficients p_{nm} can be approximated using (6) when using the second line of (12) for the array processing, while the weights α_j depend on the sampling scheme as detailed in Section III.

The choice of w_{nm}^* depends on the application. Some examples are analyzed below, but first the coefficients of single and multiple plane-wave sound-fields are briefly presented [3], [5]. Consider a unit amplitude plane wave given by $e^{i(\omega t - \mathbf{k} \cdot \mathbf{r})}$ [14], where ω, t, $\mathbf{r}$, $\mathbf{k}$ are the angular frequency, time, position and

wave number vector, respectively, with the latter denoting the wave direction of propagation. This notation is slightly modified here to $e^{i(\omega t + \mathbf{k} \cdot \mathbf{r})}$, with $\mathbf{k} = (k, \Omega_0)$ now denoting the wave arrival direction. The wave number is given by $k = \omega/c$, with c denoting the speed of sound. The spherical Fourier transform of the pressure at the microphone positions $\mathbf{r} = (r, \Omega_j)$ can now be written as [10]

$$p_{0nm}(kr) = b_n(kr) Y_n^{m^*}(\Omega_0) \qquad (13)$$

where b_n for an open sphere and a rigid sphere is given by [10]

$$b_n(kr) = \begin{cases} 4\pi i^n j_n(kr), & \text{open sphere} \\ 4\pi i^n \left(j_n(kr) - \frac{j_n'(kr_0)}{h_n'(kr_0)} h_n(kr) \right), & \text{rigid sphere} \end{cases} \qquad (14)$$

where j_n, h_n are the spherical Bessel and Hankel functions, respectively, j_n', h_n' are their derivatives, and $r_0 \leq r$ is the radius of the rigid sphere. Now consider a sound field composed of an infinite number of plane waves that arrive at the sphere from all directions Ω_0 with amplitudes $a(k, \Omega_0)$ (strictly, a is the spatial density of the amplitudes). The spherical Fourier transform of the total pressure on the sphere can be written by integrating (13) over these directions as [5]

$$p_{nm}(kr) = a_{nm}(k) b_n(kr). \qquad (15)$$

A beamforming array of order N with weights $w_{nm}^* = d_{nm}/b_n$ is analyzed next. The array output for a plane wave arriving from Ω_0 is computed by substituting the chosen weights and (13) in (12), dropping the dependance on kr for notation simplicity

$$y = \sum_{n=0}^{N} \sum_{m=-n}^{n} d_{nm} Y_n^{m^*}(\Omega_0) = d(\Omega_0) \qquad (16)$$

which as expected gives the designed array directivity at the plane wave direction. Meyer and Elko [2] proposed array weights $w_{nm}^* = d_n Y_n^m(\Omega_l)/b_n$, with Ω_l the look direction, which allowed efficient steering of the array beam.

For sound field analysis by plane-wave decomposition [3], [5]–[7], the weights are $w_{nm}^* = Y_n^m(\Omega_l)/b_n$, where Ω_l is the arrival direction of the decomposed plane wave. The array output in this case is given by substituting (15) in (12)

$$y = \sum_{n=0}^{N} \sum_{m=-n}^{n} a_{nm} Y_n^m(\Omega_l) \approx a(\Omega_l) \qquad (17)$$

which gives the sound field plane-wave decomposition at the array output. The approximation becomes equality when N tend to infinity or when the order of a_{nm} is limited to $n \leq N$. A simpler choice of $w_{nm}^* = Y_n^m(\Omega_l)$ is used to reconstruct the sound pressure on the entire sphere at positions Ω_l, which could be useful in the processing of sound recordings. The array output in this case approximates $p(\Omega_l)$.

V. RIGID VERSUS OPEN SPHERE CONFIGURATIONS

An important consideration in the design of spherical microphone arrays is the sphere configuration. An open sphere configuration [3], [4], [8] as well as microphones arranged around
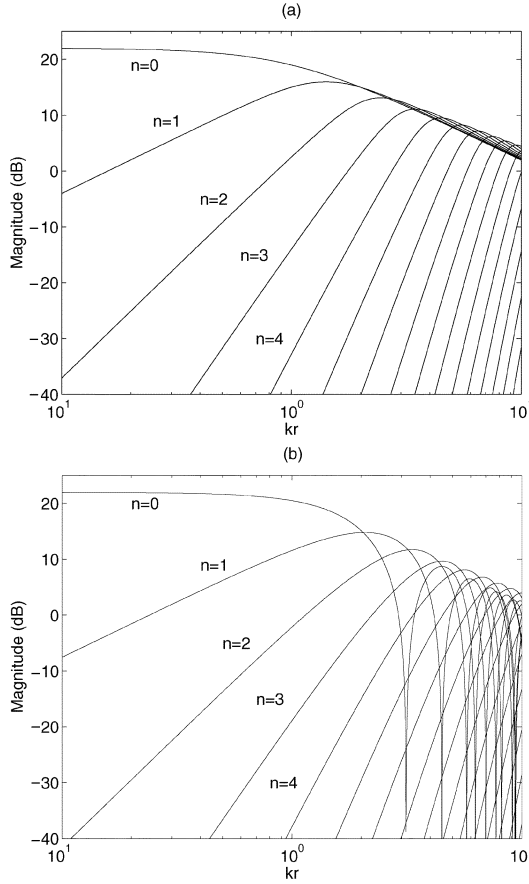
Fig. 1.   Magnitude of b_n as a function of kr for (a) rigid sphere with $kr = kr_0$, and (b) open sphere.

a rigid sphere [2] were previously suggested. The disadvantage of a rigid sphere is that it interferes with the surrounding sound field. Although scattering is accounted for in the calculation of the incident field, the scattered waves can be further reflected by other objects in the measurement environment and then be considered as additional incident waves when arriving back at the sphere. This effect might be negligible for small spheres, but may become more significant when large rigid spheres are employed and an accurate measurement of the sound field is required. However, the advantage of the rigid sphere is the improved numerical conditioning. This is because the array weights require the inversion of the function b_n, which is zero for some values of n and kr in the open sphere configuration, but not in the rigid sphere configuration [2], [5], [13], as shown in Fig. 1. One way to overcome this problem for the open sphere is to use an additional set of measurements. For example, the use of more than one sphere with varying radius was proposed in [3], while another option is to use measurement of radial velocity to complement the pressure measurement. Ways to improve open sphere measurement are the topic of current research. In this paper the analysis assumes a general b_n, which can apply to both open and rigid spheres, but rigid spheres are considered in the examples.

VI. FINITE-ORDER EFFECTS

The number of microphones is limited in practical arrays. As shown in Section III this limits the order to some value $n \leq N$

which depends on the sampling scheme. The finite order has two significant effects. First, the upper frequency is limited to around $kr < N$ due to spatial aliasing. This is further analyzed in Sections VII–IX. Second, the spatial resolution of the array is limited, which imposes limitations on the spatial performance of the array. The width of the main lobe in the beam pattern is one common feature of spatial resolution. Although this will depend on the choice of the weights, a simple expression can be derived for the case that $w_{nm}^* = Y_n^m(\Omega_l)/b_n$, i.e., for plane-wave decomposition arrays. Assuming a unit amplitude plane-wave arriving from Ω_0, the array output becomes [5]

$$
\begin{aligned}
y &= \sum_{n=0}^{N} \sum_{m=-n}^{n} Y_n^{m^*}(\Omega_0) Y_n^m(\Omega_l) \\
&= \frac{N+1}{4\pi(\cos\Theta - 1)} \left(P_{N+1}(\cos\Theta) - P_N(\cos\Theta) \right) \quad (18)
\end{aligned}
$$

where Θ is the angle between Ω_0 and Ω_l. Equation (5) and a result from Gradshteyn and Ryzhik [20, Sec. 8.915, p. 1026]) were employed in the derivation. In this case the width of the main lobe, between the first zeros, can be approximated by $2\pi/N$ [5], a measure that gives a simple dependence of spatial resolution on the order of the array. The width of the main lobe is therefore approximately proportional to $\sqrt{M}$, given M, the number of microphones, is proportional to $(N+1)^2$.

VII. ALIASING AND MEASUREMENT ERRORS EFFECTS

The performance of microphone arrays in practice is compromised by a number of factors. First, array processing involves spatial sampling, as discussed in Sections III and IV, and since a sound field such as a plane-wave field is not order-limited, aliasing errors will be present due to the higher orders [3]. Second, limited accuracy in the positioning of microphones in practical arrays will cause deviation of the actual sampling positions from the designed positions, which could introduce errors at the array output. In addition, transducer noise could also appear as noise at the array output and limit performance. Furthermore, the effect of these factors might depend on the sampling scheme chosen. Analysis of the separate effect of these factors can be a valuable tool in the design of microphone arrays in practice. Such an analysis scheme is presented in this section and its relation to white-noise-gain analysis is discussed. Note that other similar effects such as uncertainty in the microphones response could also be analyzed in a similar manner.

The output of the individual microphones at a given frequency which includes the effects as described above can be written as $p(\Omega_j') + e_j$, where Ω_j' is the actual microphone position and e_j is transducer noise. Expressing p_{nm} as in (6) and substituting in (12), the array output in this case becomes

$$
y = \sum_{n=0}^{N} \sum_{m=-n}^{n} \sum_{j=1}^{M} \alpha_j \left(p(\Omega_j') + e_j \right) Y_n^{m^*}(\Omega_j) w_{nm}^*. \quad (19)
$$

The signal and noise terms are now separated, with p, the actual pressure not necessarily order limited, replaced by applying

the inverse spherical Fourier transform and then changing the order of summations

$$
y = \sum_{n=0}^{N} \sum_{m=-n}^{n} w_{nm}^* \left(\sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} p_{n'm'} \right.
$$
$$
\times \left[\sum_{j=1}^{M} \alpha_j Y_{n'}^{m'}(\Omega_j') Y_n^{m^*}(\Omega_j) \right] + \left. \sum_{j=1}^{M} \alpha_j e_j Y_n^{m^*}(\Omega_j) \right). \quad (20)
$$

The expression in the square brackets is similar to the sampling condition as in (7), which can be extended to include errors due to aliasing (ϵ_a) and microphone positioning (ϵ_Ω), as follows:

$$
\sum_{j=1}^{M} \alpha_j Y_{n'}^{m'}(\Omega_j') Y_n^{m^*}(\Omega_j)
$$
$$
= \begin{cases} \delta_{n-n'}\delta_{m-m'} + \epsilon_{\Omega_{nmn'm'}}, & n, n' \le N \\ \epsilon_{a_{nmn'm'}} + \epsilon_{\Omega_{nmn'm'}}, & n \le N, n' > N. \end{cases} \quad (21)
$$

Given a sampling scheme (the sets α_j and Ω_j), the aliasing error term ϵ_a can be found by

$$
\epsilon_{a_{nmn'm'}} = \sum_{j=1}^{M} \alpha_j Y_{n'}^{m'}(\Omega_j) Y_n^{m^*}(\Omega_j)
$$
$$
n \le N, \ n' > N \quad (22)
$$

and given the actual positions Ω_j', the positioning error term ϵ_Ω can be found from (21) and (22) by eliminating the contribution of the aliasing error

$$
\epsilon_{\Omega_{nmn'm'}} = \sum_{j=1}^{M} \alpha_j \left[Y_{n'}^{m'}(\Omega_j') - Y_{n'}^{m'}(\Omega_j) \right] Y_n^{m^*}(\Omega_j)
$$
$$
n \le N, \ n' \ge 0. \quad (23)
$$

Substituting (21) in (20) and separating the summation over n', the array output is separated into signal and errors contributions

$$
y = \sum_{n=0}^{N} \sum_{m=-n}^{n} w_{nm}^* p_{nm} \qquad (y_s - \text{signal})
$$
$$
+ \sum_{n=0}^{N} \sum_{m=-n}^{n} w_{nm}^* \left\{ \sum_{n'=N+1}^{\infty} \sum_{m'=-n'}^{n'} p_{n'm'} \epsilon_{a_{nmn'm'}} \right\}
$$
$$
(y_a\text{-aliasing error})
$$
$$
+ \sum_{n=0}^{N} \sum_{m=-n}^{n} w_{nm}^* \left\{ \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} p_{n'm'} \epsilon_{\Omega_{nmn'm'}} \right\}
$$
$$
(y_\Omega\text{-positioning error})
$$
$$
+ \sum_{n=0}^{N} \sum_{m=-n}^{n} w_{nm}^* \left\{ \sum_{j=1}^{M} \alpha_j e_j Y_n^{m^*}(\Omega_j) \right\}
$$
$$
(y_e\text{-measurement noise}). \quad (24)
$$

Equation (24) now provides the basis for performance analysis which takes separate account of the various errors, and can

therefore be used to analyze the effects of individual factors. The expressions in the curly brackets represent un-weighted contributions with order n and degree m to each error, which when analyzed and compared can offer valuable design information. The contribution of terms with different orders to the errors is further analyzed in Section VIII. Three relative error measures can be defined which relate the power of the signal to the power of the various errors at each frequency

$$
E_a = \frac{|y_a|^2}{|y_s|^2} \qquad E_\Omega = \frac{|y_\Omega|^2}{|y_s|^2} \qquad E_e = \frac{E\left[|y_e|^2\right]}{|y_s|^2} \quad (25)
$$

where y_s, y_a, y_Ω, y_e have been defined in (24). Note that E_e can be considered as the reciprocal of the white noise gain, as discussed below. Given the sound field defined by p_{nm} and the array coefficients defined by w_{nm}^*, the effect of aliasing, positioning error and noise can be analyzed using the measures in (25) for a given sampling scheme. The relations between the signal and various errors are explored in Sections VIII and IX.

VIII. ERRORS ANALYSIS FOR A PLANE-WAVE SOUND FIELD

It is common when analyzing array performance to consider a plane-wave sound field, with the array looking in the direction of the plane wave [11], [21]. Under these conditions

$$
p_{nm} = b_n Y_n^{m^*}(\Omega_0) \qquad w_{nm}^* = \frac{c_n}{b_n} Y_n^m(\Omega_0) \quad (26)
$$

where Ω_0 is the arrival direction of the unit amplitude plane-wave and the choice of w_{nm}^* represents a beamforming array which reduces to a plane-wave decomposition array for $c_n = 1$ as discussed in Section IV. The signal power under these conditions can be written using (5) as

$$
|y_s|^2 = \left| \sum_{n=0}^{N} c_n \sum_{m=-n}^{n} Y_n^m(\Omega_0) Y_n^{m^*}(\Omega_0) \right|^2
$$
$$
= \left| \sum_{n=0}^{N} c_n \frac{2n+1}{4\pi} \right|^2 \quad (27)
$$

which reduces to $(N+1)^4/(4\pi)^2$ when $c_n = 1$. The power of the output signal in this case increases with increasing order N. The measurement noise at the array output is calculated assuming spatially uncorrelated input noise with a unit variance, and is given by

$$
E[|y_e|^2] = \sum_{j=1}^{M} \alpha_j^2 |w(\Omega_j)|^2
$$
$$
= \sum_{j=1}^{M} \alpha_j^2 \left| \sum_{n=0}^{N} \sum_{m=-n}^{n} \frac{c_n}{b_n} Y_n^m(\Omega_0) Y_n^{m^*}(\Omega_j) \right|^2
$$
$$
= \sum_{j=1}^{M} \alpha_j^2 \left| \sum_{n=0}^{N} \frac{c_n}{b_n} \frac{2n+1}{4\pi} P_n(\cos\Theta_j) \right|^2 \quad (28)
$$

where Θ_j is the angle between Ω_0 and Ω_j. The noise variance has a simpler expression for a sampling scheme with equal

weights $\alpha_j = 4\pi/M$ (note that this choice satisfies $\sum_{j=1}^{M} \alpha_j = \int_{\Omega} d\Omega = 4\pi$), computed using (7) as

$$E[|y_e|^2] = \sum_{n=0}^{N} |c_n|^2 \left\{ \frac{1}{M} \frac{2n+1}{|b_n|^2} \right\} \tag{29}$$

where the term in the curly brackets is the un-weighted contribution of order n to the overall measurement noise. It is expected from Fig. 1 that at low frequencies the high orders will have dominant contribution. Equation (29) also shows that the noise variance decreases as the number of microphones, M, increases, assuming the other parameters, such as N, remain unchanged. The white noise gain (WNG) [21], calculated as $1/E_e$ from (25), (27), and (29), represents the improvement in signal to noise ratio (SNR) at the array output compared to the input. This assumes that the SNR at the input for this case is unity, which is true for an open sphere and approximately true for a rigid sphere, as detailed in the Appendix. In addition to denoting improvement in SNR, the WNG is also used as a general measure for robustness [21]. However, to predict the explicit effect of factors such as microphone positioning errors and aliasing, additional measures as presented below are employed.

The aliasing error is computed in this case using (24) and (26) as

$$|y_a|^2 = \left| \sum_{n=0}^{N} \sum_{m=-n}^{n} \sum_{n'=N+1}^{\infty} \sum_{m'=-n'}^{n'} c_n \frac{b_{n'}}{b_n} Y_{n'}^{m'*}(\Omega_0) \right.$$
$$\left. \times Y_n^m(\Omega_0) \epsilon_{a_{nmn'm'}} \right|^2. \tag{30}$$

A more explicit and computationally efficient expression can be derived by combining (5), (22), and (30)

$$|y_a|^2 = \left| \sum_{n=0}^{N} c_n \left\{ \sum_{n'=N+1}^{\infty} \frac{b_{n'}}{b_n} \frac{2n+1}{4\pi} \frac{2n'+1}{4\pi} \right. \right.$$
$$\left. \left. \times \sum_{j=1}^{M} \alpha_j P_n(\cos\Theta_j) P_{n'}(\cos\Theta_j) \right\} \right|^2 \tag{31}$$

where Θ_j is the angle between Ω_0 and Ω_j, and the term in the curly brackets denotes the contribution of order n to the overall aliasing error. Finally, microphone positioning error in this case is computed from (24) and (26) as

$$|y_\Omega|^2 = \left| \sum_{n=0}^{N} \sum_{m=-n}^{n} \sum_{n'=0}^{\infty} \sum_{m'=-n'}^{n'} c_n \frac{b_{n'}}{b_n} Y_{n'}^{m'*}(\Omega_0) \right.$$
$$\left. \times Y_n^m(\Omega_0) \epsilon_{\Omega_{nmn'm'}} \right|^2 \tag{32}$$

with a more explicit and computationally efficient expression derived by combining (5), (23), and (32)

$$|y_\Omega|^2 = \left| \sum_{n=0}^{N} c_n \left\{ \sum_{n'=0}^{\infty} \frac{b_{n'}}{b_n} \frac{2n+1}{4\pi} \frac{2n'+1}{4\pi} \right. \right.$$
$$\left. \left. \times \sum_{j=1}^{M} \alpha_j P_n(\cos\Theta_j) \left[P_{n'}\left(\cos\Theta_j'\right) - P_{n'}(\cos\Theta_j) \right] \right\} \right|^2 \tag{33}$$

Array No.	Sampling	N	N_{max}	M	Ω_0	σ^2	$|\Delta|_{max}$
1	Equiangle	2	2	36	$(\pi/2, 0)$	1	$\approx 0.3°$
2	Gaussian	2	2	18	$(\pi/2, 0)$	1	$\approx 0.3°$
3	Uniform	2	2	12	$(\pi/2, 0)$	1	$\approx 0.3°$
4	Equiangle	2	2	36	$(\pi/2, 0)$	1	$\approx 0.3°$
5	Gaussian	2	3	32	$(\pi/2, 0)$	1	$\approx 0.3°$
6	Uniform	2	4	36	$(\pi/2, 0)$	1	$\approx 0.3°$
7	Uniform	6	6	84	$(\pi/2, 0)$	10^{-4}	$\approx 0.3°$

where Θ_j' is the angle between Ω_0 and the actual microphone position Ω_j', and the term in the curly brackets denotes the contribution of order n to the overall positioning error. The error due to inaccurate positioning of the microphones will depend on the nature of the inaccuracy. For example, if spatial sampling is achieved by rotating a single microphone by motors along θ and ϕ, then the accuracy in the motor control system will result in variation in the angles, e.g., $\theta_j' = \theta_j \pm \Delta$ and $\phi_j' = \phi_j \pm \Delta$. However, if the microphones are positioned more permanently on a rigid or open sphere, then the accuracy will be given as some distance away from the desired point on the sphere surface. This can be translated to $\theta_j' = \theta_j \pm \Delta$ and $\phi_j' = \phi_j \pm \Delta/\sin\theta$ to account for the smaller apparent radius near the poles. The latter inaccuracy is employed in the examples below.

IX. SIMULATION EXAMPLES

Several illustrative examples to the use of the analysis method developed in this paper are presented in this section. A more comprehensive analysis of array performance and the behavior of the various errors for different sampling schemes is beyond the scope of this paper but proposed for future research. The array configurations employed and other simulation details are summarized in Table I. The positions Ω_j for the equiangle sampling scheme are defined by all combination of $\theta = [15, 45, 75, 105, 135, 175]$ and $\phi = [0, 60, 120, 180, 240, 300]$ (in degrees). The sampling coordinates for the two configurations with Gaussian sampling were taken from [16], and those for the three nearly uniform sampling configurations taken from [18] (denoted designs 3.12.5, 3.36.8, and 3.84.12). All sampling configurations satisfied (7) with an error less then 4×10^{-13}.

Fig. 2(a) shows examples of the inverse WNG, $E_e = E[|y_e|^2]/|y_s|^2$, for microphone arrays 1–3 in Table I, calculated using (27) and (28). The figure shows that lowest noise or best WNG is achieved around $kr = N = 2$, which then degrades for both lower and higher kr. This is a reflection of the behavior of the $1/b_n$ term in (28) [see Fig. 1(a)]. This is also why significant degradation in the WNG would be expected for open-sphere arrays around frequencies where the b_n terms are zero. The figure
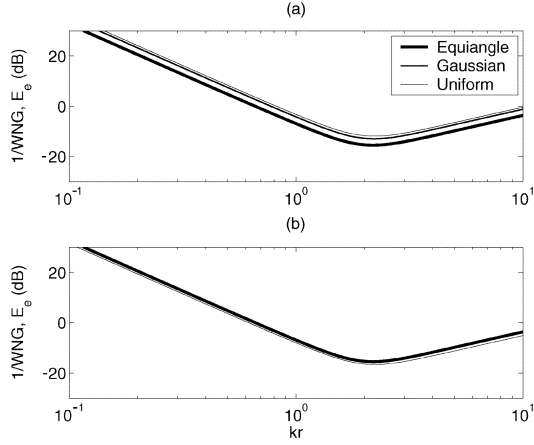
Fig. 2. Reciprocal of the white noise gain (WNG) for (a) arrays 1–3 in Table I, and (b) arrays 4–6 in Table I.
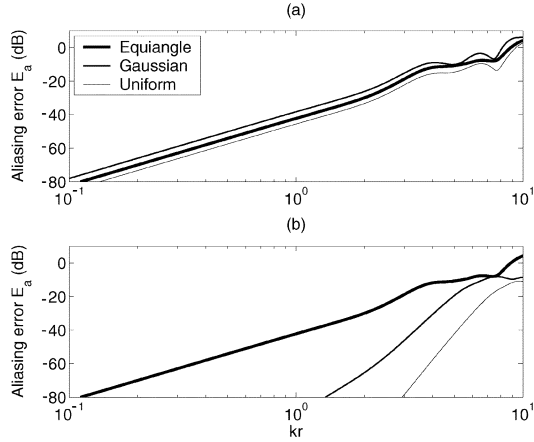

Fig. 3. Aliasing error for (a) arrays 1–3 in Table I, and (b) arrays 4–6 in Table I.

also shows that the equiangle sampling scheme has the lowest noise variance and the nearly uniform scheme the highest noise variance. A possible explanation for this difference is provided by Fig. 2(b), showing results for arrays 4–6 in Table I. Although Table I shows that these arrays allow for different maximum N, in Fig. 2(b) the same $N = 2$ was used for all three arrays. Therefore, the three arrays have the same order N, similar number of microphones M, but different sampling schemes. The figure illustrates that the noise variance is similar for all three cases, suggesting that the noise variance, and therefore the WNG, depend on the order and number of microphones, and less on the sampling configuration. Therefore, (29) can be used as a general predictor of WNG. The results also suggest that spatial over-sampling could be one way to improve the WNG.

Fig. 3(a) shows the relative aliasing error, $E_a = |y_a|^2/|y_s|^2$, calculated for arrays 1–3 in Table I using (27) and (31). In this case all three arrays show similar aliasing error. Fig. 3(b) presents the same analysis for arrays 4–6 in Table I, all designed with $N = 2$ and similar number of microphones. In this case the array with the nearly uniform sampling has significantly less aliasing error, and so does the array with the Gaussian sampling. This is because these array, although have similar number of microphones are actually designed to work with higher orders ($N_{\max}$ in Table I). This illustrates the advantage of the nearly
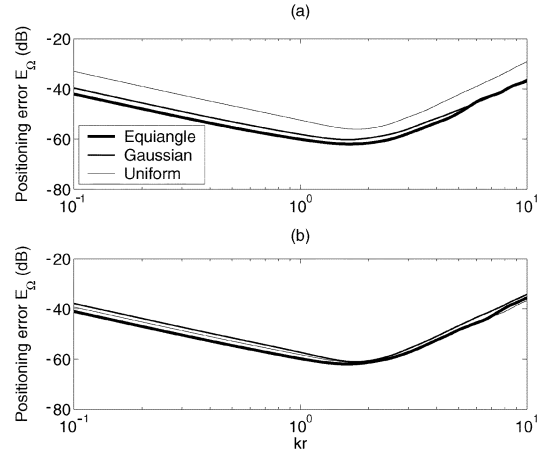

Fig. 4. Positioning error for (a) arrays 1–3 in Table I, and (b) arrays 4–6 in Table I.

uniform sampling. Although arrays with similar N and M will have similar WNG, as illustrated in Fig. 2(b), the array with the nearly uniform sampling will exhibit lower aliasing error, therefore improving high frequency performance.

Fig. 4(a) and (b) show the relative positioning error, $E_\Omega = |y_\Omega|^2/|y_s|^2$, calculated for arrays 1–3 and 4–6 in Table I, respectively, using (27) and (33). The positioning error Δ as defined in Section VIII, is uniformly distributed in the range $|\Delta| \leq 0.005 \text{ rad} \approx 0.3°$. This variation is equivalent to about $\pm 1 \text{ mm}$ on a sphere of radius 20 cm. The results presented were calculated as an average over 50 realizations of the random positioning errors. The figure shows that the errors are less then about -40 dB, which suggests that a relatively small error can be achieved with practical constrains on the accuracy of microphone positioning. The behavior of the positioning errors in Fig. 4(a) and (b) is similar to the behavior of the inverse WNG as presented in Fig. 2(a) and (b), therefore supporting the use of the WNG as a general robustness measure [21]. The difference of about 6 dB per octave between the two errors originates from the additional summation term over n' (with $b_{n'}$ in the numerator) in (33) compared to (29) (see this difference more clearly in the next example).

The final simulation example illustrates the use of the theoretical results developed in this paper. A microphone array of order $N = 6$ was designed with a nearly uniform sampling scheme [18], as detailed in Table I for array no. 7. A normally distributed measurement noise with variance -40 dB below the measured pressure is added to the microphones input, and a uniformly distributed noise in the microphones locations is also included as detailed in Table I. The sound pressure around the array was constructed from its spherical Fourier transform with $N = 50$ (simulating $N = \infty$), and then array output including the measurement noise and microphone positioning errors was simulated using (12). The various error predictions and the corresponding noise-to-signal ratios are then calculated using (27), (29), (31), and (33), with the prediction for positioning error calculated as an average over 50 realizations.

Fig. 5(a) shows the overall noise, simulated by the array processing, and the individual error predictions. The figure illustrates how the overall noise is decomposed into its contributions,
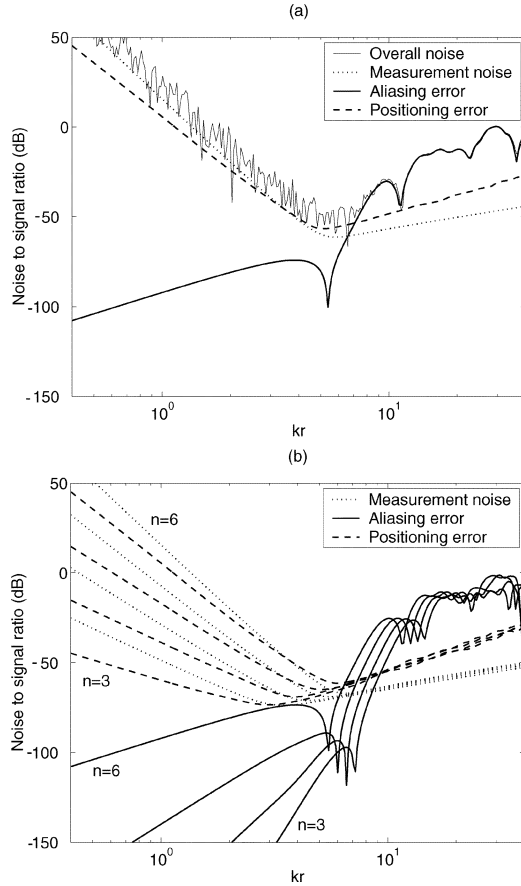
Fig. 5. Error analysis for order $N = 6$ array detailed in Table I as array no. 7. (a) Analysis of overall error with "Overall noise" denoting noise floor measured at the array output, and "Measurement noise," "Aliasing error," and "Positioning error" denoting the theoretical predictions of these effects. (b) Contributions of terms of orders $n = 3 \ldots 6$ to the measurement noise, aliasing error and microphone positioning error.

providing information on the main causes of the noise at various frequencies. At high frequencies, where $kr \geq N$, aliasing error is dominant, while at the lower frequencies, both measurement noise and positioning errors contribute almost equally. This is a useful design information, since for this example it clear that in order to reduce the noise at the low frequencies one would require to reduce the input noise level and to improve the accuracy of microphone positioning. Further, by modifying parameters such as input noise level, positioning accuracy and number of microphones, one can investigate the effect of the various factors on the overall noise floor at each frequency.

Fig. 5(a) shows that the operating frequency bandwidth is limited by aliasing at the higher frequencies and measurement errors at the lower frequencies. The overall achieved bandwidth in this case is not wide, and will not be sufficient for most applications. One way to increase the operating bandwidth, as suggested in [3], is to employ different sphere radiuses for different frequency ranges. This, however, comes at the cost of increased system complexity. Another alternative is to reduce the spatial resolution at the lower frequencies by employing a lower order N, therefore reducing the noise imposed by the higher order $1/b_n$ terms [2]. Methods for the design of broad bandwidth, high-order spherical arrays are subject of current research.

Fig. 5(b) shows the contribution of the upper four orders $n = 3 \ldots 6$ to the three noise-to-signal ratios, calculated by evaluating the terms in the curly brackets of (29), (31), and (33) for each order n. In all three cases the higher order ($n = 6$) contributed most significantly at the lower frequencies ($kr < 6$), while at the higher frequencies the contributions of the various orders are more similar.

X. CONCLUSION

This paper presented a theoretical and simulation analysis of the measurement errors of spherical microphone arrays. The effects of array order, input noise, microphone positioning accuracy and spatial aliasing were formulated individually. This facilitates efficient selection of array design parameters such as the number of microphones, transducer noise, and positioning accuracy, and allows analysis of the effect of these parameters for existing arrays. The framework of analysis developed here can therefore be useful in the design and performance analysis of practical spherical microphone arrays.

APPENDIX
SNR AT THE ARRAY INPUT

Assuming spatially uncorrelated input noise e_j with variance $\sigma^2 = 1$, then the spatial average of the noise variance, $(1/4\pi) \sum_{j=1}^{M} \alpha_j \sigma^2$, is unity for any M and sampling scheme. Note that the weight α_j has been included to support the Parseval notation as in (8), while it is also assumed that $(1/4\pi) \sum_{j=1}^{M} \alpha_j = (1/4\pi) \int_{\Omega} d\Omega = 1$ for the chosen sampling scheme. For a plane-wave sound-field propagating from Ω_0 and an open sphere the spatial average of the input signal power given by $(1/4\pi) \sum_{j=1}^{M} \alpha_j |p(\Omega_j)|^2$ is unity, but for a rigid sphere it will be affected by scattering and will depend on the sampling scheme. As a rough approximation of the average signal power the normalized integral over the entire sphere can be used

$$\frac{1}{4\pi} \sum_{j=1}^{M} \alpha_j \left| p(\Omega_j) \right|^2 \approx \frac{1}{4\pi} \int_{\Omega \in S^2} \left| p(\Omega) \right|^2 d\Omega$$

$$= \frac{1}{4\pi} \int_{\Omega \in S^2} \left| \sum_{n=0}^{\infty} \sum_{m=-n}^{n} b_n(kr) Y_n^{m^*}(\Omega_0) \right.$$
$$\left. \times Y_n^m(\Omega) \right|^2 d\Omega$$

$$= \frac{1}{4\pi} \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \left| b_n(kr) Y_n^{m^*}(\Omega_0) \right|^2$$

$$= \frac{1}{(4\pi)^2} \sum_{n=0}^{\infty} |b_n(kr)|^2 (2n+1) \qquad (34)$$

where (4) and (5) were employed in the derivation. Numerical evaluation of (34) shows that the average input signal power is unity for a rigid sphere at $kr \ll 1$, and increases up to 3 dB at $kr > 1$. Therefore, the SNR at the array input can be approximated by unity with an error smaller then 3 dB.

REFERENCES

[1] J. Meyer and G. W. Elko, "A spherical microphone array for spatial sound recordings," *J. Acoust. Soc. Amer.*, vol. 111, no. 5.2, pp. 2346–2346, 2002.

[2] ——, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," in *Proc. ICASSP*, vol. II, 2002, pp. 1781–1784.

[3] T. D. Abhayapala and D. B. Ward, "Theory and design of high order sound field microphones using spherical microphone array," in *Proc. ICASSP*, vol. II, 2002, pp. 1949–1952.

[4] G. Weinreich and E. Arnold, "Method for acoustic radiation fields," *J. Acoust. Soc. Amer.*, vol. 68, no. 2, pp. 404–411, 1980.

[5] B. Rafaely, "Plane-wave decomposition of the pressure on a sphere by spherical convolution," *J. Acoust. Soc. Amer.*, vol. 116, no. 4, pp. 2149–2157, 2004.

[6] ——, "Decomposition of reverberant sound fields into plane waves using microphone arrays," in *Proc. 5th Int. Workshop Microphone Array Systems, Theory Practice*, Erlangen-Nuremberg, Germany, May 15–16, 2003.

[7] ——, "Plane-wave decomposition by spherical-convolution microphone array," *J. Acoust. Soc. Amer.*, vol. 115, no. 5.2, pp. 2578–2578, 2004.

[8] 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. Amer.*, vol. 112, no. 5, pp. 1980–1991, 2002.

[9] J. R. Driscoll and D. M. Healy, Jr, "Computing Fourier transforms and convolutions on the 2-sphere," *Adv. Appl. Math.*, vol. 15, pp. 202–250, 1994.

[10] E. G. Williams, *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. New York: Academic, 1999.

[11] J. Meyer, "Beamforming for a circular microphone array mounted on spherically shaped objects," *J. Acoust. Soc. Amer.*, vol. 109, no. 1, pp. 185–193, 2001.

[12] E. De Witte, H. D. Griffiths, and P. V. Brennan, "Phase mode processing for spherical anttena arrays," *Electron. Lett.*, vol. 30, no. 20, pp. 1430–1431, 2003.

[13] G. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 5th ed. San Diego, CA: Academic, 2001.

[14] L. E. Kinsler, A. R. Frey, A. B. Coppens, and J. V. Sanders, *Fundementals of Acoustics*. New York: Wiley, 1982.

[15] M. J. Mohlenkamp, "A fast transform for spherical harmonics," *J. Fourier Analy. Applicat.*, vol. 5, no. 2/3, pp. 159–184, 1999.

[16] V. I. Krylov, *Approximate Calculation of Integrals*. New York: Macmillan, 1962.

[17] M. A. Armstrong, *Groups and Symmetry*. New York: Springer-Verlag, 1988.

[18] R. H. Hardin and N. J. A. Sloane, "McLaren's improved snub cube and other new spherical designs in three dimensions," *Discrete Computational Geometry*, vol. 15, pp. 429–441, 1995.

[19] V. I. Lebedev and A. L. Skorokhodov, "Quadrature formulas of orders 41, 47, and 53 for the sphere," *Russ. Acad. Sci. Dokl. Math.*, vol. 45, no. 3, pp. 587–592, 1992.

[20] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*. New York: Academic, 1980.

[21] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. Acoust. Speech Sig. Proccessing*, vol. SSP-35, no. 5, pp. 1365–1376, Oct. 1987.

Boaz Rafaely (SM'0l) received the B.Sc. degree (cum laude) in electrical engineering from Ben-Gurion University, Ben-Gurion, Israel, in 1986, the M.Sc. degree in biomedical engineering from Tel-Aviv University, Tel-Aviv, Israel, in 1994, and the Ph.D. degree from the Institute of Sound and Vibration Research (ISVR), Southampton University, Southampton, U.K., in 1997.

He was appointed Lecturer in 1997 and Senior Lecturer in 2001 at the ISVR, working on active control of sound and organizing the modular M.Sc. in applied digital signal processing. In 2002, he spent six months as a Visiting Scientist with the Sensory Communication Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, investigating speech enhancement for hearing aids. He then joined the Department of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer-Sheva, Israel, as a Senior Lecturer in 2003, where he is currently teaching acoustics and signal processing and establishing an acoustics laboratory to investigate sound fields by microphone arrays.