50. Microphone Arrays

G. W. Elko, J. Meyer

This chapter introduces various types of microphone array beamforming systems and discusses some of the fundamental theory of their operation, design, implementation, and limitations. It is shown that microphone arrays have the ability to offer directional gains that can significantly improve the quality of signal pickup in reverberant and noisy environments.

Hands-free audio communication is now a major feature in mobile communication systems as well as audio and video conferencing systems. One problem that becomes evident to users of these systems is the decrease in communication quality due to the pickup of room reverberation and background noise. In the past, this problem was dealt with by using microphones placed close to the desired talker or source. Although this simple solution has proven to be quite effective, it also has its drawbacks. First, it is not always possible or desirable to place the microphone very close to the talker's mouth. Second, by placing the microphone close to the talker's mouth, one has to deal with rapid level variation as the talker moves his or her mouth relative to the microphone. Third is the negative impact of speech plosives (airflow transients generated by plosive sounds) and forth, microphone structure-borne handling noise has a detrimental effect. Finally, for directional microphone elements, there is also a nearfield proximity effect (where the frequency response of the microphone is modulated by the relative position of the microphone to the mouth). With these issues in mind, it is of interest to investigate other potential solutions. One solution is to use beamforming microphone arrays, which can offer significant directional gain so as to result in similar audio performance to that of closely placed microphones.

50.1 Microphone Array Beamforming

The propagation of acoustic waves in space is a function of both space and time coordinates. In general, the mathematical representation of a propagating acoustic wave is a four-dimensional function: three spatial dimensions and one time variable. Acoustic wave propagation can be modeled by a linearized scalar acoustic wave equation that describes the relationships between the physical quantities of acoustic pressure and density variation of the medium [50.1],

$$\nabla^2 p = \frac{1}{c^2}\frac{\partial^2 p}{\partial t^2}, \tag{50.1}$$

where p is the instantaneous acoustic pressure fluctuation of the sound and c is the propagation speed of sound in the medium. There are a few major underlying assumptions in the derivation of (50.1), but this simplified linear model is adequate for laying the foundation for array beamforming of acoustic signals. The dimen-

sionality of the wave equation can easily be seen from (50.1), where p is a function of three space variables and one time variable. Therefore the scalar acoustic pressure field in space can be represented as $p(\boldsymbol{r}, t)$, where $\boldsymbol{r}$ is the measurement position in the acoustic field and t is the time dependence. The scalar acoustic pressure field must satisfy (50.1) at all points in space. By using the spatial Fourier transform, the acoustic field can also equivalently be represented in the wavevector–frequency domain. This representation has been widely used in the fields of geophysics, structural, underwater and aeroacoustics, and can be quite useful in the analysis and design of microphone array beamformers.

Applying the four-dimensional Fourier transform to the time–space scalar acoustic field yields,

$$P(\boldsymbol{k}, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(\boldsymbol{r}, t)\, \mathrm{e}^{-\mathrm{i}(\omega t - \boldsymbol{k}^{\mathrm{T}}\boldsymbol{r})}\, \mathrm{d}\boldsymbol{r}\, \mathrm{d}t \,, \tag{50.2}$$

where the acoustic pressure P is capitalized to indicate a transform to the frequency domain, $\boldsymbol{k}$ is the wavevector and the superscript 'T' represents the transpose operator. It may seem that this transformation has unnecessarily complicated the description of the scalar acoustic pressure field, but as will be seen later, it has transformed the field description into a form that makes the analysis of acoustic field space–time functions analogous to the field of multidimensional signal processing. Since the Fourier transform is a linear transformation, one can write an inverse relationship,

$$p(\boldsymbol{r}, t) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P(\boldsymbol{k}, \omega)\, \mathrm{e}^{\mathrm{i}(\omega t - \boldsymbol{k}^{\mathrm{T}}\boldsymbol{r})}\, \mathrm{d}\boldsymbol{k}\, \mathrm{d}\omega \,. \tag{50.3}$$

Equation (50.3) directly shows that any acoustic field $p(\boldsymbol{r}, t)$ can be represented as an infinite number of propagating plane waves with appropriate complex weighting. This interpretation can easily be seen from the equation for a single propagating plane wave with frequency ω_0 and wavevector $\boldsymbol{k}_0$,

$$p(\boldsymbol{r}, t) = A_0\, \mathrm{e}^{-\mathrm{i}(\omega_0 t - \boldsymbol{k}_0^{\mathrm{T}}\boldsymbol{r})} \,, \tag{50.4}$$

where A_0 is the plane-wave amplitude. Note that for generality, one can allow A_0 to be complex, but for this discussion it is assumed that A_0 is real. The wavenumber–frequency spectrum of this single propagating plane wave is simply,

$$P(\boldsymbol{k}, \omega) = A_0 \delta(\boldsymbol{k} - \boldsymbol{k}_0)\delta(\omega - \omega_0) \,. \tag{50.5}$$

Equation (50.5) shows the Fourier mapping of infinite continuous functions in both time and space to be a point in the wavevector–frequency space.

Now consider a general weighting function of the space–time pressure distribution with a four-dimensional linear shift-invariant filter having an impulse response $h(\boldsymbol{r}, t)$. Then, the output signal is the convolution of the space–time acoustic pressure signal and the spatial weighting function such that,

$$y(\boldsymbol{r}, t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(\boldsymbol{r} - \hat{\boldsymbol{r}}, t - \tau) p(\hat{\boldsymbol{r}}, \tau)\, \mathrm{d}\hat{\boldsymbol{r}}\, \mathrm{d}\tau \,. \tag{50.6}$$

Equation (50.6) is a filtering operation that can equivalently be represented in the wavevector–frequency domain as,

$$Y(\boldsymbol{k}, \omega) = H(\boldsymbol{k}, \omega) P(\boldsymbol{k}, \omega) \,. \tag{50.7}$$

Equation (50.6) and (50.7) show the direct analogy between the space–time wavevector–frequency representation of spatial filtering and the well-known results from multidimensional linear systems theory. Thus, by using specific spatial weighting functions, one can filter the acoustic field to investigate the propagating directions, amplitude, and phases of plane waves traveling in space. Filtering the spatiotemporal acoustic scalar pressure field is known in the field of array signal processing as *beamforming*. Thus, if one wants to *look* at plane waves propagating from the direction of the unit vector $\boldsymbol{k}_0/\|\boldsymbol{k}_0\|$, one would design a filter such that,

$$H(\boldsymbol{k}, \omega) = \delta(\boldsymbol{k} - \boldsymbol{k}_0) G(\omega) \,, \tag{50.8}$$

where $G(\omega)$ is the desired frequency response to waves propagating in the direction of the wavevector $\boldsymbol{k}_0$.

Although it may seem limiting to use plane waves as the underlying basis functions for the representation of a sound-field, it should be noted that other orthogonal representations such as spherical and cylindrical basis functions can themselves be represented as a series of plane waves.

The representation of the wavevector–frequency domain is a natural framework for analysis of the spatial and frequency filtering of beamforming arrays, as will be seen in the following.

In spatial filtering by beamforming, one deals with spatial apertures. An aperture is a region over which energy is received. Where an aperture can either be continuous, as in parabolic dishes, or discretely realized as in microphone arrays with multiple microphones, mixtures of continuous and discrete apertures are possible.

In fact, even discrete microphone arrays are a mixture since no microphone is a perfect point receiver. Although sampled aperture systems must generally perform more signal processing, they offer several advantages over continuous aperture systems. A main advantage in using sampled apertures is that they allow for the possibility of using digital signal processing on the individual array signals. Arrays can be electronically steered, can form multiple simultaneous beams, and can be made adaptive purely by electronic means.

50.1.1 Delay-and-Sum Beamforming

Delay-and-sum beamforming, also known as *classical beamforming*, is one of the simplest and oldest techniques for realizing directional array systems. Although not fundamentally limited in bandwidth, early delay-and-sum arrays were used in *narrowband* operation to focus arrays onto a particular point or direction. Since a time delay for narrowband applications can be accomplished with a unique phase shift for each element, narrowband beamforming is typically implemented with phase shifts and is therefore commonly referred to as *phased-array* beamforming. One typical application for microphone arrays is to pick up speech or acoustic signals that are wideband (the desired signals cover many octaves in bandwidth). Therefore most microphone arrays for speech acquisition are implemented as delay-sum beamformers since they are inherently wideband.

Both continuous and discrete beamformers can be seen as implementations of wavevector–frequency filtering as discussed in the previous section. It should be noted that these two beamformers are not mutually exclusive; one can, for instance, have a discrete set of continuous transducers so that the beamformer is a mix of continuous and discrete sampling. Of special interest here is the spatial and frequency response of a finite array of microphones that sample the acoustic field. A classic delay-sum beamformer uses the sum of weighted and delayed samples from an array of N discretely sampled points of the pressure field. In general, a delay-sum beamformer output y is formed as,

$$y(t) = \sum_{n=1}^{N} w_n s(\boldsymbol{r}_n, t - \tau_n) \,, \tag{50.9}$$

where the weights w_n are real and the time delay τ_n is applied to the measured signals $s(\boldsymbol{r}_n, t)$, at microphone n. If the incident field is a planewave with wavevector $\boldsymbol{k}_0$, amplitude A_0 and frequency ω_0, then,

$$s(\boldsymbol{r}_n, t) = A_0 \, \mathrm{e}^{\mathrm{i}(\omega_0 t - \boldsymbol{k}_0^{\mathrm{T}} \boldsymbol{r}_n)} \,, \tag{50.10}$$

and

$$y(t) = A_0 \, \mathrm{e}^{\mathrm{i}\omega_0 t} \sum_{n=1}^{N} w_n \, \mathrm{e}^{-\mathrm{i}(\omega_0 \tau_n + \boldsymbol{k}_0^{\mathrm{T}} \boldsymbol{r}_n)} \,. \tag{50.11}$$

The real weights w_n scale the signals measured at the positions $\boldsymbol{r}_n$. Thus, from (50.9), the origin of the term *delay-and-sum* can be seen as simply a description of how the classical delay-sum beamformer is realized.

For a causal delay-sum beamformer, the exponent in (50.11) must be less than or equal to zero for all n. Typically, one element position is selected as the spatial origin. This reference element position therefore defines the vectors $\boldsymbol{r}_n$. Since the array can be steered to any direction in space and one can select any position to define $\boldsymbol{r}_n$, an additional delay may be required in order to maintain causality. (Recall that $\boldsymbol{k}_0^{\mathrm{T}} \boldsymbol{r}_n$ can be either positive or negative depending on the direction of the incident wave relative to the array.) If an end element is selected as the spatial origin, then the causal delay is equal to the maximum time that an acoustic wave takes to transit the array. Smaller causal delays are also possible depending on which microphone position in the array is chosen as the spatial origin reference and the maximum desired steering angle.

If we set the causal time delay equal to T_0, then the output for an incident plane wave with angular frequency ω_0 and wavevector $\boldsymbol{k}_0$ can be written as,

$$y(t) = A_0 \, \mathrm{e}^{\mathrm{i}\omega_0 t} \sum_{n=1}^{N} w_n \, \mathrm{e}^{-\mathrm{i}\left[\omega_0 (T_0 + \tau_n) + \boldsymbol{k}_0^{\mathrm{T}} \boldsymbol{r}_n\right]} . \tag{50.12}$$

The maximum plane-wave response corresponds to the direction where the delays τ_n compensate for the propagation delay $\boldsymbol{k}_0^{\mathrm{T}} \boldsymbol{r}_n$ of a plane wave propagating with wavevector $\boldsymbol{k}_0$. This is done by setting the values of τ_n such that these delays offset the time lead (or lag) of a desired direction plane wave propagating over the array. Thus, setting the delays to steer the array to the direction of $\boldsymbol{k}_0$, means that the last two terms in the exponential in (50.12) cancel out. The delayed and summed output $y(t)$ is now such that all microphone signals propagating from the desired direction are added in phase. This direction corresponds to a maximum amplitude output for any selection of τ_n. Plane waves propagating from directions other than $\boldsymbol{k}_0$ will result in the addition of position-dependent phase variations, and as a result, the output amplitude will be smaller due to destructive interference.

To summarize, classical delay-and-sum beamforming uses delays between each array element that compensate for differences in the propagation delay of

the desired signal direction across the array. Signals originating from a desired direction (or location if the source is in the near field) are summed in phase, while other signals undergo destructive interference. By adjusting the weights of the delay-sum beamformer, the shape of the beam and sidelobes as well as the position of the nulls of an array can be controlled.

Uniformly Spaced Linear Arrays

While array geometries are in principle arbitrary, certain configurations are especially useful and amenable to mathematical analysis. For example, consider a simple array geometry of uniformly spaced collinear elements as shown in Fig. 50.1. Assume that an odd number of elements, numbered from $-N$ to N, are positioned on the x-axis such that element n has coordinate nd, and d is the interelement spacing. Thus, *endfire* directions lie on the x-axis, while *broadside* directions lie in the y–z plane. Much as linear time-domain systems are analyzed by their response to complex exponential time functions, an array response can be considered in terms of its response to complex exponential plane waves (as shown above).

For a collinear delay-and-sum beamformer steered towards a direction making an angle θ_0 measured clockwise from the y-axis, the relative time delays τ_n are,

$$\tau_n = n\frac{d}{c}\sin\theta_0, \tag{50.13}$$

where θ_0 is the *steering* direction, and represents the direction from which plane waves are received with maximum response.

Now, consider the unit-amplitude plane wave of frequency ω_0, arriving from a direction θ (not necessarily equal to θ_0). This wave is represented by the expression,

$$p(\boldsymbol{r}, t) = \mathrm{e}^{\mathrm{i}(\omega_0 t - \boldsymbol{k}_0^{\mathrm{T}}\boldsymbol{r})}\,. \tag{50.14}$$

Note that it has been assumed that the microphones are located along a line so the vector $\boldsymbol{r}$ points along the axis of the array. Since the array is axisymmetric about the axis of the array, the components of the wavevector can be assumed, without loss of generality, to lie in the x–y plane. With this assumption, one only needs to investigate the array response in the plane. The wavevector can be written into its three orthogonal components as,

$$\boldsymbol{k}_0 = -\frac{\omega_0}{c}\begin{pmatrix}\sin\theta\\ \cos\theta\\ 0\end{pmatrix}\,. \tag{50.15}$$

Substituting the wave field for a planewave and time delay expressions into the beamformer output expression yields,

$$y(t) = \mathrm{e}^{\mathrm{i}\omega_0 t}\sum_{n=-N}^{N} w_n\,\mathrm{e}^{\mathrm{i}\frac{\omega_0}{c}nd(\sin\theta-\sin\theta_0)}\,. \tag{50.16}$$

This expression can be simplified by removing the time-varying $\mathrm{e}^{\mathrm{i}\omega_0 t}$ term and substituting $u = \sin\theta$ and $u_0 = \sin\theta_0$ rather than by the angle itself. This convention has several advantages beyond mere compactness. First, a change of steering direction amounts to a single linear translation in u-space. Thus, the shape of the beam measured in the u-domain does not change with array steering angle, while beams measured in θ space change as a function of the steering direction. Second, one can logically divide the description of the array function into two regions: a *visible* and *invisible* region [50.2]. The visible region corresponds to values of u that have real angles in space and corresponds mathematically to $|u| \le 1$. Values of $|u| > 1$ correspond to angles that are not physical and is therefore referred to as the invisible region.

Making the substitutions for $\sin\theta$ and $\sin\theta_0$ yields an expression for the *array response*,

$$H(u, \omega) = \sum_{n=-N}^{N} w_n\,\mathrm{e}^{\mathrm{i}\frac{\omega}{c}nd(u-u_0)}. \tag{50.17}$$

Note that the description of the array response given in (50.17) has been written as a function of the two independent variables u and ω instead of the variables $\boldsymbol{k}$, t, and ω. This change of variables allows one to better visualize the response of the array in terms of angle and frequency.

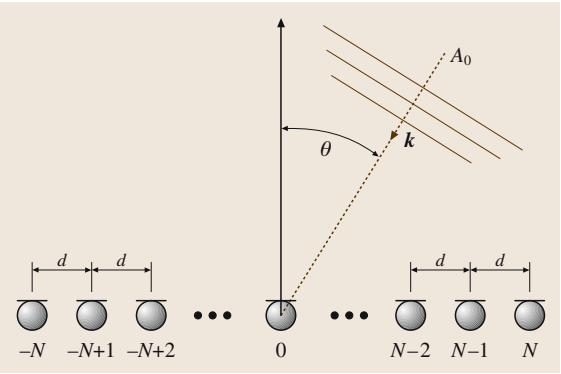

Fig. 50.1 Schematic of uniformly spaced linear array containing $2N+1$ microphones spaced by d

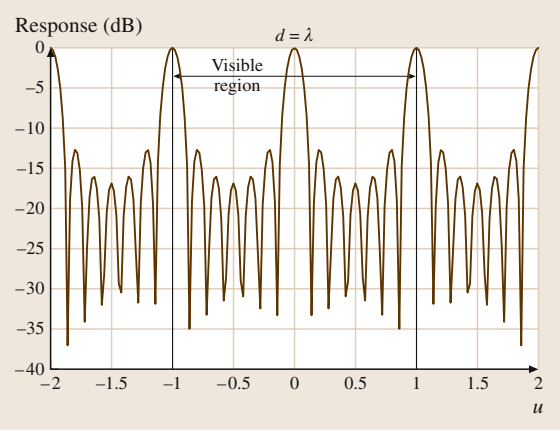

Fig. 50.2 Response of a seven-element uniformly spaced and weighted array as a function of the dimensionless variable u for element spacing equal to the wavelength of the incident acoustic signal

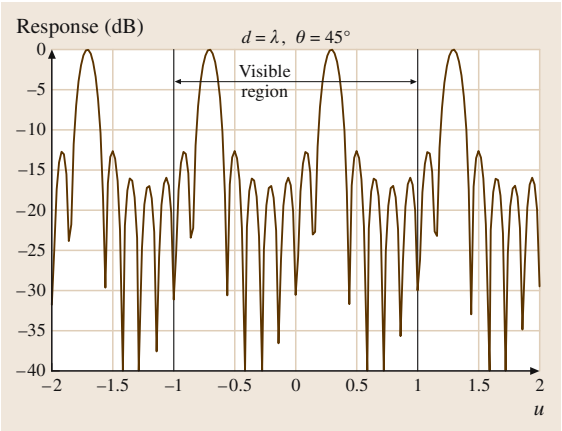

Fig. 50.3 u-response of a seven-element uniformly spaced and weighted array steered 45° as a function of the dimensionless variable u for element spacing equal to the wavelength of sound

The array response as expressed in (50.17) is a function in a two-dimensional u–ω space. For any real constant K the following identity holds,

$$H\left[\left(\frac{u}{K}+u_0\right), K\omega\right] = H(u+u_0, \omega)\,. \tag{50.18}$$

From (50.18), it can be seen that the spatial extent (beamwidth) of the array becomes commensurately smaller as the frequency increases directly proportionally to the frequency ω. Thus if the incident wave frequency is doubled, the beamwidth of the array is cut in half and vice versa. Another important aspect of the array function is that it is periodic in u for any given ω,

$$H(u, \omega) = H\left(u + \frac{2\pi mc}{\omega d}, \omega\right)\,, \tag{50.19}$$

where m is an integer. Thus, the main lobe has an infinite set of identical copies, which are called *grating lobes*. Since the *visible region* [50.2] is defined for values of $-1 \le u \le 1$, grating lobes may or may not be seen in the array response. As can be seen in (50.19), the appearance of grating lobes is a function of both microphone spacing and incident frequency.

Grating lobes are visible when another lobe appears at $|u| \le 1$. When fully visible, a grating lobe is equal in amplitude to the main lobe of the array. This phenomena is called spatial aliasing and is directly analogous to time aliasing in sampled time systems. As either the element spacing or the frequency increases, grating lobes move in an angular direction towards the main lobe and can be thought of as either a dilatation or contraction of the array function as defined in (50.19). As the spacing or frequency increases, new grating lobes become visible, appearing first along the endfire direction. If the array is not steered, grating lobes occur when the first periodic replication $(2\pi c)/\omega$ falls into the visible region $|u| \le 1$. In terms of the spacing d, grating lobes occur when

$$d \ge \frac{2\pi c}{\omega} = \lambda\,, \tag{50.20}$$

where λ is the incident sound wavelength. As shown above, the spatiofrequency response $H(u, \omega)$ is periodic, and thus *spatial aliasing* for spatiotemporal signals will occur before the grating lobe appears at $|u| = 1$. Thus, to preclude spatial aliasing while having any steering angle $|u_0| \le 1$, the grating lobe must be be out of visual space, $|u| + |u_0| \ge 2$. With this constraint,

$$d < \frac{\pi c}{\omega} = \frac{\pi c}{2\pi f} = \frac{\lambda}{2}\,. \tag{50.21}$$

Equations (50.20) and (50.21) define the *spatial sampling* requirement, which is analogous to the Nyquist sampling rate for sampled time-domain signals. If the array is steered to endfire which is the maximum steering angle in visible space, grating lobes can be avoided if adjacent lobes are separated by a distance greater than 2 in u-space (the same restriction as for no spatial aliasing for an unsteered array). It should also be noted that the lowest frequency at the onset of a grating lobe can be increased if the array is composed of directional elements that have low sensitivity in the region of $|u| \approx 1$.

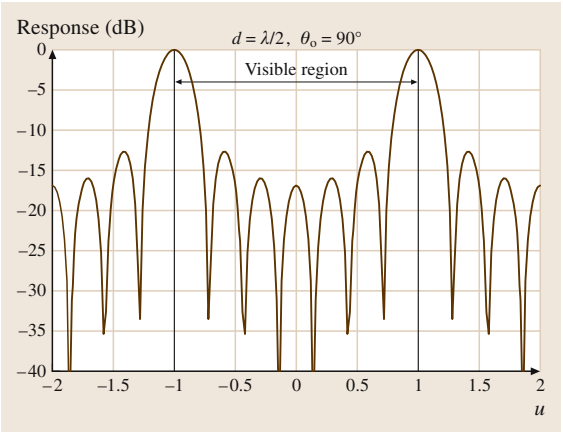

Fig. 50.4 u-response of a seven-element uniformly spaced and weighed array steered 90° as a function of the dimensionless variable u for element spacing equal to half the wavelength of sound

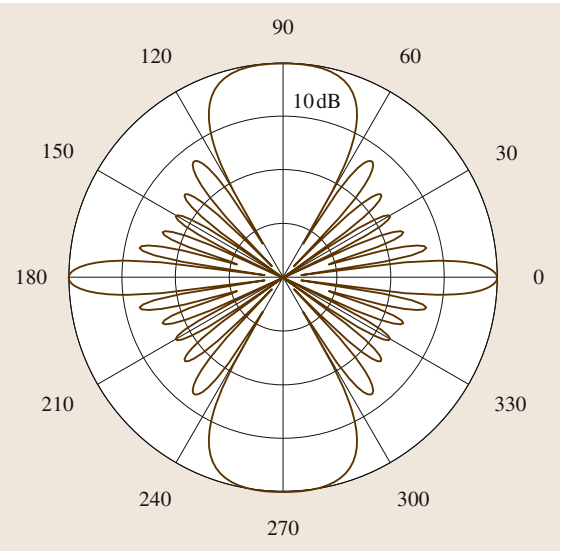

Fig. 50.5 Polar directivity pattern for a seven-element uniformly spaced and weighted unsteered array when $d = \lambda$. Note the large grating lobe at $\theta = 90°$. Also note that the directional response is axisymmetric around the $\theta = 0°$ axis

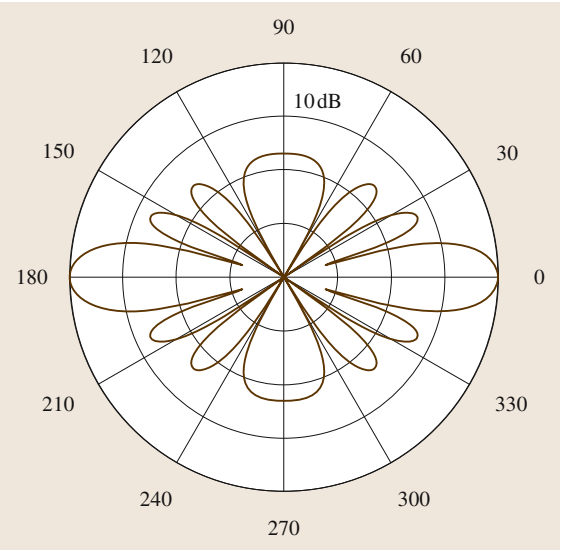

Fig. 50.6 Polar directivity pattern for a seven-element uniformly spaced and weighted unsteered array for $d = \lambda/2$. Note that compared to Fig. 50.5, there is no grating lobe in the $\theta = 90°$ direction

Uniformly Sampled and Weighted Linear Arrays

To better understand spatial aliasing and how $H(u, \omega)$ changes with increased spacing or frequency, as well as steering angle, it is instructive to examine a uniformly weighted linear array. Figure 50.2 shows the response of a uniformly weighted array as a function of the variable u for the case $d = \lambda$. The visible region is denoted on the figure as the region between the two vertical red lines. One can clearly see the grating lobes beginning to move into the visible region in the figure at $u = \pm 1$. This behavior was predicted by (50.20) for an unsteered array. Figure 50.3 shows how the response is rotated in u-space when the array is steered 45° from the unsteered direction (broadside). One can now clearly see that the grating lobe is fully within the visible region. For comparison purposes, Fig. 50.4 shows the directional response for a 90° steered array for the case of $d = \lambda/2$. Equation (50.21) predicts that for this case, a grating lobe should appear in the opposite direction (along the axis of the array but in the opposite of the steered direction). Figure 50.4 confirms that this is indeed the case as one can see the grating lobe at 270° when the array is steered to 90° (endfire).

Figures 50.5 and 50.6 show directivity patterns when $d = \lambda$ and $d = \lambda/2$ for a seven-element uniformly spaced and weighted unsteered array. It is not coincidental that there are six zeros (nulls) in the response since as can be seen in (50.17), a seven-element array response can be written as a seventh-order polynomial. As the frequency or spacing changes, the spatial response either contracts or dilates, so at lower frequencies or smaller element spacing, some or all of these nulls might not be contained within the visible region $|u| \leq 1$. One way to visualize

this effect is to plot the angular response of the array as a function of frequency.

Figure 50.7 shows a surface plot of the magnitude of the response of a seven-element array with 8 cm spacing versus angle and frequency. The colors indicate a logarithmic range limited to 40 dB. The narrowing of the main lobe as frequency increases is clearly evident in the figure. It can also be seen that the array is essentially omnidirectional below 500 Hz and that the first spatial aliasing lobe appears at around 4 kHz while a second grating lobe appears at around 8 kHz. These two frequencies correspond to sound with wavelengths of 8 and 16 cm, respectively. Thus grating lobes appear for the unsteered array at frequencies where the element spacing is an integer multiple of the incident sound wavelength, as predicted by (50.20). Six distinct zeros can be seen in the response between the grating lobes in Fig. 50.7 and it can also be seen that they gradually disappear from the visible region below 4 kHz. No nulls occur in the visible region below 500 Hz.

The angular position of the first nulls on either side of the main lobe can be computed for an unsteered uniformly weighted array by noting that (50.17) can be rewritten

$$H(u,\omega) = \frac{\sin(\hat{N}\omega u d/2c)}{\sin(\omega u d/2c)}\,, \tag{50.22}$$

where $\hat{N} = 2N+1$ is the total number of elements, The first null occurs when the argument of the numerator of (50.22) is $\pm\pi$. The beamwidth, $\mathrm{BW}_{\mathrm{null}}$ is twice the angle between the main lobe and the first null

$$\begin{aligned}\mathrm{BW}_{\mathrm{null}} &= 2\sin^{-1}\left(\frac{2\pi c}{\hat{N}\omega d}\right)\\ &= 2\sin^{-1}(\lambda/L)\,,\end{aligned} \tag{50.23}$$

where λ is the wavelength of the incident sound and $L = \hat{N}d$ is close to the array length $[(\hat{N}-1)d]$. (Note that other definitions for beamwidth are also used, such as the angles where the response falls to half the mainlobe level.) If the array length L is much larger than the incident wavelength, then one can use the small-angle approximation and obtain,

$$\mathrm{BW}_{\mathrm{null}} \approx \frac{2\lambda}{L}\,. \tag{50.24}$$

Equations (50.23) and (50.24) show that the beamwidth is proportional to the wavelength and inversely proportional to the length of the array.

Although the above analysis is only for uniformly weighted and spaced linear arrays, the general qualitative relationships between the array length, frequency, and beamwidth are similar for weighted and nonuniformly sampled arrays.

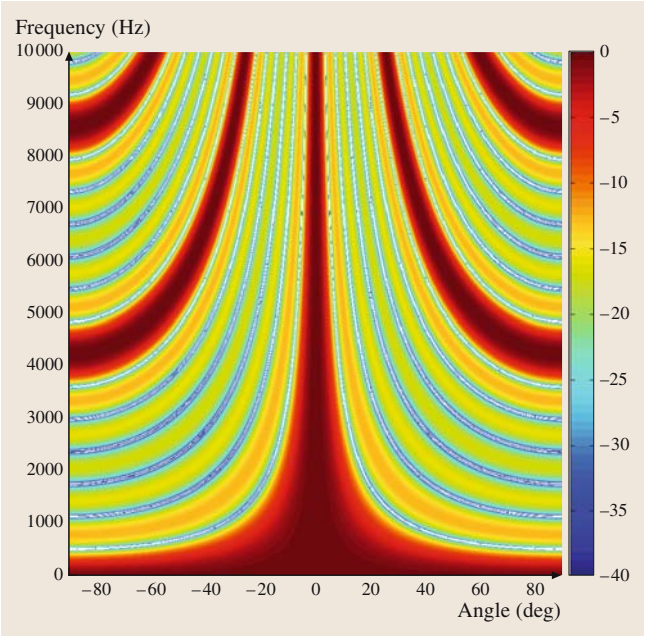

Fig. 50.7 Response of a seven-element uniformly spaced and weighted array for an element spacing of 8 cm

Analogy with FIR Filter Design

The discussion thus far expresses the beamformer behavior in terms of the array weights w_n. The opposite process, finding the weights for a given desired behavior, is also extremely important. This process is now shown to be related to finite impulse response (FIR) filter design.

Consider the Fourier transform of the array weights,

$$X(\mathrm{e}^{\mathrm{i}\omega_{\mathrm{t}}}) = \sum_{n=-N}^{N} w_n\,\mathrm{e}^{-\mathrm{i}\omega_{\mathrm{t}}n}\,, \tag{50.25}$$

where $\omega_{\mathrm{t}} = (u-u_0)\omega_0 d/c$. Comparing this with (50.17) yields the identity

$$H\left(u_0 + \frac{\omega_{\mathrm{t}}c}{\omega_0 d}, \omega_0\right) = X(\mathrm{e}^{\mathrm{i}\omega_{\mathrm{t}}})\,, \tag{50.26}$$

for any ω_0. From (50.26), it can be seen that finding the weights w_n for a given $H(u,\omega_0)$ is equivalent to an FIR filter design problem for a given desired frequency response $X(\mathrm{e}^{\mathrm{i}\omega_{\mathrm{t}}})$. Thus, like FIR filter design, array beamformer design is based on finding the appropriate coefficients of a finite-order polynomial that match a design goal in some desired way.

Classical delay-and-sum beamformer element weights are constants and therefore not a function of frequency. Since the spatial response of a delay-and-sum beamformer is a function of frequency, where the beamwidth of the array is commensurately smaller as the frequency increases, it is of interest to find a more-flexible design method that allows control of the beam pattern as a function of frequency. An obvious modification is to generalize the delay-and-sum beamformer by replacing the set of scalar weightings by a set of general filters. For obvious reasons, this generalized beamformer architecture is widely known as a filter-and-sum beamformer.

50.1.2 Filter-and-Sum Beamforming

As discussed in the previous section, it is desirable to enable the control of a beamformer design over a wide frequency range. The analysis of the delay-and-sum beamformer can be extended to a discrete-time filter-and-sum beamformer. Thus, without a loss in generality, the elemental filters are assumed to be FIR filters of length M. The beamformer output sequence, for integer κ, is

$$y[\kappa] = \sum_{n=-N}^{N} \sum_{m=0}^{M-1} p\left[\boldsymbol{r}_n, (\kappa - m)\mathrm{T}\right]\ h_n[m]\,, \quad (50.27)$$

where κ is the time sample, $h_n[m]$ is the impulse response of the filter associated with the n–th element (i. e., the n-th elemental filter), the number of microphones is $2N+1$, T is the sampling period, and $p(\boldsymbol{r}, t)$ is the incident wave field. As before, we consider the complex exponential plane wave,

$$p(\boldsymbol{r}, \kappa\mathrm{T}) = \mathrm{e}^{\mathrm{i}(\omega_0 \kappa \mathrm{T} - \boldsymbol{k}_0^{\mathrm{T}} \boldsymbol{r})}\,, \quad (50.28)$$

for some frequency ω_0 and wavevector $\boldsymbol{k}_0$. A uniformly spaced beamformer with spacing d has an output

$$y[\kappa] = \mathrm{e}^{\mathrm{i}\omega_0 \kappa \mathrm{T}} \sum_{n=-N}^{N} \sum_{m=0}^{M-1} h_n[m]\, \mathrm{e}^{\mathrm{i}\left(-\omega_0 m\mathrm{T} + \frac{\omega_0}{c} u_0 n d\right)}\,, \quad (50.29)$$

where $u_0 = \sin\theta_0$ as before.

Once again, the time-dependent portion is omitted to give the array response,

$$H(u, \omega) = \sum_{n=-N}^{N} \sum_{m=0}^{M-1} h_n[m]\, \mathrm{e}^{-\mathrm{i}\left(\omega m\mathrm{T} - \frac{\omega}{c} u n d\right)}\,. \quad (50.30)$$

From (50.30) it can now be seen that a filter-sum beamformer design problem is equivalent to a two-dimensional FIR filter design problem.

Two-Dimensional Filter Design

Just as delay-and-sum beamformer design is related to one-dimensional (1-D) FIR filter design, the filter-and-sum beamformer design is related to two-dimensional (2-D) FIR filter design that can be seen by considering an $N \times M$ function of 2-D Dirac delta functions, weighted according to the elemental filter coefficients,

$$f(x, y) = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} \delta(x - n, y - m) h_n[m]\,. \quad (50.31)$$

Now, the discrete 2-D Fourier transform of this function is:

$$X(\omega_1, \omega_2) = \sum_{n=-N}^{N} \sum_{m=0}^{M-1} h_n[m]\, \mathrm{e}^{-\mathrm{i}(\omega_1 n + \omega_2 m)}\,. \quad (50.32)$$

Note that this function is periodic with period 2π in both the ω_1 and ω_2 dimensions. Combining this with (50.30) yields

$$H\left(\frac{\omega_1 c\mathrm{T}}{\omega_2 d}, \frac{\omega_2}{\mathrm{T}}\right) = X(\omega_1, \omega_2)\,. \quad (50.33)$$

Equation (50.33) implies that a filter-and-sum beamformer can be conceptualized as a 2-D filter whose Fourier transform is closely related to the desired array response. Note that $X(\omega_1, \omega_2)$ is exactly the familiar wavenumber–frequency response [50.3], after converting the wavenumber and frequency variables $k_x = \|\boldsymbol{k}\| \sin\theta$, ω to the quantities $\omega_1 = k_x d$, $\omega_2 = \omega\mathrm{T}$, which have units of radians.

50.1.3 Arrays with Directional Elements

A microphone element frequency response can be effectively corrected with a single inverse filter placed at the output of the beamformer. Microphone elements with varying directivity patterns cannot generally be compensated for, but one can mathematically model their behavior using the *pattern multiplication* theorem [50.2]. This principle states that the total far-field response can be obtained from the far-field response of an equivalent array constructed using omnidirectional

elements, multiplied by the far-field response of the individual microphones. Thus, the total array response is the product of the two array responses,

$$H(u,\omega) = H_{\mathrm{m}}(u,\omega)H_{\mathrm{a}}(u,\omega)\,,$$

where $H_{\mathrm{m}}(u,\omega)$ is the far-field response of a single microphone element, and $H_{\mathrm{a}}(u,\omega)$ is the far-field response of the equivalent array composed of omnidirectional elements. (Note that all of the directional elements must be identical and oriented in the same manner).

50.2 Constant-Beamwidth Microphone Array System

Due to hardware limitations, early digital beamforming microphone arrays were implemented as delay-sum beamformers [50.4]. As discussed previously in (50.19), a delay-and-sum beamformer has a decreasing beamwidth as the frequency increases. In fact, (50.23) showed that these functions are inversely related. Narrowing of the beamwidth as frequency increases can result in an undesired low-pass-colored output from an array if it is placed into a typical room with reverberation. This is due to the larger beamwidth at lower frequencies that allows more input power from a reverberant acoustic field. To minimize the coloration of a desired wide-band signal, it is desirable to consider beamformers whose directivity is constant over frequency.

By examining the results given in (50.23) and (50.24), an obvious solution towards attaining a constant-beamwidth beamformer design would be to design a filter-sum beamformer where the ratio of the acoustic wavelength to effective array length was held constant or nearly constant.

One proposed solution to this problem was to split the array into subarrays that covered different

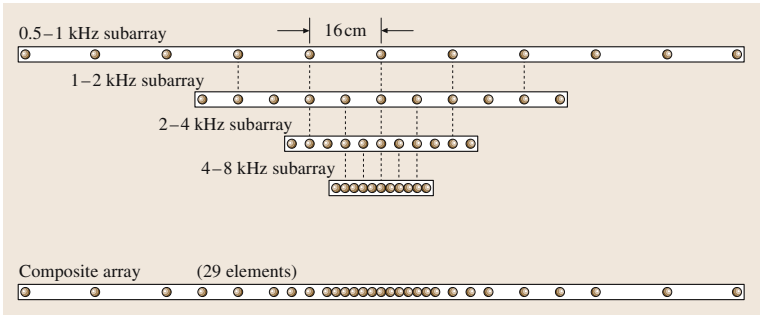

Fig. 50.8 Diagram of nested four-subarray microphone array

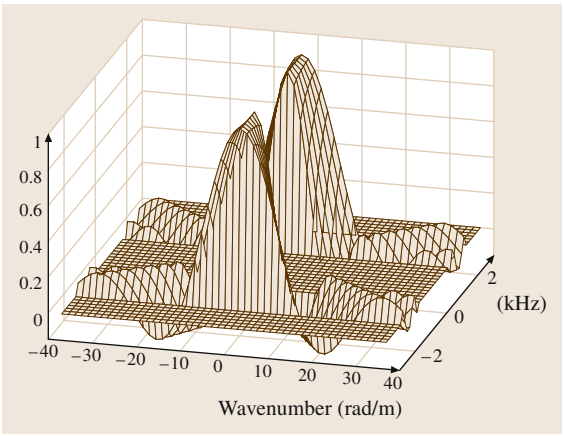

Fig. 50.9 Frequency–wavenumber design for a constant-beamwidth beamformer over one subarray

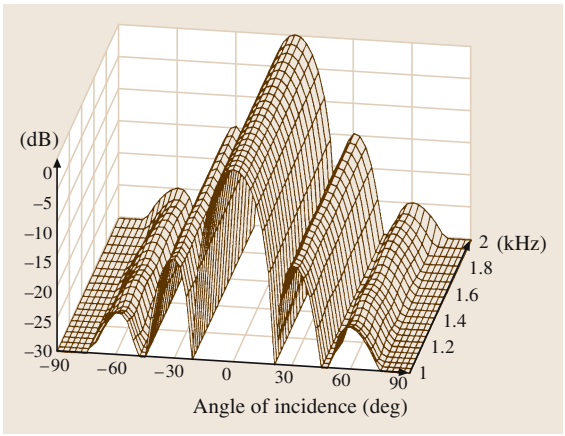

Fig. 50.10 Array response over one octave for a constant directivity array using a fan filter as shown in Fig. 50.9

frequency bands [50.4]. Some early microphone arrays were hybrid digital systems composed of three harmonically-nested subarrays covering a bandwidth from 300 Hz to 4 kHz [50.4, 5], but were later modified to include four harmonically nested subarrays (200–8 kHz). A four-band harmonically nested subarray is shown diagrammatically in Fig. 50.8. One interesting observation that can be made here is that, by harmonically nesting the subarrays, many elements are commonly shared in the array. Thus, the total number of array elements can be significantly reduced.

Further refinements were made on the multi-octave arrays to allow experimentation with a filter-sum configuration [50.6, 7]. These were developed to allow investigation into array algorithms for constant-beamwidth and dynamic control of the array beamwidth [50.7, 8]. Figure 50.9 shows the frequency–wavenumber response of the *fan filter* used to keep the beamwidth constant over the bandwidth of one octave subarray. A basic detail to notice here is that, as the frequency rises, the wavenumber bandwidth commensurately increases so that the beamwidth remains essentially constant. Figure 50.10 shows the array response over the 1–2 kHz octave when using the fan-filter design as shown in Fig. 50.9. It can be seen that the directivity is effectively constant over the whole octave. Similar fan filters are used for all octaves of the microphone array resulting in a constant beamwidth over most of the frequency range [50.6, 7].

The fan-filter design is equivalent to constant directivity beamformers based on the use of different low-pass filters on each element where the low-pass cutoff frequency is proportional the position of the element relative to the center of the array [50.9]. Intuitively the low-pass design acts to shorten the length of the array inversely proportional to frequency, thereby exactly counteracting the inverse relationship between array beamwidth and frequency.

50.3 Constrained Optimization of the Directional Gain

Directional gain is a common measure used in the design and analysis of microphone arrays. Optimization of the directional gain may not yield practical or realizable array designs, and therefore it is necessary to constrain the optimization to attain a useful design.

The directivity factor metric is typically used in quantifying the change in the acoustic signal-to-noise ratio (SNR) of a beamforming array in an isotropic noise field. The directivity factor (Q) expresses the ratio of sound energy received from the steered direction to the average energy received from all directions. Mathematically, the definition is,

$$Q(r,\theta_0,\phi_0)=\frac{4\pi\,|y(r,\theta_0,\phi_0)|^2}{\int_0^{2\pi}\int_0^{\pi}|y(r,\theta,\phi)|^2\sin\theta\,\mathrm{d}\theta\,\mathrm{d}\phi}\,, \tag{50.34}$$

where $y(r,\theta,\phi)$ is the response to a wave arriving from spherical coordinate directions (r,θ,ϕ), and (θ_0,ϕ_0) the steering directions. It can be shown that this function can be maximized for any array geometry by solving for the maximum of the ratio of two Hermitian forms [50.10]. Although it is not always the goal to maximize the directivity factor, this quantity is typically used to quantify beamforming acoustic SNR gains for far-field sources in isotropic noise fields. It is standard engineering practice to express the directivity factor in decibels, and this form is called the directivity index. The directivity index (DI) is defined as

$$\mathrm{DI}(r,\theta_0,\phi_0)=10\log\left[Q(r,\theta_0,\phi_0)\right]. \tag{50.35}$$

Optimization of the directional gain at frequencies where the array size is smaller than the acoustic wavelength results in array designs that are referred to as superdirectional arrays. Superdirectional arrays obtain directivity by forming the differences between subarray beampatterns. Superdirectional arrays have element weighting functions that oscillate in and out of phase between elements and can have directional indices (DI) potentially twice as high as classical delay-sum beamformers [$20\log(N)$ versus $10\log(N)$] [50.11]. Pattern-differencing beamformers result in a reduced acoustic signal output relative to microphone self-noise and electronics. Thus the array output SNR that can be significantly compromised. This problem has been known in the past and many authors have written on the subject [50.12]. The earliest publication containing a solution that constrains the gain to deal with the problem was a paper by *Gilbert* and *Morgan* [50.13]. However, it was not until a publication by *Cox* et al. [50.14] that a practical design methodology emerged. The following development follows that of Cox et al.

Consider a linear array of N sensors with desired signal and noise spatial cross-correlation matrices de-

noted by $\boldsymbol{P}$ and $\boldsymbol{Q}$, respectively. If one assumes that the signal and noise signals are stationary and independent processes, then the spatial cross-correlation matrix for the array can be written as

$$\boldsymbol{R}(\omega) = \sigma_{\mathrm{s}}^2 \boldsymbol{P}(\omega) + \sigma_{\mathrm{n}}^2 \boldsymbol{Q}(\omega)\,, \tag{50.36}$$

where the frequency domain is denoted by the explicit use of the temporal frequency ω. Thus the input signal-to-noise ratio is simply $\sigma_{\mathrm{s}}^2/\sigma_{\mathrm{n}}^2$. The output power spectrum of the beamformer that combines each microphone signal by a weight $w(\omega)$ can be written

$$Y = \boldsymbol{w}^{\mathrm{H}} \boldsymbol{R} \boldsymbol{w}, \tag{50.37}$$

where $\boldsymbol{w}$ is the complex weight vector containing the beamformer weights for each microphone element at frequency ω_0, and $^{\mathrm{H}}$ denotes the complex conjugate transpose. Also note that the explicit frequency dependence has been omitted for compactness of the expression. The array gain (equivalent to the directivity index for an isotropic noise field) can now be defined as the relative gain in SNR at the beamformer output, and can be written as,

$$G = \frac{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{P} \boldsymbol{w}}{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{Q} \boldsymbol{w}}\,, \tag{50.38}$$

Note that in general, $\boldsymbol{Q}$ can also contain components of sensor self-noise.

In order to quantify the gain in SNR through a beamformer, *Cox* [50.14] proposed a measure called *white-noise gain* (WNG), which simplifies the above equations by assuming that the noise matrix can be assumed as an identity matrix. An expression of the WNG is then simply

$$G_{\mathrm{w}} = \frac{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{P} \boldsymbol{w}}{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{w}} \tag{50.39}$$

From (50.39), it can be seen that lower values of WNG imply a higher relative sensitivity to white-noise signals in each microphone relative to the sensitivity of acoustic signals. Therefore, the term WNG is best understood when interpreted as the gain of the array *over* white noise.

A constrained design can be obtained by the use of Lagrange multipliers such that

$$\begin{aligned} &\arg\min_{\boldsymbol{w}} \left\{ \frac{1}{G} + \epsilon \frac{1}{G_{\mathrm{w}}} \right\} \\ &= \arg\min_{\boldsymbol{w}} \left\{ \frac{\boldsymbol{w}^{\mathrm{H}} [\boldsymbol{Q} + \epsilon \boldsymbol{I}] \boldsymbol{w}}{\boldsymbol{w}^{\mathrm{H}} \boldsymbol{P} \boldsymbol{w}} \right\}\,. \end{aligned} \tag{50.40}$$

Equation (50.40) is a generalized eigenvalue problem and the constrained optimum weight can also be found [50.10] as the eigenvector corresponding to the largest eigenvalue of the middle term in the numerator in (50.40). For a single desired source propagating from a known direction θ relative to the array, one can obtain the optimum weight as,

$$\boldsymbol{w}_{\mathrm{opt}} = [\boldsymbol{Q} + \epsilon \boldsymbol{I}]^{-1} \boldsymbol{s}^*(\theta), \tag{50.41}$$

where $\boldsymbol{s}^*$ is the complex-conjugate signal across the array generated by the desired signal at angle θ. Equation (50.41) can be seen to be the constrained optimum beamformer that operates by the process of *whiten then match*. The *whitening* is done by the inverse of the term in the brackets where areas of higher spatial noise are commensurately attenuated by the beamformer. The second term within the brackets is a regularization of the inverse of the noise correlation matrix and acts as a constraint on the inverse. Matching to waves propagating from the desired direction is accomplished by the steering vector $\boldsymbol{s}^*$ that steers the array in the direction of the desired sound source.

By adjusting the Lagrange multiplier between the values of 0 and ∞, one can obtain a monotonic change in directional gain from an unconstrained array for $\epsilon = 0$ to that of a uniformly weighted array ($\epsilon \approx \infty$).

50.4 Differential Microphone Arrays

Differential microphone arrays are part of a regime of arrays that are of general interest due to their small size relative to the acoustic wavelength. For a given number of microphone elements in an array, differential arrays theoretically have the potential to attain maximum directional gain [50.11]. However, as will be shown later in this section, there are some significant caveats in realizing high directional gain with small differential arrays.

Earlier discussion on delay and filter-sum beamformers showed that these arrays operate as FIR spatial low-pass filters. As a result, directional gain is obtained when the array length is large or on the order of the incident sound wavelength. On the contrary, differential arrays operate as spatial high-pass filters, and as such, can be developed and analyzed differently than the delay and filter-sum arrays. (To be precise, a differential array does fall into the general filter-sum beamformer

category. However, since the analysis of differential array operation can be developed more compactly, it will be assumed that these arrays are different.)

Differential array output can be understood as a finite-difference approximation to the sum of spatial derivatives of the scalar acoustic pressure field. The term *first-order* differential array applies to any array whose response is proportional to the combination of two components: a zeroth-order (acoustic pressure) signal and another proportional to the first-order spatial derivative of a scalar acoustic pressure field. Similarly, the term n-th-*order* differential array is used for arrays that have a response proportional to a linear combination of signal derived from spatial derivatives up to, and including order n.

Before discussing various implementations of n-th-order finite-difference arrays, expressions are developed for the n-th-order spatial acoustic pressure derivative in a direction $\boldsymbol{r}$. Since realizable differential arrays are approximations to acoustic pressure differentials, equations for general order differentials provide significant insight into the operation of these arrays.

The acoustic pressure field for a propagating acoustic plane wave can be written

$$p(\boldsymbol{r}, t) = A_0 \, \mathrm{e}^{\mathrm{i}(\omega_0 t - \boldsymbol{k}_0^{\mathrm{T}} \boldsymbol{r})} = A_0 \, \mathrm{e}^{\mathrm{i}(\omega_0 t - k_0 r \cos\phi)} \,, \tag{50.42}$$

where A_0 is the plane-wave amplitude. The angle ϕ is the angle between the position vector $\boldsymbol{r}$ and the wavevector $\boldsymbol{k}_0$ (note that the angle ϕ here is not the same as the standard spherical coordinate angle used to describe the angle in the x–y plane). Dropping the time dependence and taking the n-th-order spatial derivative along the direction of the position vector $\boldsymbol{r}$ yields:

$$\frac{\mathrm{d}^n}{\mathrm{d}r^n} p(k, r) = A_0 (-\mathrm{i}k \cos\phi)^n \, \mathrm{e}^{-\mathrm{i}kr\cos\phi} \,. \tag{50.43}$$

The plane-wave solution is valid for the response to sources that are *far* from the microphone array. The term *far* implies that the distance between the source and receiver is many times the square of the relevant dimension divided by the acoustic wavelength where the relevant dimension is the greater of the source or receiver dimension. Using (50.43) one can interpret that the n-th-order differential has a bidirectional pattern component with the shape of $(\cos\phi)^n$. It can also be seen that the frequency response of a differential microphone is high-pass with a slope of $6n$ dB per octave. If the far-field assumption is relaxed, the response of the differential system to a point source located at the coordinate origin is

$$p(k, r) = A_0 \frac{\mathrm{e}^{-\mathrm{i}(kr\cos\phi)}}{r} \,, \quad r > 0. \tag{50.44}$$

The n-th-order spatial derivative in the radial direction r is

$$\frac{\mathrm{d}^n}{\mathrm{d}r^n} p(k, r, \phi) = A_0 \frac{n!}{r^{n+1}} \, \mathrm{e}^{-\mathrm{i}kr\cos\phi} (-1)^n \sum_{m=0}^{n} \frac{(\mathrm{i}kr\cos\phi)^m}{m!} \,. \tag{50.45}$$

As can be seen in (50.45), a fundamental property for differential arrays is that the general n-th-order array response is a weighted sum of bidirectional terms of the form $\cos(\phi)^n$.

Implicit in the development of differential arrays is the inherent assumption that the true differential can be approximated by discrete microphones that are used to estimate the pressure differentials by finite differences. For reasonable estimation of the differential by finite differences of the acoustic pressure, the microphones must be placed such that the element spacing is much less than the acoustic wavelength over the desired frequency range. With the small-spacing requirements, one can see both the main advantage and disadvantage of differential arrays: they can attain high directional gain with small size, but also have commensurately small WNG values, indicating a high sensitivity to self-noise as well as microphone amplitude and phase mismatch.

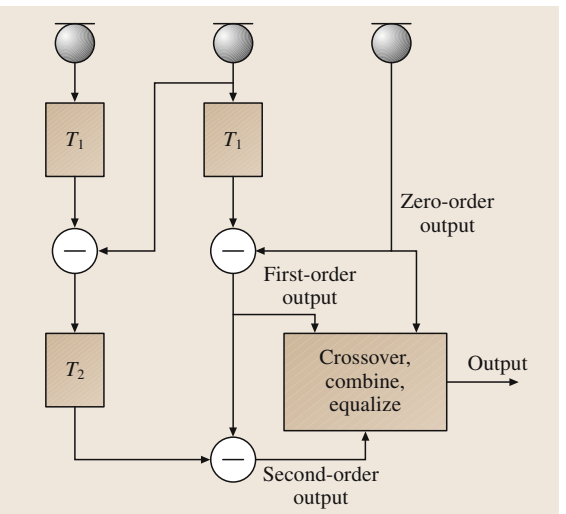

Fig. 50.11 Construction of a generalized second-order differential array as first-order differential combinations

An n-th-order differential array can be written as a sum of the n-th-order spatial differences and lower-order terms. An n-th-order array can also be written as the product of n first-order response terms as

$$y_n(\omega,\phi) = A_0 \prod_{i=1}^{n} \left(1 - \mathrm{e}^{-\mathrm{i}\omega(T_i + d_i/c\cos\phi)}\right) , \quad (50.46)$$

where the d_i relate to the microphone spacings, and the T_i relate to the chosen time delays. There is a design advantage in expressing the array response in terms of the products of first-order terms: it is now simple to represent higher-order systems as cascaded systems of lower order. Figure 50.11 shows how differential arrays can be constructed for up to second order. Extension of the design technique to higher orders is straightforward.

Values of T_i can easily be determined from the microphone spacing and the desired null angle. The ordering of T_i is not important as long as $\omega T_i \ll \pi$. If again it is assumed that $\boldsymbol{k}d_i \ll \pi$ and $\omega T_i \ll \pi$, then (50.46) can be approximated as

$$y_n(\omega,\phi) \approx A_0 \omega^n \prod_{i=1}^{n} (T_i + d_i/c\cos\phi) . \quad (50.47)$$

Equation (50.47) can be further simplified by making some simple substitutions for the arguments in the product term. Setting $\alpha_i = T_i/(T_i + d_i/c)$, then

$$y_n(\omega,\phi) \approx A_0 \omega^n \prod_{i=1}^{n} [\alpha_i + (1-\alpha_i)\cos\phi] . \quad (50.48)$$

If the product in (50.48) is expanded, a power series in $\cos\phi$ can be written for the response of the n-th-order array to an incident plane wave

$$\begin{aligned} &y_n(\omega,\phi) \\ &= A_0 G \omega^n \left(a_0 + a_1\cos\phi + a_2\cos^2\phi + \ldots \right. \\ &\quad \left. + a_n \cos^n\phi\right) , \end{aligned} \quad (50.49)$$

where the constant G is an overall gain factor. The only frequency-dependent term in (50.49) is ω^n. Thus the frequency response of an n-th-order differential array can be easily compensated by a low-pass filter whose frequency response is proportional to ω^{-n}. By choosing a structure that places only a delay behind each element in a differential array, the coefficients in the power series in (50.49) are independent of frequency, resulting in an array whose beampattern is independent of frequency.

To simplify the following exposition on the directional properties of differential arrays, it is assumed that the amplitude factor can be neglected. Also, since the directional pattern described by the power series in $\cos\phi$ can have any general scaling, the directional response can be written as solely a function of ϕ:

$$\begin{aligned} y_n(\phi) &= a_0 + a_1\cos\phi + a_2\cos^2\phi + \ldots \\ &\quad + a_n\cos^n\phi . \end{aligned} \quad (50.50)$$

Without loss of generality, the coefficients a_i can be defined to attain a normalized response at $\phi = 0°$, which implies

$$\sum_{i=0}^{n} a_i = 1. \quad (50.51)$$

In general, n-th-order differential microphones have, at most, n nulls (zeros). This follows directly from (50.50) and the fundamental theorem of algebra.

As an instructive example, let us examine the specific case for a second-order array,

$$y_2(\phi) = a_0 + a_1\cos\phi + a_2\cos^2\phi . \quad (50.52)$$

Equation (50.52) can also be factored into two first-order terms and written as

$$y_2(\phi) = [\alpha_1 + (1-\alpha_1)\cos\phi][\alpha_2 + (1-\alpha_2)\cos\phi] , \quad (50.53)$$

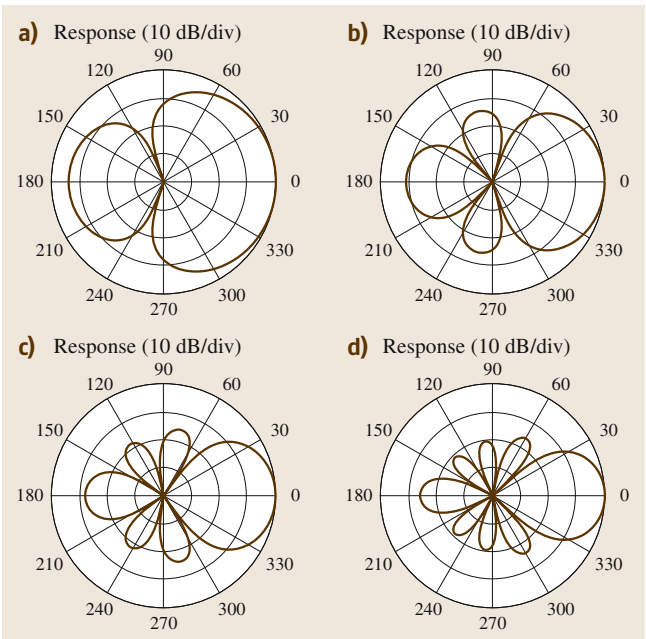

Fig. 50.12a–d Directional patterns that have maximum directional gain for differential microphone arrays for up to forth-order arrays

where

$$a_0 = \alpha_1\alpha_2,$$
$$a_1 = \alpha_1(1-\alpha_2)+\alpha_2(1-\alpha_1),$$
$$a_2 = (1-\alpha_1)(1-\alpha_2)\,, \tag{50.54}$$

or

$$\alpha_1 = a_0 + a_1/2 \pm \sqrt{(a_0+a_1/2)^2 - a_0}\,,$$
$$\alpha_2 = a_0 + a_1/2 \mp \sqrt{(a_0+a_1/2)^2 - a_0}\,. \tag{50.55}$$

As shown, the general form of the second-order system is the sum of second-, first-, and zero-order terms. If certain constraints are placed on the values of a_0 and a_1, it can be seen that there are two nulls (zeros) in the interval $0 \le \phi < \pi$. The array response pattern is axisymmetric about $\phi = 0$. These zeros can be explicitly found at the angles ϕ_1 and ϕ_2:

$$\phi_1 = \cos^{-1}\left(\frac{-\alpha_1}{1-\alpha_1}\right)\,, \tag{50.56}$$

$$\phi_2 = \cos^{-1}\left(\frac{-\alpha_2}{1-\alpha_2}\right)\,, \tag{50.57}$$

where now α_1 and α_2 can take on either positive or negative values where the magnitude of the inverse cosine is less than or equal to one. If the resulting beam pattern is constrained to have a maximum at $\phi = 0°$, then the values of α_1 and α_2 are also constrained. One interesting thing to note is that negative values of α_1 or α_2 correspond to a null moving into the front half-plane. Negative values of α_1 for the first-order microphone can be shown to have a rear-lobe sensitivity that exceeds the sensitivity at 0°. Since (50.53) is the product of two first-order terms, emphasis of the rear lobe caused by a negative value of α_2 can be counteracted by the zero from the term containing α_1. As a result, a beampattern can be found for the second-order microphone that has maximum sensitivity at $\phi = 0°$ and a null in the front-half plane. This result also implies that the beamwidth of a second-order microphone with a negative value of α_2 is narrower than that of the second-order dipole ($\cos^2(\phi)$ directional dependence). Figure 50.12 shows the beam patterns that have the highest attainable directivity for first through forth-order differential arrays [50.11].

50.5 Eigenbeamforming Arrays

The scalar acoustic pressure sound field obeys the Helmholtz equation [50.1], which states that the distribution of acoustic pressure and particle velocity on a surface uniquely defines the sound field within this surface if no sources or obstacles are enclosed within the surface. One can also use the Helmholtz equation to compute the external sound field if there are no obstacles or sources external to the measured surface. Thus, the interior and exterior sound field is uniquely determined if the sound pressure and particle velocity on a closed surface are known. Therefore it follows that only the sound pressure and particle velocity on a closed surface is required to capture all information of the sound field. This observation leads to a different beamformer implementation where, instead of forming beams in the classical sense by combining filtered individual microphone signals, one can use spatially decomposed *eigenbeams* to realize a general beamformer.

50.5.1 Spherical Array

Although the array surface can be any general surface, one can greatly simplify the analysis and array construction by using microphones placed on a rigid surface. Using a rigid body has the advantage that the radial particle velocity on its surface is zero. This greatly simplifies the general solution since only the sound pressure needs to be measured. Using a spherical geometry also simplifies the mathematics and allows for a design that puts equal weight on all directions due to its spherical symmetry. Theoretically many other shapes are possible, but the most amenable are shapes that are congruous to separable coordinate systems like cylindrical, oblate,

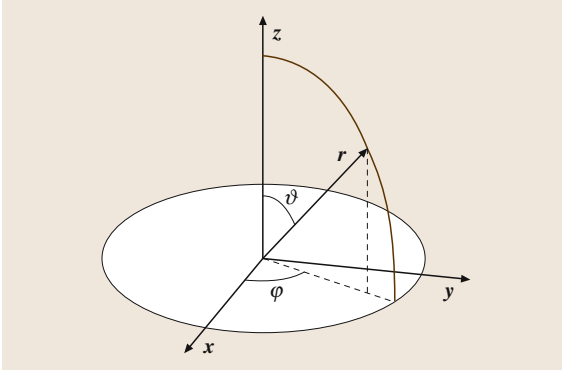

Fig. 50.13 Definition of the spherical coordinate system

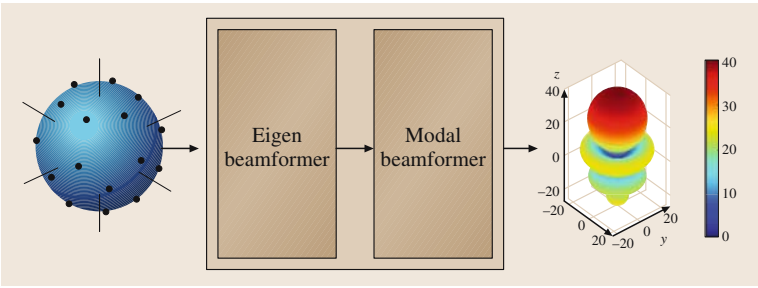

Fig. 50.14 Block diagram of the functional blocks of the spherical array

prolate, etc. Fig. 50.13 shows the notation for the standard spherical coordinate system that is used for the spherical array.

50.5.2 Eigenbeamformer

The overall structure of the spherical array can be decomposed into two main cascaded sections as shown in Fig. 50.14: an eigenbeamformer and a modal beamformer [50.15, 16].

The task of the eigenbeamformer is to transform the microphone signals into an orthonormal beam space. Since these beams are characteristic of the sound field in a similar way as eigenvectors are for a matrix, they are referred to as eigenbeams. Hence, the beamforming preprocessor is referred to as an eigenbeamformer. The second stage is a *modal* beamformer that linearly combines the eigenbeams to form a desired output beam pattern. In the following sections, it is shown that a spherical harmonic decomposition of the sound field leads to a computationally simple realization of the modal beamformer.

To begin an analysis of the spherical eigenbeamformer, we first need to introduce the orthonormal spherical harmonic eigenfunctions $Y_n^m(\vartheta, \varphi)$ which have order n and degree m, and are functions of the spherical angles ϑ, φ. The spherical harmonics Y_n^m are defined as [50.17],

$$Y_n^m(\vartheta, \varphi) = \sqrt{\frac{(2n+1)}{4\pi}\frac{(n-m)!}{(n+m)!}} P_n^m(\cos\vartheta)\mathrm{e}^{\mathrm{i}m\varphi}, \tag{50.58}$$

where $P_n^m(\cos\vartheta)$ are the standard associated Legendre functions. To calculate the response of the spherical eigenbeamformer to a planewave with incident angle ϑ, φ, and wavenumber k, it is judicious to express the plane wave in terms of spherical harmonics. By matching the boundary conditions of the rigid sphere of radius a to the impinging plane wave from the direction (ϑ, φ), one can compute the total acoustic pressure on the surface at location $(a, \vartheta_\mathrm{s}, \varphi_\mathrm{s})$ as

$$p_\mathrm{s}(\vartheta_\mathrm{s}, \varphi_\mathrm{s}, ka, \vartheta, \varphi) = 4\pi \sum_{n=0}^{\infty} i^n b_n(ka) \sum_{m=-n}^{n} Y_n^m(\vartheta, \varphi) Y_n^{m^*}(\vartheta_\mathrm{s}, \varphi_\mathrm{s}) . \tag{50.59}$$

The b_n are the *modal coefficients*, defined as

$$b_n(ka) = \left[j_n(ka) - \frac{j_n'(ka)}{h_n^{(2)'}(ka)} h_n^{(2)}(ka) \right] , \tag{50.60}$$

where $j_n(ka)$ are the spherical Bessel functions, and the prime indicates the derivative with respect to its argument, and the function $h_n^{(2)}(k)$ is the spherical Hankel of the second kind.

From the previous development, it can be seen that the rigid spherical baffle has a unique pressure field for plane waves impinging from different angles. A desired beam in a specific direction can be formed by filtering the surface pressure by a general weighting of the acoustic pressure on the surface. The spherical eigenbeamformer operates by decomposing this general surface weighting function into spherical harmonics. Thus, the output of an eigenbeamforming microphone to an incident plane wave for a spherical harmonic weighting function of $Y_{n'}^{m'}(\vartheta_\mathrm{s}, \varphi_\mathrm{s})$ can be written

$$\begin{aligned} F_{n',m'}(\vartheta, \varphi, ka) &= \int_{\Omega_\mathrm{s}} p_\mathrm{s}(\vartheta_\mathrm{s}, \varphi_\mathrm{s}, ka, \vartheta, \varphi) Y_{n'}^{m'}(\vartheta_\mathrm{s}, \varphi_\mathrm{s}) \mathrm{d}\Omega_\mathrm{s}, \\ &= 4\pi i^{n'} b_{n'}(ka) Y_{n'}^{m'}(\vartheta, \varphi) \end{aligned} \tag{50.61}$$

where $\mathrm{d}\Omega_\mathrm{s}$ represents an integration over the surface of the sphere. The far-field directivity of the microphone has the same directional dependence as the applied spherical harmonic sensitivity function, namely $Y_{n'}^{m'}$. Thus, it can now be seen that forming a desired beam

can be accomplished by using spherical harmonics as the basis functions expansion of the general surface weighting and then appropriately weighting and summing the outputs of these basis functions. The paragraph below on eigenbeams gives a detailed analysis of these eigenbeams. One final thing to note here is that the modal coefficients b_n introduce a frequency dependance. In order to combine different eigenbeams appropriately to form a desired beam, the b_n coefficients must be equalized so that they have the same amplitude and frequency response.

Discrete Orthonormality

In order to realize a spherical array, the continuous aperture has to be sampled. To achieve the same result as obtained in (50.61), the sample positions must fulfill the following *discrete orthonormality* condition:

$$A_{nm}\sum_{s=0}^{S-1} Y_n^m(\vartheta_s,\varphi_s)Y_{n'}^{m'}(\vartheta_s,\varphi_s)=\delta_{nn'}\delta_{mm'} \quad . \quad (50.62)$$

where the A_{nm} are scale factors that are required to fullfill the discrete orthonomality condition. Although is not trivial to find a set of sensor locations that fulfill this discrete orthonormality, one sensor arrangement that fulfills this constraint up to fourth-order spherical harmonics are sample points at the center of the 32 faces of a truncated icosahedron [50.16].

The resulting structure of the eigenbeamformer can be derived from (50.62): for a specific eigenbeam Y_n^m the microphone signals are weighted by the sampled values of the corresponding surface sensitivity, $Y_n^m(\vartheta_s,\varphi_s)$. It can therefore be seen that the eigenbeamformer is effectively a set of simple delay-sum beamformers per eigenbeam (with no delay required).

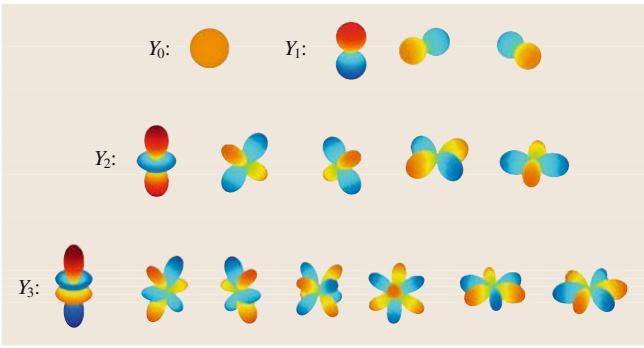

Fig. 50.15 Real and imaginary components of the spherical harmonics from zero to third order

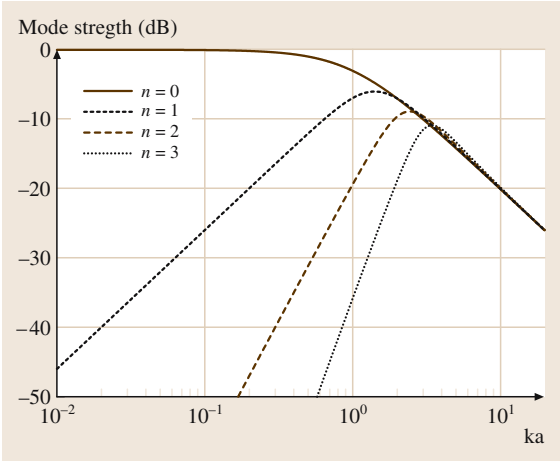

Fig. 50.16 Modal coefficients b_n for the first four orders

Besides the orthonormal constraint, spatial aliasing has to be taken into account when sampling a discrete aperture as in the design of classical beamforming arrays. Since there are $(N+1)^2$ spherical harmonics for a spatial resolution of order N, a minimum of $(N+1)^2$ sample locations are required to distinguish the spherical harmonics.

Eigenbeams

The outputs of the eigenbeamformer are a set of orthonormal beampatterns that are referred to as eigenbeams. These eigenbeams represent a spatially orthonormal decomposition of the sound field and represent a complete description of the original sound field up to the highest-order spherical harmonic used in the orthonormal expansion.

Figure 50.15 shows some example eigenbeams. From (50.58), it can be seen that the elevation dependence follows the Legendre function while the azimuth dependence has a sine-cosine dependance. The order n determines the number of zeros in the ϑ-direction while two times the degree m gives the number of zeros in the φ-direction.

The total number of eigenbeams required to capture N-th-order spherical harmonics of the sound field is $(N+1)^2$. As noted previously, the maximum achievable directional gain is $20\log(N+1)$. Thus, to obtain a maximum directional gain of 12 dB for any arbitrary direction, one would require eigenbeams up to third order. Note here that the term *order* has the same context as that used in differential arrays.

Similar to differential arrays, eigenbeams have the desirable quality that their directional response is fre-

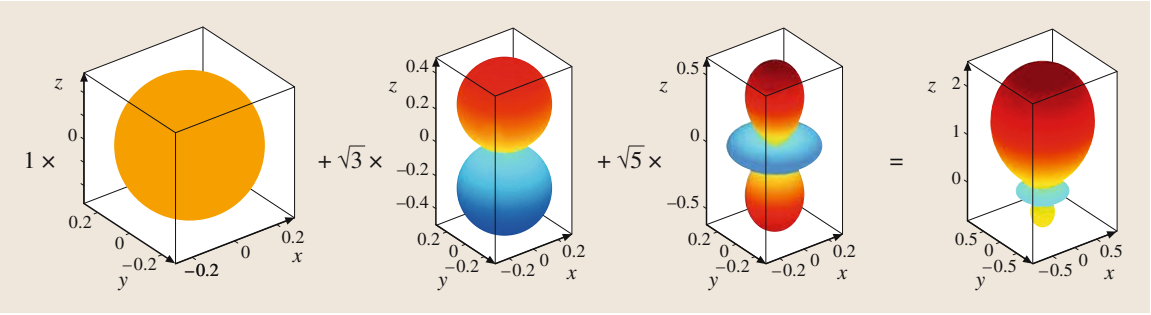

Fig. 50.17 Generating a second order hypercardioid pattern (note that only the directional properties are shown)

quency independent; however, as shown in Fig. 50.16 and (50.60), the magnitude response of the eigenbeams are dependent on the eigenbeam order and have high-pass responses that are $6n$ dB per octave. Thus one needs to equalize the different eigenbeams when combining them to realize a desired beampattern.

50.5.3 Modal Beamformer

The modal beamformer forms the second stage of the overall eigenbeamformer processing structure and comprises a simple weight-and-sum beamformer

$$D(\vartheta,\varphi)=\sum_{n=0}^{N} c_n Y_n(\vartheta,\varphi)\,, \tag{50.63}$$

where the beamformer multiplies each input beam by a factor c_n and adds up all weighted beams. As an example, Fig. 50.17 shows the generation of a second-order hypercardioid pattern steered along the z-axis.

To steer the resulting beam pattern towards a desired look-direction $[\vartheta_0,\varphi_0]$ one can use the Legendre polynomial addition theorem [50.1],

$$P_n(\cos\Theta) = \sum_{m=-n}^{n}\frac{(n-m)!}{(n+m)!}P_n^m(\cos\vartheta)P_n^m(\cos\vartheta_\mathrm{s})\,\mathrm{e}^{\mathrm{i}m(\varphi-\varphi_\mathrm{s})} \tag{50.64}$$

and (50.63) to write the overall steered array response as

$$D_{\vartheta_0\varphi_0}(\vartheta,\varphi) = \underbrace{\sum_{n=0}^{N} c_n}_{\text{combine}} \underbrace{\sum_{m=-n}^{n}\sqrt{\frac{(n-m)!}{(n+m)!}}P_n^m(\cos\vartheta_0)\,\mathrm{e}^{-\mathrm{i}m\varphi_0}}_{\text{steer}} \times Y_n^m(\vartheta,\varphi)\,. \tag{50.65}$$

Although modal beamforming arrays seem quite different from classical arrays, they are equivalent in operation but offer a more-intuitive and efficient way to realize beamforming arrays. As noted earlier, eigenbeamforming is not limited to using a spherical harmonic expansion to realize a modal beamformer, and any other spatially orthogonal basis function expansion can be used. Modal beamforming offers the capability of reducing the number of signals that need to be stored for processing since, in principal, the number of eigenbeams required can be much less than the number of microphone elements. Modal beamforming also has the desirable property that forming multiple simultaneous beams is computationally very efficient since the modal beamformer requires M multiply-adds, where M is the number of eigenbeams to form each unique output. Finally, the inherent computational simplicity of modal beamformers leads to efficient dynamic and adaptive microphone array implementations based on multiple beams and not on individual elements.

50.6 Adaptive Array Systems

The microphone arrays described in the previous sections are beamformers whose spatial responses are static since the weights, once designed, are fixed. Designs of fixed beamformers are done with assumptions of the signal and noise temporal and spatial statistics. This approach often leads to good designs when the a priori

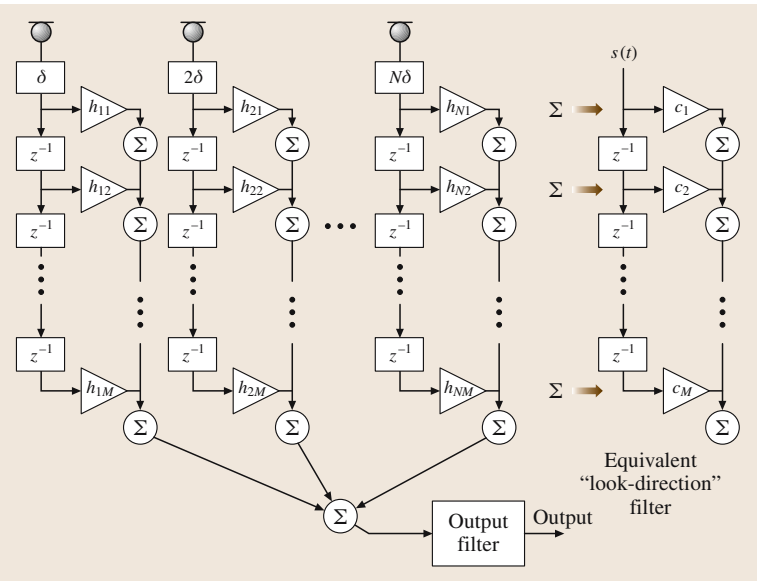

Fig. 50.18 Block diagram of a broadband adaptive microphone array implemented as an FIR filter-sum beamformer

assumptions match the acoustic conditions in which the array is placed. However, in acoustic and noise environments that dynamically change, it can be advantageous to use arrays that self-optimize their performance based on the sound field that the array is placed. Arrays that self-optimize their spatial response are generally referred to as adaptive arrays. There are many forms of adaptive array systems [50.18], but only a constrained broadband array is described here since most other schemes are based on similar approaches.

50.6.1 Constrained Broadband Arrays

Adaptive beamformers typically optimize a quadratic metric that is related to the output SNR. A common optimization is to minimize the output power under the constraint that the desired signal is not affected or only minor modifications are imparted to the signal. Thus one needs to minimize the output power under a constraint that the desired signal is not harmed. In most cases, like that of speech pickup in a room, the statistics of the desired source are not precisely known or are highly nonstationary. For this case, common adaptive array designs are based on the assumption that the position or direction of the desired source is known. The position can be found by beam-steering techniques based on power or known signal statistics, or by more sophisticated methods such as eigendecomposition [50.3] or other time-delay estimation direction-finding schemes [50.19]. It is further assumed that the locations of the desired source and noise sources change slowly. In a typical scenario, there may be a desired source at one angle and undesired interferers at other angles. In view of these requirements, adaptive optimization of the array output can result in better performance than designs based on assumed noise field models.

A block diagram of an adaptive filter-sum array is shown in Fig. 50.18. Each microphone feeds a weighted tapped delay line (FIR filter) whose intermediate outputs are summed. If there are N microphones and M taps on each delay line, there are a total of NM tap weights h_{nm}. All weights are adapted to optimize the output according to some specified criterion which, as mentioned above, is typically based on minimizing the output power under the constraint that the desired source direction is not affected or only slightly degraded.

The considerations governing the choice of N and M for a uniform linear array are described as follows. A linear combination of the microphone signals can produce at most $N-1$ controllable nulls in the directivity pattern of the array at a single frequency. If an M-tap FIR filter is placed behind each microphone, then the same number of nulls can be produced at M frequencies ($M/2$ if the weights are real). Thus, for broadband sources, N is governed by the expected number of noise sources and M by the frequency range over which the desired signal has appreciable energy. The goal of the adaptive algorithm is to search for the optimum location of the nulls under the constraints that signals arriving from the desired direction are not too distorted in frequency response.

Adaptive arrays of this type have been considered in the past. Constrained optimization of an adaptive array like the one shown in Fig. 50.18 can be written in matrix form as

$$\min_{\boldsymbol{h}} \boldsymbol{h}^{\mathrm{T}} \boldsymbol{R} \boldsymbol{h} \quad \text{subject to} \quad \boldsymbol{C}\boldsymbol{h} = \boldsymbol{c} \,, \tag{50.66}$$

where $\boldsymbol{R}$ is the input cross-correlation matrix, the $M \times N$ matrix $\boldsymbol{C}$ describes the M linearly independent constraints, and $\boldsymbol{c}$ is the set of M constraining values. *Frost* [50.20] developed an algorithm that minimizes the output power of the array while maintaining a flat magnitude and linear phase response for signals arriving from the desired source direction. The Frost algorithm uses a *hard constraint* such that the output contains only a delayed version of the signal arriving from the *look direction* (the look direction is the direction that is known a priori and is set by the delay values δ in Fig. 50.18). Constraints are enforced by noting that for signals propagating from the *look direction*, the signals moving through the FIR tapped-delay lines are all in phase. The matrix $\boldsymbol{C}$ sums all of the columns to form an equivalent vector $\boldsymbol{c}$ which is the response of the array to the look direction. The look-direction equivalent filter is the column sum and is shown in Fig. 50.18 as the rightmost FIR filter. For the Frost hard constraint of flat magnitude and linear phase, the vector $\boldsymbol{c}$ must be a Dirac delta function (where all taps but one tap are zero and the single nonzero tap is set to the desired gain of signals propagating from the look direction). Although the Frost algorithm is straightforward, one can extend the degrees of freedom by softening the constraint, which is useful when the number of taps is small.

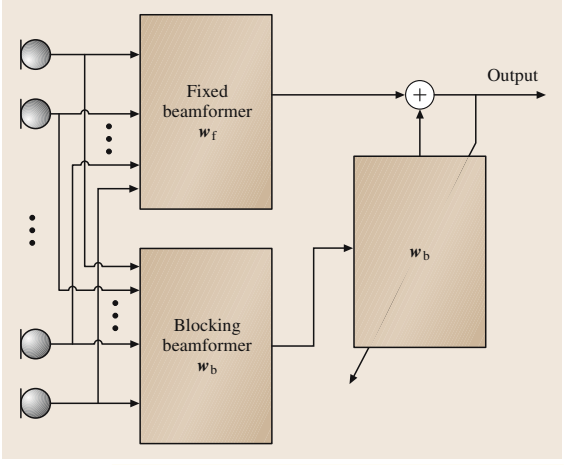

Fig. 50.19 Schematic of a generalized sidelobe canceler (GSC) adaptive array

One variant of the Frost beamformer that can simplify the computational cost by forcing the constraint into the front-end of the array processing, was proposed by *Griffiths* and *Jim* [50.21]. Their solution was to preform a set of $N-1$ difference beams that have nulls in the look direction. These *prebeams* are then used to adjust an adaptive filter-sum beamformer to minimize the total output power. Since the prebeams do not contain any look-direction signal, they implicitly invoke the look-direction constraint when combined adaptively with the main adaptive beamformer. Figure 50.19 shows a general schematic of how the Griffiths–Jim beamformer operates. In Fig. 50.19 the fixed filter-sum beamformer is denoted as $\boldsymbol{W}_{\mathrm{f}}$, the subtractive prebeamformer *blocking matrix* as $\boldsymbol{W}_{\mathrm{b}}$, and the adaptive beamformer as $\boldsymbol{W}_{\mathrm{a}}$. The general structure shown in Fig. 50.19 is known as a *generalized sidelobe canceler* (GSC) [50.18] and this structure was actually the basis of the first adaptive beamformers, which predated the development of the Frost beamformer. Intuitively, the GSC structure is appealing since it allows one to directly visualize the operation of the adaptive beamformer.

A GSC beamformer generally has one fixed main beam steered in the known look direction, and one or more secondary beams that have nulls in the look direction. These secondary beams are used to identify spatial noise sources not in the look-direction and then linearly combined with the main beam to adaptively minimize the output power or some other defined performance measure. Specific works on GSC-based beamformers for microphone arrays have been presented by *Hoshuyama* et al. [50.22], and *Herbordt* et al. [50.23].

It is well known that the Frost beamformer and GSC beamformers are sensitive to multiple paths [50.24], which manifest themselves as *signal cancelation*. Signal cancelation is due to correlated off-axis signals that cancel the desired look-direction signal by destructive interference within the filter-sum beamformer. This effect can be easily visualized by considering the operation of a GSC beamformer. In GSC beamforming, reflected look-direction signals pass the *blocking matrix* and can therefore be used by the adaptive beamformer to cancel the look-direction signal. The exact amount of signal cancelation depends on the geometry of the array as well as the acoustic environment. One practical implementation to mitigate this multipath effect is to use a smaller number of taps so that reflected signals do not enter into the adaptive FIR tapped-delay lines [50.24].

Kaneda and *Ohga* [50.25] replaced the Frost *hard* constraint with a softer constraint. Since they were

mainly interested in speech signals and realized that small perturbations to magnitude and phase were not detectable on a speech signal, their *soft* constraint allowed for some deviation in magnitude and phase for signals propagating from the look direction.

The *soft* constraint can be seen if one writes the transfer function $G(\mathrm{i}\omega)$ from the desired source to the output,

$$D = \int |G(\mathrm{i}\omega) - 1|^2 \, \mathrm{d}\omega \le D_0, \tag{50.67}$$

where D_0 is some specified positive number.

Neither the Frost or Kaneda approaches exploited the fact that the quality of speech is less sensitive to the phase of the transfer function from the desired source direction to the array output. In order to capitalize on this property, a different soft-constraint algorithm was proposed by *Sondhi* et al. [50.24] that places a constraint on only the magnitude of the transfer function. This modified constraint results in an output magnitude response variation defined as

$$D = \int \left[|G(\mathrm{i}\omega)|^2 - 1 \right]^2 \mathrm{d}\omega. \tag{50.68}$$

Fortunately a computationally simple algorithm can be found that uses a gradient search to minimize the normalized output power with the constraint that

$$D \le D_0 \, . \tag{50.69}$$

A constraint such as (50.69) is difficult to impose directly. Therefore a search to minimize another cost function C defined as

$$C = P + \epsilon D \tag{50.70}$$

leads to a solution, where P is the normalized output power and ϵ is a Lagrange multiplier. The value D obtained at the optimum setting of the filters is then implicitly a function of the parameter ϵ. By varying ϵ, a desired value of the distortion is realized.

Figure 50.18 contains a block labeled *output filter*. The purpose of this filter is to compensate for the frequency response deviation that occurs in the softly constrained adaptive beamformer. In order for the array to steer nulls at frequencies where wavelengths are longer than the array length, the array must resort to pattern differencing, which results in a commensurate high-pass response in the look direction. If one constrains the array too tightly, the adaptive algorithm focuses on maintaining flat array response and not on minimizing the output power. By allowing larger-magnitude distortion at lower frequencies, the adaptive algorithm focuses on minimizing the output power at these lower frequencies. The inverse filter acts to compensate the commensurate low-frequency roll-off distortion. An approach that accomplishes a similar result was introduced by *Gooch* and *Shynk* [50.26], who proposed an adaptive array containing both pole and zero filters. This can effectively handle the constrained adaptive array and the frequency response change due to superdirectional operation.

50.7 Conclusions

This chapter has introduced various types of beamforming microphone array systems and discussed some of the fundamental theory of their operation, design, and implementation. With the continuing trend of lower cost electronics, it is now feasible to implement microphone arrays in a variety of low-cost consumer audio communication devices. It was shown that appropriately designed microphone arrays have the ability to offer directional gains that can significantly improve the quality of hands-free speech pickup in reverberant and noisy environments.

References

50.1 P.M. Morse, K.U. Ingard: *Theoretical Acoustics* (McGraw Hill, New York 1968)

50.2 B.D. Steinberg: *Principles of Aperture and Array System Design* (Wiley, New York 1976)

50.3 D.H. Johnson, D.E. Dudgeon: *Array Signal Processing: Concepts and Techniques* (Prentice-Hall, Upper Saddle River 1993)

50.4 J.L. Flanagan, J.D. Johnston, R. Zahn, G.W. Elko: Computer-steered microphone arrays for sound transduction in large rooms, J. Acoust. Soc. Am. **78**, 1508–1518 (1985)

50.5 W. Kellermann: A self steering digital microphone array, Proc. IEEE ICASSP **1991**, 3581–3584 (1991)

50.6 R.J. Lustberg: *Acoustic Beamforming Using Acoustic Arrays* (MIT Masters Thesis, MIT Cambridge 1993)

50.7 T.C. Chou: *Broadband Frequency-Independent Beamforming* (MIT Masters Thesis, MIT Cambridge 1994)

50.8 M.M. Goodwin: *Implementation and Applications of Electroacoustic Array Beamformers* (MIT Masters Thesis, MIT Cambridge 1992)

50.9 D.B. Ward, R.A. Kennedy, R.C. Williamson: FIR filter design for frequency invariant beamformers, IEEE Signal Process. **3**, 69–71 (1996)

50.10 D.K. Cheng, F.I. Tseng: Gain optimization for arbitrarily antenna arrays, IEEE Trans. Antennas Prop. **AP-13**, 973–974 (1965)

50.11 G.W. Elko: Differential microphone arrays. In: *Audio Signal Processing for Next Generation Multimedia Communication Systems*, ed. by Y. Huang, J. Benesty (Kluwer Academic, Dordrecht 2004)

50.12 N. Yaru: A note on super-gain antenna arrays, Proc IRE **39**, 1081–1085 (1951)

50.13 E.N. Gilbert, S.P. Morgan: Optimum design of antenna arrays subject to random variations, Bell System Tech. J. **34**, 637–663 (1955)

50.14 H. Cox, R. Zeskind, T. Kooij: Practical supergain, IEEE Trans. ASSP **34**, 393–398 (1986)

50.15 J. Meyer, G.W. Elko: A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield, IEEE ICASSP **2002**, 1781–1784 (2002)

50.16 J. Meyer, G.W. Elko: Spherical Microphone Arrays for 3D Sound Recording. In: *Audio Signal Processing for Next Generation Multimedia Communication Systems*, ed. by Y. Huang, J. Benesty (Kluwer Academic, Dordrecht 2004)

50.17 E.G. Williams: *Fourier Acoustics* (Academic Press, San Diego 1999)

50.18 R.A. Monzingo, T.W. Miller: *Introduction to Adaptive Arrays* (Wiley, New York 1980)

50.19 L. Chen, J. Benesty, Y. Huang: Time delay estimation in room acoustic environments: an overview, EURASIP Appl. Signal Process. **2006**, ID26503 (1972)

50.20 O.L. Frost: An algorithm for linear constrained adaptive array processing, Proc. IEEE **60**, 926–935 (1972)

50.21 L.J. Griffiths, C.W. Jim: An alternative approach to linearly constrained beamforming, IEEE Trans. Antennas Prop. **AP-30**, 27–34 (1982)

50.22 O. Hoshuyama, A. Sugiyama, A. Hirano: A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters, IEEE Trans. Signal Process. **47**, 2677–2684 (1999)

50.23 W. Herboldt, W. Kellermann: Analysis of blocking matrices for generalized sidelobe cancellers for non-stationary broadband signals, Proc. IEEE ICASSP **2002**, 4187–4191 (2002)

50.24 M.M. Sondhi, G.W. Elko: Adaptive optimization of microphone arrays under a nonlinear constraint, Proc. IEEE ICASSP **2002**, 981–984 (2002)

50.25 Y. Kaneda, J. Ohga: Adaptive microphone array system for noise reduction, IEEE Trans. ASSP **34**, 1391–1400 (1986)

50.26 R.P. Gooch, J.J. Shynk: Wide-band adaptive array processing using pole-zero digital filters, IEEE Trans. Antennas Prop. **AP-34**, 355–367 (1985)