ON THE CONDITIONING OF THE PROPAGATION FUNCTION IN A NEAR-FIELD LOUDSPEAKER ARRAY

Liang Wang, Andy W. H. Khong and Woon-Seng Gan

School of Electrical and Electronic Engineering, Nanyang Technological University
email: wangliang, andykhong and ewsgan@ntu.edu.sg

ABSTRACT
This paper presents an algorithm for computing the weights of a loudspeaker array based on a given set of listening locations and their respective desired sound field distribution. We achieve this by introducing the E-norm condition number to control the desired sound field distribution for our application. We show that the conditioning of the problem is determined by this E-norm condition number and that part of the computation of this number is independent of frequency. Exploiting this intrinsic property and through the use of an optimization algorithm, we develop an efficient algorithm for the computation of array weights to achieve desired sound field distribution in the spatial domain. The proposed algorithm can also be used to search for a set of alternative listening locations that give rise to a well-conditioned solution for the loudspeaker weights.

1. INTRODUCTION
The use of a loudspeaker array system for far-field beam-forming has been an active area of research since the middle of last century. The objective of such an array is to project sound in the far-field. One of the applications of such a loudspeaker array is the deployment of a sound reproduction system for a live concert in an open field or a large hall. In such cases, it is important to project the sound energy in a beam in order to reach a group of targeted audience in the far-field of the array. To meet such a demand, loudspeakers are often constructed in a vertical line array. This allows users to achieve long throw characteristics of a straight-line array [1].

One of the early attempts in controlling such loudspeaker array using digital signal processing techniques for controlling directivity pattern was presented in [2]. More recently, the idea of wave field synthesis (WFS) [3] [4] is presented. The theory is based on the Huygens principle which states that every wavefront can be decomposed into a superposition of elementary spherical wavefronts emitted from secondary sources. In a WFS array, each loudspeaker is independently controlled in order to operate as a secondary source. It is also useful to note that a large number of loudspeakers are often needed in implementing such WFS techniques. The authors of [5][6] presented beamforming techniques using a near-field loudspeaker array constructed from a limited number of loudspeakers.

In this paper, we develop an algorithm for a loudspeaker array using a relatively small number of loudspeakers (16 or less) as compared to WFS. The aim is to compute the loudspeaker weights in order to recreate a desired sound field at predefined listening locations that are determined by the users. In order to achieve this, we first formulate the problem using the propagation matrix of a loudspeaker array. In addition, we show that the conditioning of this problem is determined by the properties of this propagation matrix. We then exploit such properties and propose an algorithm to determine if the predefined listening locations are optimal for sound field recreation. In the event that a non-optimal solution is found, our algorithm will propose a new set of listening locations within the vicinity of the original location in the spatial domain for best reproduction of the sound field. As an alternate application, we can, through the use of this algorithm, select the optimal placement of the loudspeakers in order to achieved a predefined sound field.

2. FAR-FIELD AND NEAR-FIELD ARRAY
Since different techniques are used for far and near-field loudspeaker array design, it is important to review the classification of far and near-field array. In theory, the distance separating near and far-fields is defined by a location in space where the path length differences to all points on the surface of the loudspeaker perpendicular to this location are the same. However, in the case of practical loudspeakers, this distance is infinite. For such practical loudspeakers, this transition point can then be defined by the distance at which the loudspeaker’s three-dimensional radiation balloon no longer changes with increasing distance from the source with regard to frequency. It can also be defined by the distance from the source where the radiated level begins to follow the inverse-square law for all radiated frequencies. These two methods for determining the transition point require accurate anechoic measurements of the loudspeakers. A simplified rule of determining the far and near-field transition point for a single loudspeaker is given by the distance from the loudspeaker to the listening location where the path length difference from points on the loudspeaker to the listening location are within one-quarter wavelength at the highest frequency of interest for a propagating wave. The left panel of Fig. 1 shows an example of this transition distance d, where R is the radius of a loudspeaker diaphragm, and λ is the wavelength at the highest frequency of interest. The transition distance can therefore be computed using

\[ \sqrt{d^2 + R^2} - d = \frac{\lambda}{4} \]  

As oppose to a single loudspeaker, it is important to note that when determining the transition distance for a loudspeaker array, the vibrating surface of the array spans over all the loudspeakers. As a result of this, the radius of the loudspeaker diaphragm R in (1) is now replaced by half the length of the linear loudspeaker array. Consequently, the transition distance increases significantly. The right panel of Fig. 1 shows an illustrative example of a typical line array loudspeaker. The corresponding transition distance is plotted in Fig. 2.

As can be seen from Fig. 2, at a typical distance from 2 to 10 m away from the loudspeaker array, the listeners are in the near-field of the loudspeaker array across most of the frequencies. It is therefore important to note that for a...
not negligible as shown in Section 2.

From the listening location to each of the loudspeakers are differences in distance direction. However in the near-field such as occur in most scenarios for a loudspeaker array, the differences in distance are temporarily omitted for clarity of presentation. The determination of the weight vector can be computed using

$$\hat{w} = G^H y. \quad (8)$$

However, in most scenarios, the number of loudspeakers is not equal to the listening locations. For the case when $N > M$, the array weight vector can be calculated by

$$\tilde{w} = yG^+. \quad (9)$$

where $G^+$ is the pseudo-inverse of the propagation matrix

$$G^+ = (G^H G)^{-1} G^H. \quad (10)$$

Similarly, if $N < M$, the array weight vector can be computed using

$$\tilde{w} = yG^+, \quad (11)$$

and

$$G^+ = G^H (GG^H)^{-1}. \quad (12)$$

As can be seen, the determination of the weight vector $\tilde{w}$ is limited by the conditioning of the matrix $G^H G$ or $GG^H$. Any ill-conditioning of this matrix will render the computation of $\tilde{w}$ being susceptible to numerical errors, it is therefore important to determine how well-conditioned a particular system is before computing $\tilde{w}$. The techniques introduced in the following section apply to both $G^H G$ and $GG^H$. We will use $G^H G$ for illustration.

3. THE PROPOSED ARRAY SYSTEM DESIGN

We now formulate the problem of finding a set of loudspeaker array weights so as to generate a required sound field at a predefined position. As described in Section 2, our focus is on algorithms for the near-field. We consider the problem of controlling the sound field in a few listening locations within a pre-determined area of interest. As discussed in Section 1, several works in the literature have addressed this problem but mainly for the far-field case. In particular, several methodologies employ the well-known “delay-and-sum” technique for beamforming purpose. To achieve this, a vector of complex coefficients is applied to the loudspeakers to selectively project the undistorted signals towards a desired direction. However in the near-field such as occur in most scenarios for a loudspeaker array, the differences in distance from the listening location to each of the loudspeakers are not negligible as shown in Section 2.

We first consider an array of $N$ loudspeakers with a weight vector given by

$$w = [w_1 \; w_2 \; \ldots \; w_N], \quad (2)$$

for which we would like to compute. The frequency indices are temporarily omitted for clarity of presentation. The desired response at the $M$ listening locations are defined as

$$y = [y_1 \; y_2 \; \ldots \; y_M]. \quad (3)$$

Exploiting the near-field characteristics for the array, we model the transfer function from the $i^{th}$, $i = 1, \ldots, N$, loudspeaker to the $j^{th}$, $j = 1, \ldots, M$, listening location using the Green’s function [7]

$$G_{ij} = \frac{1}{4\pi r_{ij}} e^{-jk\omega r_{ij}}, \quad (4)$$

where $r_{ij}$ is the distance from the $i^{th}$ loudspeaker to the $j^{th}$ listening location, $k = \omega/c$ is the wave number and $c$ is the speed of sound in air. Since the acoustic waves from each of the loudspeakers are superimposed at each of the listening location, we assume the signal input matrix is a unity matrix. Consequently, the response at the $j^{th}$ listening location is given by

$$y_j = \sum_{i=1}^{N} w_i G_{ij}. \quad (5)$$

Formulating above using matrix notation, we obtain

$$y = wG. \quad (6)$$

The $N \times M$ propagation matrix

$$G = \begin{bmatrix} G_{11} & G_{12} & \cdots & G_{1M} \\ G_{21} & G_{22} & \cdots & G_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ G_{N1} & G_{N2} & \cdots & G_{NM} \end{bmatrix} \quad (7)$$

is therefore a collection of all the Green’s function from each loudspeaker in the array to each listening location. It is useful to note that when the number of loudspeakers is equal to the number of listening locations ($N = M$), the array weight vector can be computed using

$$\hat{w} = yG^{-1}. \quad (8)$$

However, in most scenarios, the number of loudspeakers is not equal to the listening locations. For the case when $N > M$, the array weight vector can be calculated by

$$\tilde{w} = yG^+. \quad (9)$$

where $G^+$ is the pseudo-inverse of the propagation matrix

$$G^+ = (G^H G)^{-1} G^H. \quad (10)$$

Similarly, if $N < M$, the array weight vector can be computed using

$$\tilde{w} = yG^+, \quad (11)$$

and

$$G^+ = G^H (GG^H)^{-1}. \quad (12)$$

As can be seen, the determination of the weight vector $\tilde{w}$ is limited by the conditioning of the matrix $G^H G$ or $GG^H$. Any ill-conditioning of this matrix will render the computation of $\tilde{w}$ being susceptible to numerical errors, it is therefore important to determine how well-conditioned a particular system is before computing $\tilde{w}$. The techniques introduced in the following section apply to both $G^H G$ and $GG^H$. We will use $G^H G$ for illustration.
4. CONDITIONING OF THE PROPAGATION FUNCTION

As shown in Section 3, and reported in [6], the condition number of the matrix $G^H G$ is critical in achieving a practical and stable solution to the loudspeaker weight vector. If the condition number is too high, a small error in calculating the array weight vector or any slight movement of the listening locations can result in a significant difference between the desired and achieved sound pressure level. To address this issue, we further examine the condition number of the matrix $G^H G$. Using (7), matrix $G^H G$ can be defined as

$$
G^H G = \begin{bmatrix}
G^H G_{11} & G^H G_{12} & \cdots & G^H G_{1M} \\
G^H G_{21} & G^H G_{22} & \cdots & G^H G_{2M} \\
\vdots & \vdots & \ddots & \vdots \\
G^H G_{M1} & G^H G_{M2} & \cdots & G^H G_{NM}
\end{bmatrix},
$$

where

$$
G^H G_{ij} = \sum_{n=1}^{N} G_{in} G_{nj}.
$$

One of the most popular ways of determining the conditioning of a problem is to exploit the $l_2$-norm of the matrix. The $l_2$-norm condition number of the matrix $G^H G$ is defined as

$$
\chi_2[G^H G] = \left\| G^H G \right\|_2 \left\| (G^H G)^{-1} \right\|_2 ,
$$

where $\left\| \cdot \right\|_2$ is the $l_2$-norm. Although the computation of $\chi_2[G^H G]$ is popular, we consider an example case where there are $N = 10$ loudspeakers and $M = 10$ desired listening locations. With this, it is then required to compute the condition number $\chi_2[G^H G]$ of a $10 \times 10$ matrix for each frequency. For the case where the frequency of interest spans across a wide audio range, the computation of the $l_2$-norm condition number can be demanding.

In order to reduce computational load of the algorithm, we propose to employ the E-norm condition number [8]. As will be shown below, the motivation of using this E-norm condition number is to increase the efficiency in determining the conditioning of $G^H G$ across a wide range of frequencies. The E-norm of a $M \times M$ matrix $R$ can be defined as

$$
\| R \|_E = \left\{ \frac{1}{M} \text{tr} \left[ R^H R \right] \right\}^{1/2}.
$$

Similar to the Frobenius norm, the E-norm belongs to an entrywise norm. Employing (16), we further define

$$
\| R^{1/2} \|_E = \left\{ \frac{1}{M} \text{tr} \left[ R \right] \right\}^{1/2},
$$

and

$$
\| R^{-1/2} \|_E = \left\{ \frac{1}{M} \text{tr} \left[ R^{-1} \right] \right\}^{1/2}.
$$

The E-norm condition number of $R^{1/2}$ is then defined as

$$
\chi_E \left[ R^{1/2} \right] = \| R^{1/2} \|_E \| R^{-1/2} \|_E^{-1}.
$$

It has been shown in [8] that $\chi_E \left[ R^{1/2} \right]$ varies monotonically with $\chi_2[R]$. As a result, $\chi_E \left[ R^{1/2} \right]$ can be used as a measure of $\chi_2[R]$. Using the definitions above, we obtain the E-norm condition number matrix as follows

$$
\chi_E \left[ (G^H G)^{1/2} \right] = \left\| (G^H G)^{1/2} \right\|_E \left\| (G^H G)^{-1/2} \right\|_E^{-1},
$$

where

$$
\left( (G^H G) \right)^{1/2} = \left\{ \frac{1}{M} \text{tr} \left[ G^H G \right] \right\}^{1/2}
$$

and

$$
\left( (G^H G)^{-1} \right)^{1/2} = \left\{ \frac{1}{M} \text{tr} \left[ (G^H G)^{-1} \right] \right\}^{1/2}.
$$

In order to compute the E-norm condition number defined in (20), we compute $\text{tr}(G^H G)$ as required by (21). This can be obtained by

$$
\text{tr}(G^H G) = \sum_{m=1}^{M} G^H G_{mm}
= \sum_{m=1}^{M} \sum_{n=1}^{N} G_{nm}^* G_{mn}.
$$

This implies that $\text{tr}(G^H G)$ can be obtained by the summation of the squared elements in the propagation matrix $G$. More importantly, as will be shown in Section 5, $\text{tr}(G^H G)$ is constant across frequency for a fixed array setup. Consequently, the first term on the right hand of (20) is independent of the operating frequency and hence, we can compute (23) for only one operating frequency in order to determine how well-condition $G^H G$ is across the entire audible frequency range of interest. This results in substantial savings in computation since we do not need to compute this condition number across the entire frequency range of interest.

To compute (22), we exploit the eigenvalue decomposition (EVD) of $G^H G$ defined by

$$
G^H G = QAQ^{-1},
$$

where $Q$ is the square matrix whose columns are the eigenvectors of $G^H G$, and $A$ is the diagonal matrix whose diagonal elements are the corresponding eigenvalues. The inverse of the matrix can then be expressed as

$$
(G^H G)^{-1} = QA^{-1}Q^{-1}.
$$

As a result, the trace of $(G^H G)^{-1}$ can be found as

$$
\text{tr} \left\{ (G^H G)^{-1} \right\} = \text{tr} \left\{ QA^{-1}Q^{-1} \right\}
= \text{tr} \left\{ QQ^{-1}A^{-1} \right\}
= \text{tr} \left\{ A^{-1} \right\}
= \sum_{i=0}^{K} \frac{1}{\lambda_i},
$$

where $\lambda_i$ and $K$ are the eigenvalues and rank of the matrix $G^H G$. Using the above technique, condition number can be obtained given any listening locations. Furthermore, in the event when our algorithm detects that the matrix $G^H G$ is ill-conditioned for a particular set of predetermined listening locations, our proposed algorithm then employs an optimization method to search for listening locations that are near the desired listen location that is better conditioned for the computation of $\tilde{w}$. With this new solution, users can then decide if the proposed listening locations are acceptable for their listening application.
5. SIMULATION RESULTS

We illustrate the performance of our proposed algorithm in the context of sound field reproduction using a linear loudspeaker array. As shown in Fig. 3, we employ eight loudspeakers positioned linearly with an array length of 2 m. As an illustration, the number of listening locations is selected as $M = 3$. It is important to note that the number of listening locations is not limited to three and can be extended to accommodate more users. Furthermore, for smoother transition between the listening locations, a larger number is suggested [6]. The technique used to generate the loudspeaker array weights remains the same.

Before investigating on the placement of the listening points, the following simulation illustrates the property of the technique developed in Section 4. Figure 4 shows the variation of the $l_2$-norm condition number $\chi^2$ of $G^H G$ and the square of E-norm condition number $\chi^2$ of $G^H G$ across frequencies for the same array and listening locations across different frequencies. It can be seen that $\chi^2$ is monotonic with the conventional $l_2$-norm condition number $\chi^2$ of the matrix $G^H G$. This implies that any variation in $\chi^2$ will give a good indication of the conditioning of the system in order to compute the weights $\hat{w}$ given by (9) or (11).

Figure 5 shows how $\log_{10} \chi^2$ varies with frequencies for different spacing ranging from 0.1 to 0.5 m between the listening locations as shown in Fig. 3. We can see from this result that $\chi^2$ is high for the low frequency range. In addition, $\chi^2$ reduces with increasing spacing between the listening locations. Similar results can be obtained by keeping the listening point static while changing the inter-spacing between the loudspeaker units in the array. This result implies that spatial diversity is required in order to achieve good reproduction of a sound field particularly in the low frequencies. For higher frequencies beyond 3 kHz however, this dependence is not as significant.

Figure 6 shows the variation of $\log_{10} \chi^2$ across frequencies for different loudspeaker spacing. Similar to the previous case of varying the spacing between listening location, we note that $\log_{10} \chi^2$ reduces with increasing loudspeaker spacing. The two results of Figs. 5 and 6 illustrate the mutual relationship between spacings of the listening locations and the loudspeaker units. In addition to the above, it can be seen from Fig. 5 and 6 that $\chi^2$ is dependent on the loudspeaker array setup, listening locations and the operating frequencies. It is important to note that, since the operating frequency is largely dependent on the source and not within user’s control, we focus on the placement of the listening locations rather than the operating frequency. This implies that we can determine the conditioning of the problem by searching within a spatial location as will be shown later in this section. Furthermore, as the setup of the listening locations and loudspeaker units are mutually related, the techniques developed in the previous section can be used for our algorithm to search for better placement of listening locations or loudspeaker array.

In order to illustrate the motivation and benefit of the proposed algorithm, we plot the first term of (20) as shown in Fig. 7 using the same setup. It can be seen that the trace of the matrix $G^H G$ is constant and is independent of the operating frequency. As a result, for each setup, this term only needs to be computed once for all the operating frequencies and hence computational efficiency is achieved.

It is important to note that $\chi^2$ is dependent
Figure 6: Variation of $\log_{10} \lambda_2 \left( \left( G^H G \right)^{1/2} \right)$ with different frequencies for different inter-loudspeaker spacing.

Figure 7: Trace of the matrix $G^H G$ across frequencies for a fixed array setup.

on the listening locations predefined. In the case when a set of predefined listening locations renders an ill-conditioned matrix $G^H G$, we utilize the techniques developed in Section 4, to find listening locations that give rise to a better conditioned matrix. This allows one to compute $\hat{w}$ according to (9) or (11). User can also set a search range within the vicinity of the desired positions. To illustrate this example, we show, in Fig. 8, a scenario where the search range is set as maximum of 1.5 m away from the predefined initial listening point. In this experiment, the loudspeaker array setup is shown in Fig. 3. We consider an example case where the operating frequency is 1 kHz. Using our proposed E-norm condition number, an iterative algorithm similar to that presented in [9] was implemented to search for a minimum of the E-norm condition number. In addition, the search was constrained to be within 1.5 m from the initial listening location. The numbers shown in Fig. 8 depict the evolution of the suggested listening positions at every iteration of the optimization algorithm. It can be seen from this result that the optimal position obtained is at the position where the local minimum of E-norm condition number is achieved. This position corresponds to the vicinity of the initial point selected. With this new listening location, users can now consider moving to this new location within the vicinity of the initial location in order to enjoy the predefined sound field set by him or her.

6. CONCLUSIONS

In this paper we present a novel approach to search for an optimal array configuration to avoid the ill-condition problem in near-field loudspeaker array design. Conventionally, the condition number is computed using $l_2$-norm. We proposed an efficient way of computing the E-norm condition number. Results obtained using both $l_2$-norm and E-norm condition number are shown. From the results, it can be seen that using the proposed E-norm condition number significantly reduces the computational complexity. Additional experiments suggest that the results obtained using the E-norm condition number is equivalent to that using the $l_2$-norm condition number.

REFERENCES