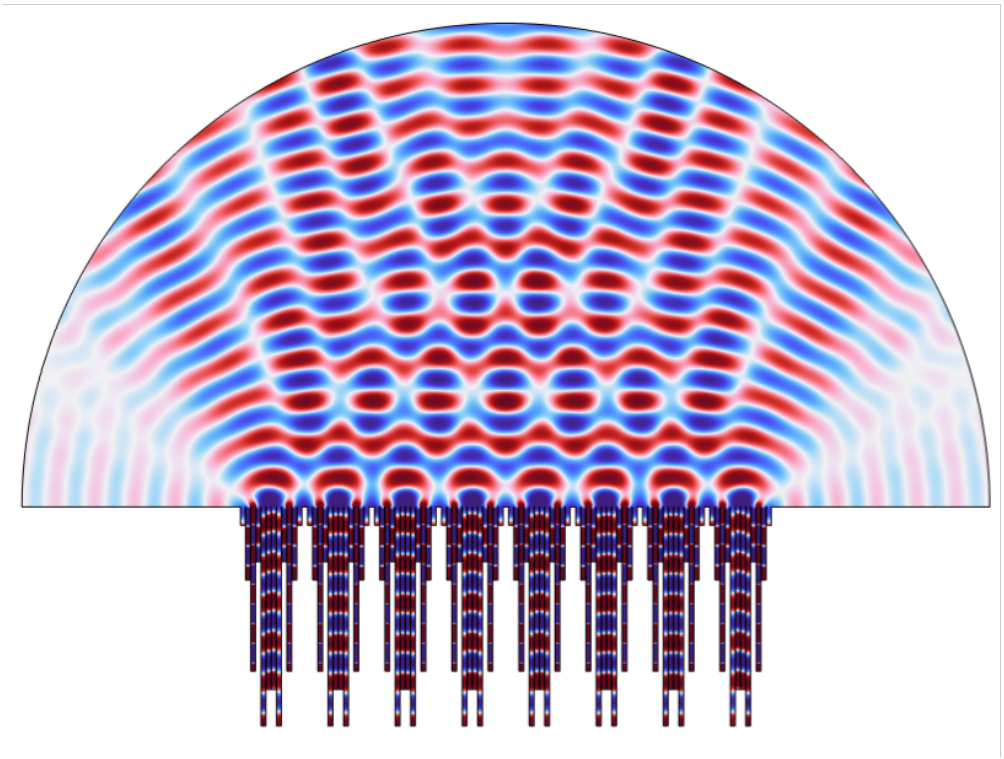


# Enlargement of Perceived Size of Sound Source using Diffusive Geometries

Master's thesis in Master Programme Sound and Vibration

ZHIWEN XU

DEPARTMENT OF ARCHITECTURE AND CIVIL ENGINEERING

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2024
www.chalmers.se

MASTER'S THESIS ACEX30

# Enlargement of Perceived Size of Sound Source using Diffusive Geometries

ZHIWEN XU


CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Architecture and Civil Engineering
*Division of Applied Acoustics*
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2024

Enlargement of Perceived Size of Sound Source using Diffusive Geometries
ZHIWEN XU



Supervisor: JensAhrens, Division of Applied Acoustics
Examiner: JensAhrens, Division of Applied Acoustics

Master's Thesis 2024
Department of Architecture and Civil Engineering
Division of Applied Acoustics
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 72 443 9873

Typeset in LaTeX
Chalmers Reproservice
Gothenburg, Sweden 2024

Enlargement of perceived size of sound source using diffusive geometries
ZHIWEN XU
Department of Architecture and Civil Engineering
Division of Applied Acoustics
Chalmers University of Technology

# Abstract

Interaural cross correlation, which quantifies the similarity of binaural signals, is essential for immersive and rendered auditory environment. Previous research has proven that reducing interaural correlation leads to expansion of perceived size of a source. The Schroeder diffuser is a prevalently used acoustic design to create pleasant acoustic environment since it was firstly developed. It helps to eliminate many acoustic problems in many scenarios through scattering waves in multiply directions. In this work, the Schroeder diffuser is employed to manufacture natural sound field with reduced interaural correlation so that models of source enlarger can be studied further.

To explore the properties of the sound field of the Schroeder diffuser, the produced sound field is simulated and a method of representing the sound field with circular harmonic expansion is applied. The characteristics of the sound field of the Schroeder diffuser are investigated to discover their connectivity with low interaural correlation in order that any other arbitrary sound field can be manipulated to achieve it. In light of such discovery, two models of simple sound fields which are impulse field and plane wave field are utilized to produce the sound field with low interaural correlation with proper modification engaged. These simple models are capable of saving computational cost as simulating the Schroeder diffuser is pretty resource and time consuming.

**Keywords:** Spatial Audio, Diffusive Geometry, Schroeder Diffuser, Circular Harmonic Expansion, Binaural Analysis, Digital Signal Processing

# Contents


**1 Introduction** **1**

1.1 Interaural cross correlation [1] . . . . . 1

1.2 Simulation work [2] . . . . . 1

**2 Theory** **3**

2.1 Schroeder diffuser . . . . . 3

2.1.1 Design of Schroeder diffuser [4] . . . . . 3

2.1.2 Autocorrelation function . . . . . 4

2.2 Acoustic Virtual Environment [7] . . . . . 4

2.2.1 Sound field Synthesis . . . . . 4

2.2.2 Acoustic wave equation [7] . . . . . 5

2.2.3 Polar Coordinate [6] [7] . . . . . 6

2.2.4 Circular harmonic expansion [7] . . . . . 7

2.3 Interaural cross correlation [8] . . . . . 11

**3 Method** **13**

3.1 Schroeder diffuser simulation . . . . . 13

3.1.1 Simulation method [9] . . . . . 13

3.1.2 Simulation example from previous work [2] . . . . . 14

3.2 Circular harmonics expansion algorithm . . . . . 16

3.3 Artificial Field Manufacturing . . . . . 19

**4 Results** **20**

4.1 Schroeder diffuser simulation . . . . . 20

4.1.1 Autocorrelation . . . . . 24

4.1.2 Inter-aural cross correlation . . . . . 25

4.2 Circular harmonic expansion . . . . . 26

4.3 Artificial field . . . . . 33
4.3.1 Impulse response . . . . . 33
4.3.2 Plane wave . . . . . 36

**5 Discussion** **40**

**6 Conclusion** **41**

# 1 Introduction

Spatial audio technology, which creates immersive auditory environment in 3-D space, is quite different from traditional stereo audio that only conveys left and right. However, it is the binaural signal that finally brings virtual feeling to our auditory system, whatever technologies are employed.

The binaural signal has to be decorrelated so that people's brains can detect the differences between signals played at two ears and render virtual sense. Interaural cross correlation (IACC) has been referred widely to quantify the similarity between binaural signals. Previous research has discussed the relationship between IACC and appparent source width (ASW) that reducing IACC leads to increase of ASW [1]. ASW describes the perceived size of a sound image such as the width of orchestra in concert hall. It's employed here to represent the perceived size of the sound source.

The Schroeder diffuser is a well developed acoustic design to create pleasant acoustic environments. It's able to scattering incident waves into multiply directions in order that acoustic problems caused by specular reflections can be purged from many scenarios. In this work, the Schroeder diffuser is utilized to manufacture natural sound field that may convey low interaural cross correlation in such manner that models of source enlarger can be studied with it as a basis. The simulations of the Schroeder diffusers are performed to explore the properties of its scattering sound field and find its connectivity with low IACC. Benefit from circular harmonics expansions method, such nature field by the Schroeder diffuser can be easily decomposed and reconstructed with circular harmonic coefficients, which further enables any arbitrary sound field such as impulse field and plane wave field to be manipulated to achieve low IACC.

## 1.1 Interaural cross correlation [1]

In a previous work the performance of IACC to predict ASW was test with a stero set-up. Three experiments were performed to test if IACC predicts ASW properly: 1) loudspeaker presentation in a listening room; 2) headphone presentation of the same listening room; 3) headphone presentation of an anechoic listening condition. All the three experiments are conducted under $0.25\,\mathrm{kHz}$ and $1\,\mathrm{kHz}$ with and without cross-talk.

The measurements showed that the IACC were consistent with measured data without cross-talk, whereas the prediction with cross-talk didn't account for the measurement. Since cross-talk is not considered in this work. The IACC is still a great predictor of perceived size.

## 1.2 Simulation work [2]

Another previous work performed simulations of diffusive geometries, which is the foundation of simulation part in this work. Simulations of a 2-D diffuser with $(8 \times 8)$ wells are performed to study its scattering characteristics, and autocorrelation diffusion coefficient is

employed to estimate how diffused a scattering field is. Detailed theoretical derivations and implementations are presented, which will be discussed in later section 3.

# 2 Theory

## 2.1 Schroeder diffuser

Schroeder diffuser is an optimized surface [3] proposed by Schroeder, which has ability to scatter sound wave omnidirectionally from whichever direction the incident wave comes [4]. It's constructed by digging a series of equal-width wells with different depths on a wall. The depths of such wells are determined by quadratic-residue sequence [5], therefore Schroeder diffuser is also called quadratic-residue diffuser (QRD). In a theoretical environment without air friction, sound field only depends on the travelling routes of component sound waves because only phase differences of such component waves would bring changes to net field. Schroeder diffuser reconstructs the sound field exactly by changing the propagation paths of incident waves.

### 2.1.1 Design of Schroeder diffuser [4]

The valid bandwidth of Schroeder diffuser depends on physical dimensions including width and depths of wells. Based on previous studies of acoustic diffusers, the width $w$ of wells is defined as:

$$w = \frac{\lambda_{min}}{2} \tag{1}$$

where $\lambda_{min}$ is the wavelength of highest frequency for optimal diffusion. The lower frequency limit $f_0$ is determined by designer, which also control the depth of wells $D_n$:

$$D_n = \frac{\lambda_0}{2N} s_n \tag{2}$$

where $\lambda_0 = c_0/f_0$, $N$ is a prime number chosen by designer, and $s_n$ is quadradic-residue sequence. For 1D Schroeder diffuer the sequence is:

$$s_n = n^2 \bmod N \tag{3}$$

while that of a 2D Schroeder diffuser is:

$$s_{n,m} = (n^2 + m^2) \bmod N \tag{4}$$

where $n$ and $m$ are non-negative integers which determines the depth of $n\mathrm{th}$ and $m\mathrm{th}$ well on the wall.

### 2.1.2 Autocorrelation function

The autocorrelation function is frequently employed to investigate the spatial similarity of energy scattered from diffuser. Diffuser which creates more diffused field would return a higher value from autocorrelation function. The diffusion coefficient can be derived from the function through taking an average on it, and the equation of it is :

$$d_\Psi = \frac{(\sum_{i=1}^{n} 10^{L_i/10})^2 - \sum_{i=1}^{n}(10^{L_i/10})^2}{(n-1)\sum_{i=1}^{n}(10^{L_i/10})^2} \tag{5}$$

where $\Psi$ represents the incidence angle, $L_i$ is the sound pressure level in $i\mathrm{th}$ direction, and the scattered waves radiate in $n$ directions. $d_\Psi$ will be a positive value located between $1/n$ and 1, and larger value means more diffused field.

**Sound Pressure Level** The mathematical definition of sound pressure level is [6]:

$$L_\mathrm{p} = 10\log_{10}(\frac{\mathrm{p}}{2\times 10^{-5}\,\mathrm{Pa}})^2 \tag{6}$$

where $\mathrm{p}$ represents sound presusre.

## 2.2 Acoustic Virtual Environment [7]

Virtual environment has become a common concept in both video and audio industries. It's an immersive sensory feeling created by stimulating people's watching and hearing systems. It's usually a holistic experience combining the technologies of both video virtual environment and acoustics virtual environment. In this section, the scope is narrowed down to acoustic virtual environment, where studies are focused on audio reproduction technologies for rendering methods.

The most common rendering technologies are head related methods and room related methods. Head related methods only create the sound signlas at both ears, while room related methods creates a sound field in a room in the virtual environment so that listeners can pick whichever positions they want to listen at.

### 2.2.1 Sound field Synthesis

Sound field synthesis is a newly developed room related method. The main idea of sound field synthesis is to decompose the desired sound field with a set of elementary sound sources, of which the superposition of emitted fields is an excellent approximation of such desired field. The elementary sources are distributed continuously in space to supply information, thus in theory, the synthesized field depends highly on the density and layout of

sources. Larger room usually requires more elementary sources, otherwise the deviations between superpositions and desired field would be unacceptable. In this thesis, study is focused on 1D diffuser, and the exterior sound field is in a 2D surface. As a result, all the results and analysis is based on 2D polar coordinate.

### 2.2.2 Acoustic wave equation [7]

Acoustic wave equation is the patial-differential equation that describes the relations between acoustic properties and space and time, from which the propagation properties of acoustic waves can be solved and studied. Characteristics of a sound field are usually represented by three fundamental physical properties:

- $p(\boldsymbol{r}, t)$, sound pressure
- $\boldsymbol{v}(\boldsymbol{r}, t)$, partical velocity
- $\rho(\boldsymbol{r}, t)$, mass density

where $\boldsymbol{r} = (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ is the displacement vector in 3-D Cartesian coordinate, $t$ is the time at observation.

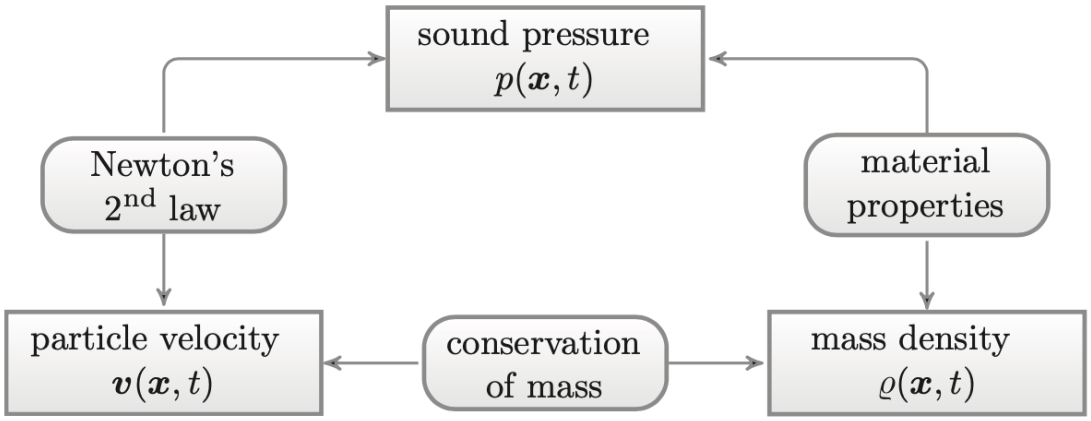



Figure 1: Physical principles between three fundamental acoustic properties [7].

These three fundamental acoustic properties are related with each other with the physical principles shown in figure 1. And three equations can be used to represent such principles:

- Equation of motion, derived from Newton's 2nd law, describes relationship between sound pressure $p(\boldsymbol{r}, t)$ and particle velocity $\boldsymbol{v}(\boldsymbol{r}, t)$.
- Equation of continuity, equivalent to conservation of mass, connects particle velocity $\boldsymbol{v}(\boldsymbol{r}, t)$ and mass density $\rho(\boldsymbol{r}, t)$ by considering the mass flux through a volume element
- Equation of state for a gas, referring to thermodynamics properties, solves the association between $p(\boldsymbol{r}, t)$ and $\rho(\boldsymbol{r}, t)$ with heat capacity of medium gas.

**Wave equation** Combine the three equations one can conclude the general form of wave equation as,

$$\nabla^2 p(\boldsymbol{r}, t) - \frac{1}{c^2}\frac{\partial^2 p(\boldsymbol{r}, t)}{\partial t^2} = 0 \tag{7}$$

based on which the solutions are acquired in COMSOL Multiphysics. And the mathematical form of laplacian operator is $\nabla^2$ is

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{8}$$

### 2.2.3 Polar Coordinate [6] [7]

In this thesis, the physical image of sound field is constructed in a 2-D plane, and for the convenience of using circular harmonic expansion, the wave equation has to be transformed into polar coordinate.

In polar coordinate the displacement $\boldsymbol{r}$ is described with two coefficients $r$ and $\alpha$, which are the distance from the origin point and angle distance from $0$, respectively.

As a result, the displacement $\boldsymbol{r}$ can be transformed into

$$\boldsymbol{r} = \begin{bmatrix} x \\ y \end{bmatrix} = r \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix} \tag{9}$$

and

$$\boldsymbol{r} = |r|^2 = x^2 + y^2 \tag{10}$$

$$\tan\alpha = \frac{y}{x} \tag{11}$$

The laplacian operator in polar coordinate is

$$\nabla^2 = \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2}{\partial \alpha^2}, \tag{12}$$

which brings the form of wave equation into

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial p(\boldsymbol{r}, t)}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 p(\boldsymbol{r}, t)}{\partial \alpha^2} - \frac{1}{c^2}\frac{\partial^2 p(\boldsymbol{r}, t)}{\partial t^2} = 0 \tag{13}$$

### 2.2.4 Circular harmonic expansion [7]

In previous section 2.2.2, the wave equation 7 is derived and rearranged in polar coordinate 13. The general solution of wave equation is

$$p(\boldsymbol{r}, t) = \tilde{P}(\boldsymbol{k}, \omega) e^{i(\omega t + \boldsymbol{k}^{\mathrm{T}} \cdot \boldsymbol{r} + \phi)} \tag{14}$$

where $\omega$ is the angular frequency of wave and $\phi$ is initial phase. $\boldsymbol{k}$ denotes the wave vector at position of interest. The wave vector $\boldsymbol{k}$ can be expressed by wave number $k = \dfrac{2\pi}{\lambda}$ and unit wave vector $\hat{\boldsymbol{n}}$

$$\boldsymbol{k} = \|\boldsymbol{k}\| \hat{\boldsymbol{n}} = k \hat{\boldsymbol{n}}, \tag{15}$$

and the unit vector is defined as

$$\hat{\boldsymbol{n}} = \left( \frac{x}{\sqrt{x^2 + y^2 + +z^2}}, \frac{y}{\sqrt{x^2 + y^2 + +z^2}}, \frac{z}{\sqrt{x^2 + y^2 + +z^2}} \right). \tag{16}$$

**Fourier transformation** As one can find in equation 14, the solution of wave equation is a function of both position $\boldsymbol{r}$ and time $t$. The properties of sound wave such as sound pressure fluctuates with spatial position and time. As one of the most fundamental methods regarding signal processing, Fourier transformation is capable of decomposing signals into constituent frequencies or wave numbers, respectively to time and spatial positions. Because any arbitrary sound field can be composed by several sinusoidal waves from some directions.

As a consequence, applying Fourier transformation to either time or positions will help to indicate the components of such waves through investigating its spectrum. This is quite essential when evaluating properties of signals as spectrum reports how different components contribute to the total signal.

For both spatial position and time, there's a Fourier transformation pair. The transformation relationships can be expressed as

$$\boxed{p(\boldsymbol{r}, t) \overset{\mathscr{F}_t}{\longleftrightarrow} P(\boldsymbol{r}, \omega) \overset{\mathscr{F}_{\mathbf{k}}}{\longleftrightarrow} \tilde{P}(\boldsymbol{k}, \omega)} \tag{17}$$

The pair for time is

$$P(\boldsymbol{r}, \omega) = \mathscr{F}_t\{p(\boldsymbol{r}, t)\} = \int_{-\infty}^{\infty} p(\boldsymbol{r}, t) e^{-i\omega t} \, \mathrm{d}t \tag{18}$$

$$p(\boldsymbol{r}, t) = \mathscr{F}_t^{-1}\{P(\boldsymbol{r}, \omega)\} = \frac{1}{2\pi} \int_{-\infty}^{\infty} P(\boldsymbol{r}, \omega) e^{i\omega t} \, \mathrm{d}\omega \tag{19}$$

while for spatial positions the pair becomes

$$\tilde{P}(\boldsymbol{k},\omega) = \mathscr{F}_{\boldsymbol{r}}\{P(\boldsymbol{r},\omega)\} = \int_{-\infty}^{\infty} P(\boldsymbol{r},\omega)e^{-ik^{\mathrm{T}}\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{r} \tag{20}$$

$$P(\boldsymbol{r},\omega) = \mathscr{F}_{\boldsymbol{r}}^{-1}\Big\{\tilde{P}(\boldsymbol{k},\omega)\Big\} = \frac{1}{(2\pi)^d}\int_{-\infty}^{\infty} P(\boldsymbol{k},\omega)e^{ik^{\mathrm{T}}\cdot\boldsymbol{r}}\,\mathrm{d}\boldsymbol{k} \tag{21}$$

where $d$ is the degree of freedom.

The mechanism of Fourier transformation adopts the orthogonality of components that the inner product of two orthogonal vectors equals to 0.

**Monofrequency plan wave** As discussed above, any complex sound field can be regarded as a superposition of many plance waves. Based on the general solutino of wave equation, consider a monofrequency plane wave with frequency $\omega_0$ propagating in a fixed direction $\varphi_0$. In polar coordinate such wave can be described as

$$p(\boldsymbol{r},t,\omega_0,\varphi_0) = \tilde{P}(\omega_0,\varphi_0)e^{i(\omega_0 t+\boldsymbol{k}_0^{\mathrm{T}}\cdot\boldsymbol{r})} \tag{22}$$

where the initial phase $\phi$ is eliminated for convenience. Refer to the coordinate relationship above 9, and at meantime unit wave vector

$$\hat{\boldsymbol{n}}_{\boldsymbol{0}} = \begin{bmatrix}\cos\varphi_0\\ \sin\varphi_0\end{bmatrix}, \tag{23}$$

the inner product between wave vector and position is calculated as

$$\boldsymbol{k_0}^{\mathrm{T}}\cdot\boldsymbol{r} = k_0\hat{\boldsymbol{n}}_{\boldsymbol{0}}{}^{\mathrm{T}}\cdot\boldsymbol{r} = \frac{\omega_0}{c}\|\hat{\boldsymbol{n}}_{\boldsymbol{0}}\|\cdot\|\boldsymbol{r}\|\cos(\alpha-\varphi_0) \tag{24}$$

label $\gamma = \alpha - \varphi_0$, then the plane wave described in equation 22 becomes

$$p(\boldsymbol{r},\alpha,t,\omega_0,\varphi_0) = \tilde{P}(\omega_0,\varphi_0)e^{i(\omega_0 t+\frac{\omega_0}{c}r\cos\gamma)}. \tag{25}$$

Since it's in polar coordinate, the pressure has a angular periodicity that

$$p(\boldsymbol{r},\alpha,t,\omega_0,\varphi_0) = p(\boldsymbol{r},\alpha+2\pi,t,\omega_0,\varphi_0) \tag{26}$$

**Bessel functions and orthogonal expansion** To simplify the calculation of function $p(\boldsymbol{r}$,-$\alpha,t,\omega_0,\varphi_0)$, a common method is separating the function into the products of several single

variable functions, or they can be called as orthogonal functions because their inner product equals to 0 over a specific interval and are similar to the concept of orthogonal vectors. These functions are exactly the orthogonal vectors in mathematical manners, while orthogonal vectors in linear algebra can be visualized more easily as they're perpendicular with each other.

In this case the Bessel functions are employed to split the exponential term in equation 25 so that the function is expanded with orthogonal functions. Bessel functinons are the solutions of Bessel's differential equation:

$$x^2\frac{d^2y}{dx^2} + x\frac{dy}{dx} + (x^2 - a^2)y = 0 \tag{27}$$

and it's easy to find that Bessel's differential equation is a second-order normal linear differential equation, which means that there must be two linear solutions. Here the first kind Bessel functions are referred to with its definition using integral representation:

$$i^n J_n(kr) = \frac{1}{2\pi}\int_0^{2\pi} e^{i(kr\cos\gamma - n\gamma)}\,\mathrm{d}\gamma, \quad n \in \mathbb{Z} \tag{28}$$

the exponential term $e^{ikr\cos\gamma}$ is included in the integral, so Bessel's functions can be regarded as the Fourier series coefficients of it, and it can be calculated inversely with the manner of Fourier transformation. Since $n$ is integers, integral has to be turned into summation:

$$e^{ikr\cos\gamma} = \sum_{n=-\infty}^{+\infty} i^n e^{in\gamma} J_n(kr) \tag{29}$$

$$= \sum_{n=-\infty}^{+\infty} i^n e^{in(\alpha-\varphi_0)} J_n(kr) \tag{30}$$

finally equation 25 turns out to be

$$p(\boldsymbol{r}, \alpha, t, \omega_0, \varphi_0) = \tilde{P}(\omega_0, \varphi_0) \sum_{n=-\infty}^{+\infty} i^n e^{i\omega_0 t} e^{in\alpha} J_n(\frac{w_0}{c}r) e^{-in\varphi_0} \tag{31}$$

**Broadband plane waves** Broadband wave can be regarded as a superposition of many monofrequency plane wave with different amplitude. Here the propagation direction of wave is still $\varphi_0$ but the frequency $\omega_0$ is no longer a single value, but a variable which contains all the constituent frequencies. The total pressure of such broadband wave can be calculated through integrating monofrequency plane wave over frequency, which can be described as

$$p_{bb}(\boldsymbol{r}, \alpha, t, \varphi_0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} p(\boldsymbol{r}, \alpha, t, \omega_0, \varphi_0)\, \mathrm{d}\omega_0 \tag{32}$$

insert the equation of monofrequency plane wave 25

$$\begin{aligned} p_{bb}(\boldsymbol{r}, \alpha, t, \varphi_0) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{P}(\omega_0, \varphi_0) e^{i(\omega_0 t + \frac{\omega_0}{c} r\cos\gamma)}\, \mathrm{d}\omega_0 && (33)\\ &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \tilde{P}(\omega_0, \varphi_0) e^{i\frac{\omega_0}{c} r\cos(\alpha-\varphi_0)} e^{i\omega_0 t}\, \mathrm{d}\omega_0 && (34)\\ &= \mathscr{F}_t^{-1}\Big\{ \tilde{P}(\omega_0, \varphi_0) e^{i\frac{\omega_0}{c} r\cos(\alpha-\varphi_0)} \Big\} && (35) \end{aligned}$$

apply Fourier transformation on both side of the equation

$$P_{bb}(\boldsymbol{r}, \alpha, \omega, \varphi_0) = \tilde{P}(\omega, \varphi_0) e^{i\frac{\omega}{c} r\cos(\alpha-\varphi_0)} \tag{36}$$

refer to equation 30 and then the final series expansion is

$$P_{bb}(\boldsymbol{r}, \alpha, \omega, \varphi_0) = \sum_{n=-\infty}^{+\infty} \tilde{P}(\omega, \varphi_0) i^n e^{in(\alpha-\varphi_0)} J_n(\frac{\omega}{c} r) \tag{37}$$

**General sound field** In previous sections the expansions of monofrequency and broadband sound waves have already been discussed in mathematical manners. However, the theoretical description is still not universal enough as the realistic sound field is usually impossible to be constructed by monodirectional wave. With the experience of expanding the scope from monofrequecy wave to broadband waves, the monodirectional wave can be similarly broaden to general sound field through generalizig the propagating directions of sound waves to all angular range $\varphi_0 \in [0, 2\pi]$.

Implement the integral over propagating direction $\varphi_0$

$$P_g(\boldsymbol{r},\alpha,\omega) = \frac{1}{2\pi}\int_0^{2\pi} P_{bb}(\boldsymbol{r},\alpha,\omega,\varphi_0)\,\mathrm{d}\varphi_0 \tag{38}$$

$$= \frac{1}{2\pi}\int_0^{2\pi} \sum_{n=-\infty}^{+\infty} \tilde{P}(\omega,\varphi_0) i^n e^{in(\alpha-\varphi_0)} J_n(\frac{\omega}{c}r)\,\mathrm{d}\varphi_0 \tag{39}$$

$$= \sum_{n=-\infty}^{+\infty} \frac{1}{2\pi}\int_0^{2\pi} \tilde{P}(\omega,\varphi_0) e^{-in\varphi_0}\,\mathrm{d}\varphi_0\, i^n e^{in\alpha} J_n(\frac{\omega}{c}r) \tag{40}$$

$$= \sum_{n=-\infty}^{+\infty} \breve{P}_n(\omega) i^n e^{in\alpha} J_n(\frac{\omega}{c}r) \tag{41}$$

where $\breve{P}_n(\omega)$ is the Fourier coefficients of complex amplitude $\tilde{P}(\omega,\varphi_0)$ with pair

$$\breve{P}_n(\omega) = \frac{1}{2\pi}\int_0^{2\pi} \tilde{P}(\omega,\varphi_0) e^{-in\varphi_0}\,\mathrm{d}\varphi_0 \tag{42}$$

$$\tilde{P}(\omega,\varphi_0) = \sum_{l=-\infty}^{+\infty} \breve{P}_l(\omega) e^{il\varphi_0} \tag{43}$$

Finally, the general sound field $P_g(\boldsymbol{r},\alpha,\omega)$ has is characterised by a Fourier series

$$P_g(\boldsymbol{r},\alpha,\omega) = \sum_{n=-\infty}^{+\infty} \acute{P}_n(\boldsymbol{r},\omega) e^{in\alpha} \tag{44}$$

with its Fourier coefficients written as

$$\acute{P}_n(\boldsymbol{r},\omega) = \breve{P}_n(\omega) i^n J_n(\frac{\omega}{c}r) \tag{45}$$

## 2.3 Interaural cross correlation [8]

The interaural cross correlation function is useful when evaluating the perceived size of source. It can be calculated with pressure signal in time domain with definition as

$$\rho_{t_1,t_2}(\tau) = \frac{\int_{t_1}^{t_2} p_l(t) p_r(t+\tau)\,\mathrm{d}t}{\sqrt{\int_{t_1}^{t_2} p_l^2(t)\,\mathrm{d}t \int_{t_1}^{t_2} p_r^2(t)\,\mathrm{d}t}} \tag{46}$$

and interaural cross correlation (IACC is defined as the maximum value of the absolute value of this function over $\tau$:

$$\mathrm{IACC}_{t_1,t_2} = \max|\rho_{t_1,t_2}(\tau)|, \tau \in (-1,1)\,\mathrm{ms} \tag{47}$$

# 3 Method

In this section, the methods for simulation and algorithm implementation are discussed in a detail way. As mentioned before, the Schroeder diffusers are firstly simulated with computer-aid software, and circular harmonics are then applied to reconstruct the scattering sound field with abstract sound waves, and finally the sound field are manufactured totally artificially to explore more possibilities of designing diffusers. Since such diffusers are not really manufactured, simulations are regarded as predictions of functionality of them. However, the following procedures which are realized with algorithms are more general and have more potential to be applied in real products.

## 3.1 Schroeder diffuser simulation

### 3.1.1 Simulation method [9]

Simulation work are very common in either academical research or industrial R&D. It's very reliable regarding predicting the functionality of theoretical model or industrial products. With built-in functions and theoretical models, simulation software are able to calculate a variety of properties of research objects such as motions status, heat transfer, vibrations and so on. There are many commercial software on the market now to fulfill diverse demands from different research area, but almost all of them employ finite element method (FEM) to solve the systems of interest.

FEM is a well-developed numerical technique to study successive systems. A systematic description of FEM can be found in a book [9]. It's an approximation of real scenarios as the impossibility to construct a completely accurate replacement of them and such approximation is applied through interpolation. A general procedure of FEM is listed below,

- **Discretization** Discertize the domain into smaller elements. The size and shape of elements are controller by the problem itself.
- **Interpolation** Use interpolation function to approximate solutions within elements with values at know nodes.
- **Element Equation Formulation** Formulate equations of elements based on physical principles which control the behaviours of the systems.
- **Equation System Assembly** List the element equations together to form a equation system so that the entire domain is represented.
- **Boundary Condition Application** Apply boundary condition to describe the internal or external constrain onto the domain.
- **Solution** Solve the equation system with various numerical methods.

In commercial software, such procedures are integrated together for more fluent workflow. What researcher needs to do are building demanded objects, dividing objects with proper size of meshes, choosing physical principles and setting correct parameters.

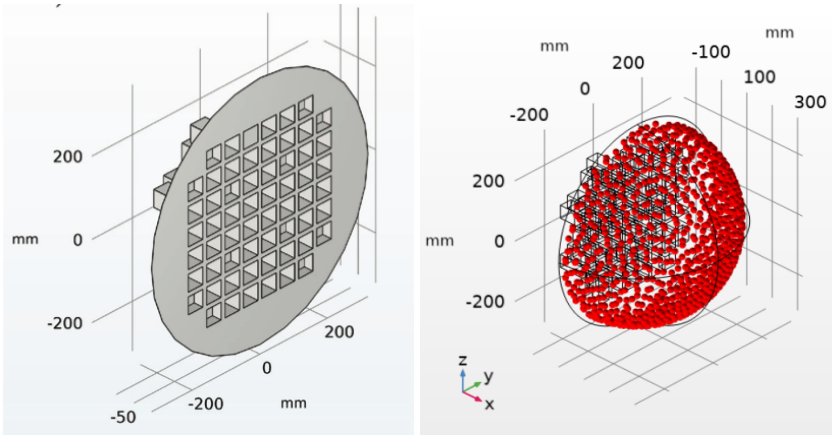
In this thesis, **COMSOL Multiphysics** is used for simulations works. Since the physics image is constructed in a 2-D polar plane, the space dimension in model builder is chosen to be 2-D as a matter of course. 'Pressure Acoustics, Frequency Domain' is chosen as physics interface as it's suitable for linear acoustics simulations. Study type is set as 'Frequency Domain' to acquire results with spectral characteristics.

The sound field produced by diffuser is composed of incident wave and scattered wave. To study the effect of scattered wave, incident wave has to be deduced from total field. In geometries setup, diffuser is placed at the origin position of coordinate because the scattered wave can be regarded to be stimulated by diffuser as a source. This brings much convenience to analyses in polar coordinate. The diffuser is simulated to be placed in an air domain with enough large size that stable sound field can be constructed and prevent the influence of near field effect. At the outer boundary of air domain, a perfect matched boundary is set to absorbed all the energy radiated as it's impossible to simulate an infinitely extending field. When it comes to solving of field, exterior field calculation is activated.

According to the principles discussed in section 2.1.1, diffusers are designed and simulated with settings described above. The solutions exported from simulations are pressure response in frequency domain, in which the information of scattering such as traveling distance and influence of diffuser on input signal is stored. Consider the scattering process of diffuser as a linear time invariant (LTI) system, these frequency responses (FR) of pressure exported are exactly transfer function of the systems.

### 3.1.2 Simulation example from previous work [2]

As mentioned in section 1, a $(8 \times 8)$ diffuser was simulated and autocorrelation diffusion coefficients of were calculated in a previousr work. The geometry and sampling grid are shown in figure 2. The diffuser is compared with a flat hard wall in figure 3 with autocorrelation diffusion coefficient, which represents how diffused the scattering field is. The axes in the figures are corresponding coordinates in figure 2, which means the $1^{\text{st}}$ and $3^{\text{rd}}$ plots in figure 3 are the autocorrelations observing from side view (parallel to +y direction) and the $2^{\text{nd}}$ and $4^{\text{th}}$ plots are looking from from view (-x direction). Each circle in the plots represents an incident direction, and the distance between circle's position and origin point is the value of autocorrelation diffusion coefficient. So the circle farther away from origin point means the scattered wave field with incident wave from such direction is more diffused.




(a) Diffuser geometry (b) Sampling grid

Figure 2: Diffuser geometry and sampling grid in related work [2]

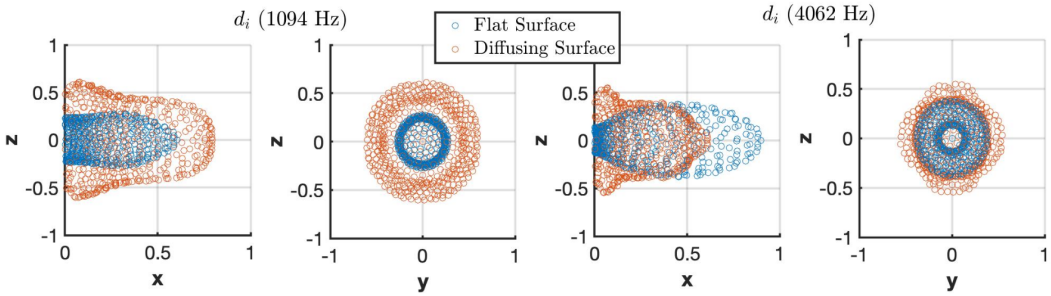



Figure 3: Autocorrelation diffusion coefficient [2]

The simulation of this related work is replicated and the autocorrelation diffusion coefficients are plotted in figure 4 and 5. During the replication, the half-sphere of grid only contains 100 points which are evenly distributed. The accuracy of autocorrelation diffusion coefficient might not be as good as that in figure 3, but the overall shape of distribution is reasonable.

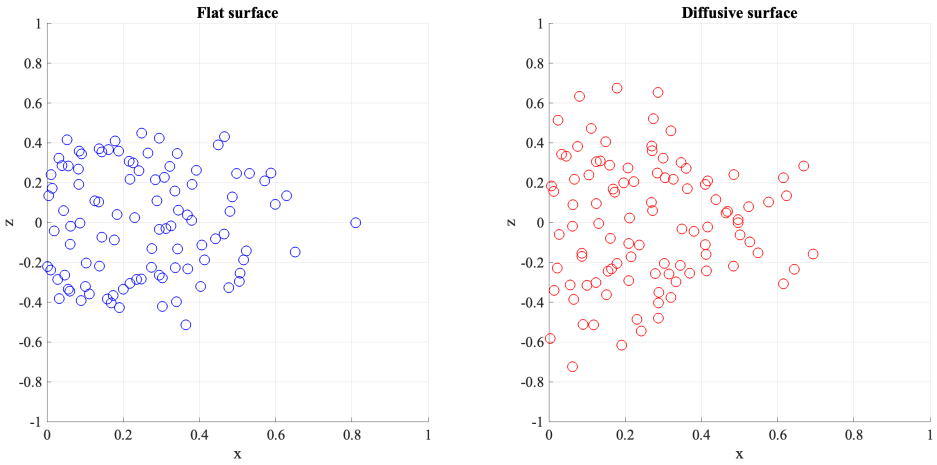



Figure 4: Autocorrelation diffusion coefficient in 1096 Hz

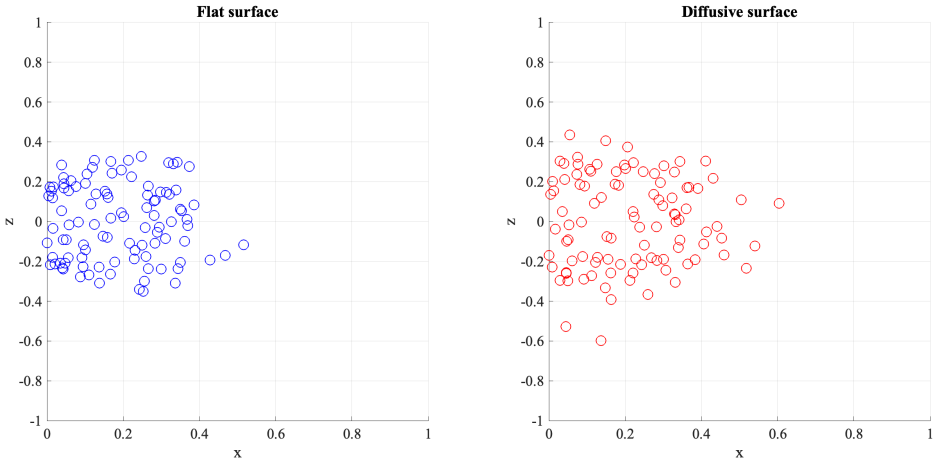



Figure 5: Autocorrelation diffusion coefficient in $4062\,\mathrm{Hz}$

## 3.2 Circular harmonics expansion algorithm

Circular harmonic expansion is similar to Fourier series expansion, and both of them are expansions of periodic functions. Fourier series are expansions of real line functions. With proper chosen Fourier coefficients, any periodic real line functions can be expressed with the sum of Fourier series. Analogously, circular harmonics are expansions of periodic functions defined on circles, which are more appropriate to be defined in polar coordinate with two variables, radial and angular displacement. And the expansion equation is derived using Fourier expansion theory. Even further, periodic functions defined on a surface of sphere can also be expanded with series, which are called spherical harmonics.

The circular harmonics are formulated as [7]

$$CH(m) = e^{im\theta} \tag{48}$$

where $m$ is the order index.

In following figure 6 and 7, circular harmonics of lowest 4 orders are visualized for both real part and imaginary part. In order to plot all values with same radial displacement range, absolute values of circular harmonics are taken, but the real sign of them are labelled with different colors. Red curves represent positive values and blue are for negative.

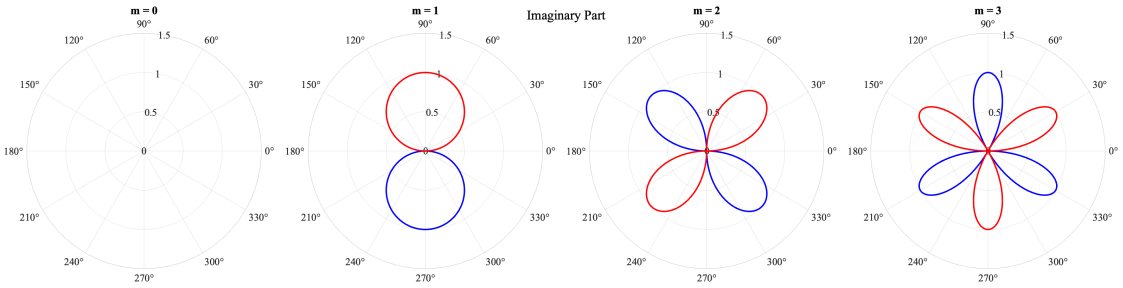



Figure 6: Imaginary part of circular harmonics

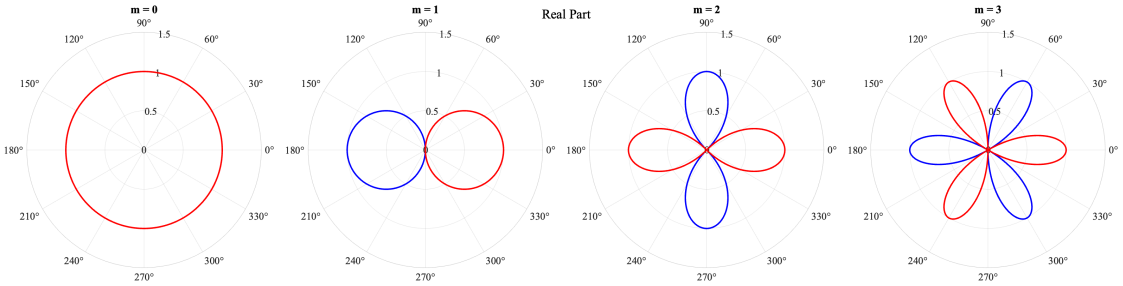



Figure 7: Real part of circular harmonics

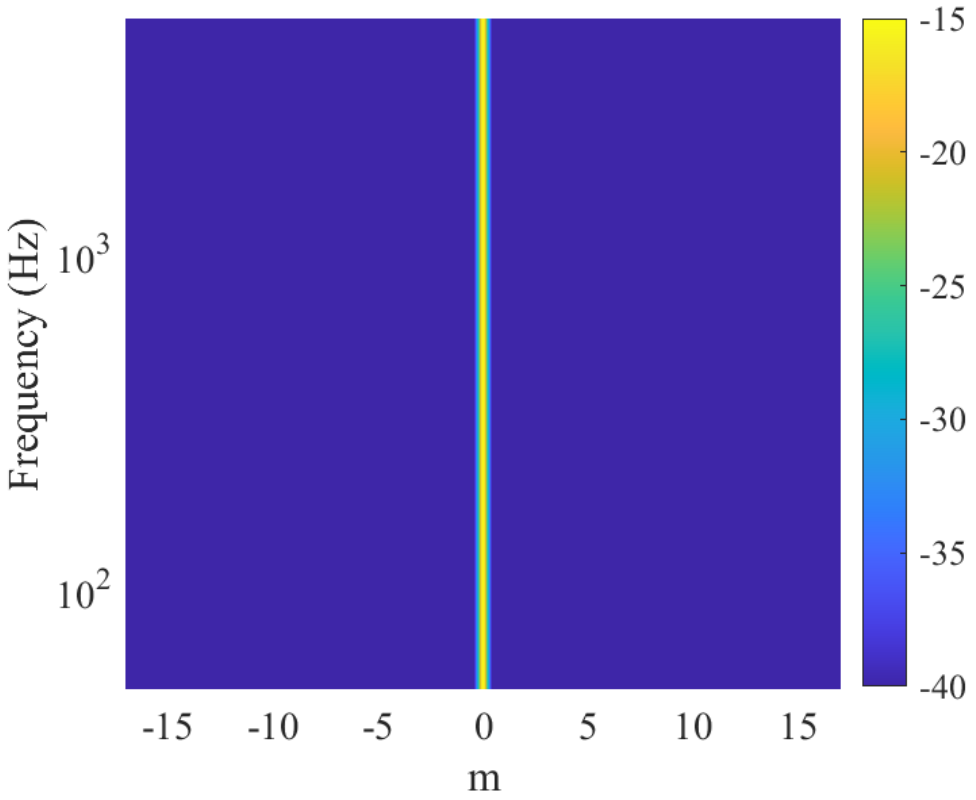



Figure 8: Circular harmonics coefficient of a point source

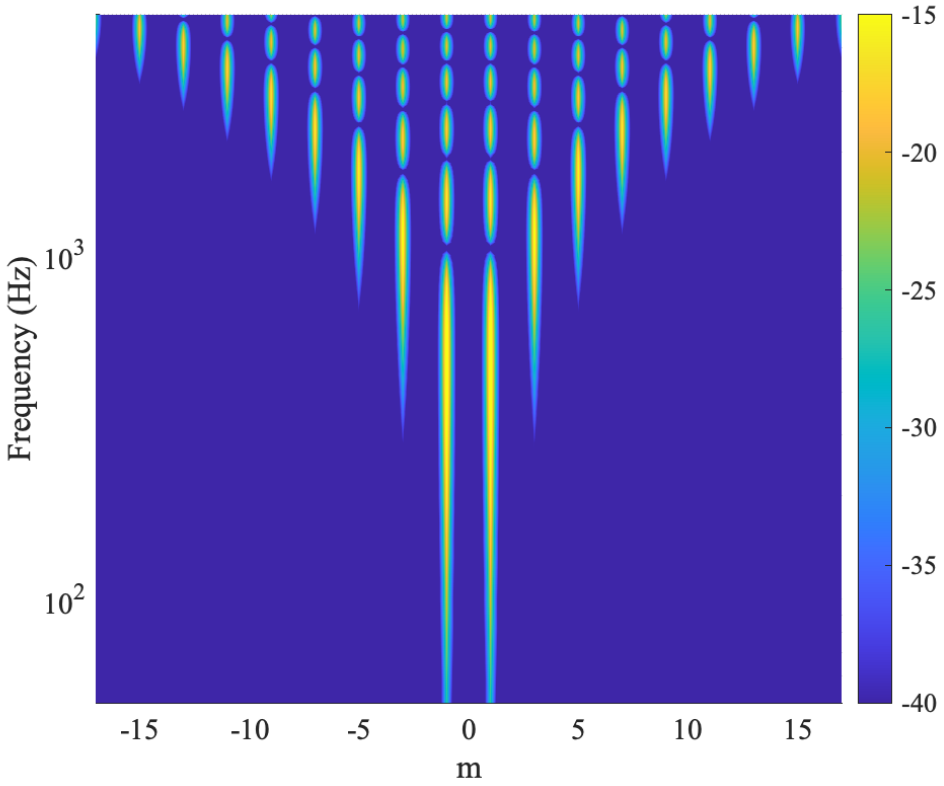



Figure 9: Circular harmonics coefficient of a dipole

As shown in figure 7, circular harmonics of $0^{\text{th}}$ order is a circle, and it's the same as the wavefront from a point source. Thus the circular harmonics coefficients of point source should only contain the $0^{\text{th}}$ order component. Figure 8 is the circular harmonic coefficients of a point source and one can conclude that it's compatible with expectation. Figure 9 is the circular harmonic coefficients of a dipole horizontally placed at origin point, where only circular harmonics of odd order contribute to total sound field. Dipole with two point sources vibrating out of phase doesn't stimulate symmetric constituent to their central line.

To achieve the circular harmonics expansion of a circular defined function, expansion coefficients have to be calculated with a certain method. According to equation 43, coefficients $\breve{P}_n(\omega)$ of circular harmonics can be calculated with complex amplitude $\tilde{P}(\omega, \varphi_0)$ along a circle, which can be exported from simulation results. Theoretically, pressure distribution among the whole domain can be manipulated with circular harmonics and their coefficients. The reconstructed field won't be totally equivalent to that got from simulations because it's a approximation using several points along a circular grids. But the accuracy is acceptable with proper chosen density of points. The algorithm to perform such calculations will be developed with MATLAB.

## 3.3 Artificial Field Manufacturing

During previous two sections, methods of performing simulations and expanding wavefront with circular harmonics are described and discussed in detail. But both of them ask for getting real data of sound pressure from simulations. In this section, two artificial manufacturing methods are introduced to realize the analogous functionality without any original pressure data. But the pressure data from simulations will be utilized as reference targets. Such two methods also employ the idea of circular harmonics expansions, because the manufactured field is calculated with coefficients that are acquired from new methods.

The first method is to create an arbitrary impulse response and apply adjustments on it to make it fulfilled. Since the circular harmonics coefficients for any specific order can be seen as a frequency response, the primary focus is thus to modify the arbitrary impulse response in such manner that the spectrum of it gets closer to that of the Schroeder diffusers. After which the impulse response becomes an artificial transfer function to be employed by listening signals. Besides, some random deviations such as Guassian noise will be added to explore more possibility to enhance its functionality further. Because almost any virtual listening experience comes from decorrelation between two ears, and extra deviations are possible to bring more decorrelation to binaural signals.

The second method is generating an plane wave with circular harmonics either. The circular harmonic coefficients of plane wave can be directly calculated with Bessel functions.

# 4 Results

In this thesis, methods to enlarge the perceptive size of sound source with diffusive geometries are explored. Several diffusers are designed and analysed to find dependence of their functionality on their designing characteristics. Each diffuser is seen as a LTI system that its scattering influence on sound field will be exported as transfer function in the form of frequency response of pressure. Proper audio samples are introduced to test the functionality of the diffusers, where headphone is required to play such audio signals as the virtual listening experience comes from the decorrelation of two ears. Moreover, algorithm to calculate circular harmonics coefficients and reconstruct sound field from coefficients is developed. This enables researcher to manipulate sound field with desired circular harmonic coefficients. And more importantly, diffusers are further simplified into abstract sources such as point sources which can build similar fields to that of diffusers with proper adjustments on their vibration characteristics.

## 4.1 Schroeder diffuser simulation

According to equation 3, the quadratic-residue sequence, which influences the depths of wells, are determined by a self-chosen prime number $N$ by designer. The dimensions of diffusers such as width and depths of wells are defined by calid working frequency band of the diffusers.

Here in this thesis, prime number $N$ is set to be 15. So the sequence is calculated:

$$s_n = 0, 1, 4, 9, 3, 12, 10, 10, 12, 3, ... \tag{49}$$

The higher limit for optimal diffusion is $f_{max} = 5\,\mathrm{kHz}$, which results in the width of wells to be

$$w = \frac{\lambda_{min}}{2} = \frac{1}{2} \cdot \frac{c}{f_{max}} = \frac{343\,\mathrm{m/s}}{2 \times 5000\,\mathrm{Hz}} = 3.43\,\mathrm{cm} \tag{50}$$

and the lower limit of frequency band is chosen as $f_0 = 100\,\mathrm{Hz}$ and determines the depths of wells together with sequence $s_n$ with following relationship

$$D_n = \frac{\lambda_0}{2N} s_n = \frac{1}{2 \times 13} \times \frac{343\,\mathrm{m/s}}{100\,\mathrm{Hz}} s_n = 13.2\,\mathrm{cm} \cdot s_n \tag{51}$$

With all the parameters above, Schroeder diffusers can be designed with different number of wells $n$. Since the diffusers are placed in 2-D polar coordinate with a 1-D sequence, they are denoted as $(n \times 1)$ diffusers in following descriptions.

Firstly, a $(8 \times 1)$ diffuser with 8 wells is designed to see how Schroeder diffuser works on

incident waves and influences the net sound field. A diagram of this diffuser is shown in figure 10, with necessary parameters labelled on.

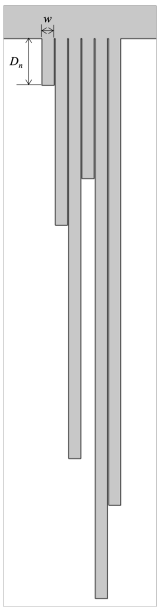



Figure 10: $(7 \times 1)$ Schroeder diffuser with N = 13

To simulate the listening experience when people are enclosed by the scattering sound field, a listener is located at $3.9\,\mathrm{m}$ away from the diffuser and $60°$ from normal direction, with location $(3.9\,\mathrm{m}, 30°)$ in polar coordinate. Two sampling points are placed in the simulations domain to simulate two ears and the distance between them is set as $18\,\mathrm{cm}$. The diagram of listening position is shown in fugure 11.

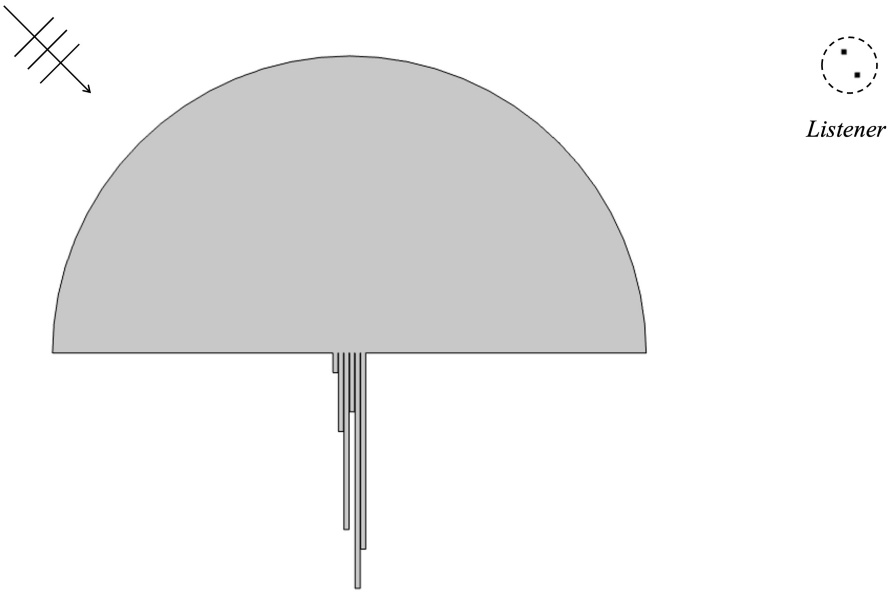
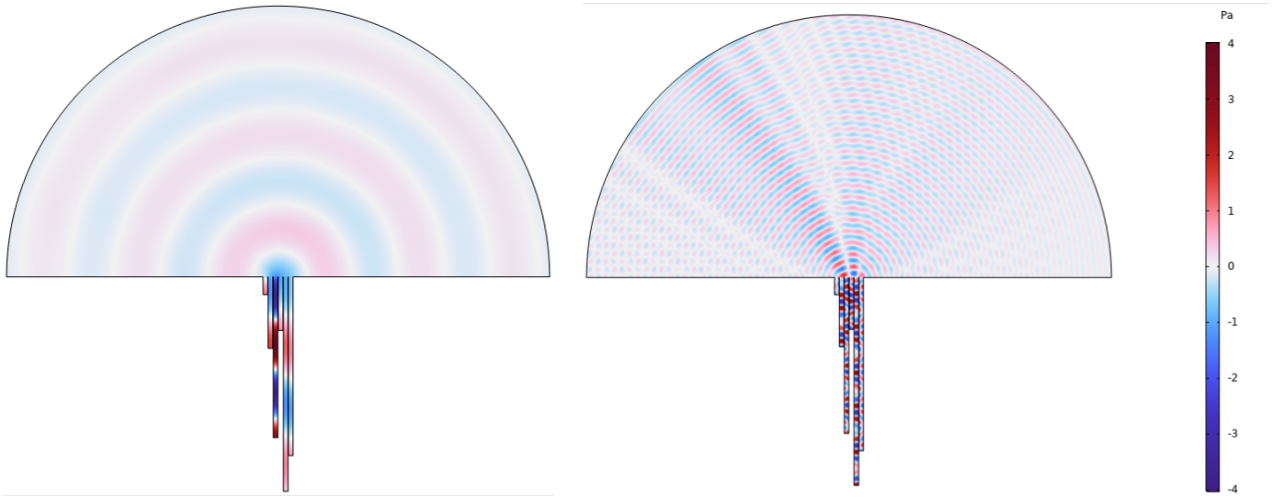



Figure 11: Simulation setup $(7 \times 1)$ Schroeder diffuser with N = 13 and listener positions

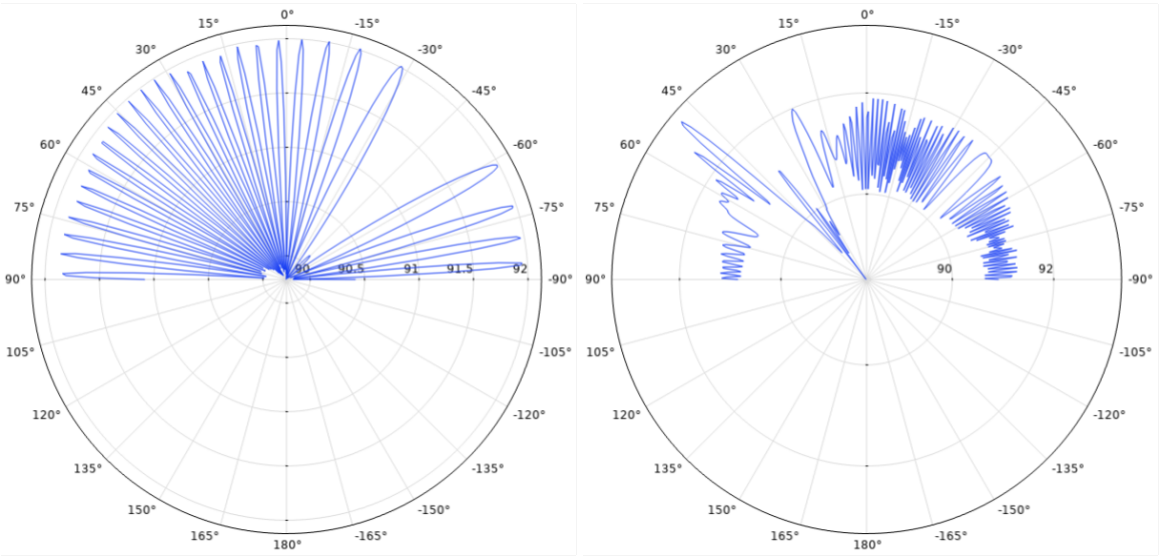
The half circle here is an air domain as the diffuser is assumed to be placed in air. The horizonal bottom boundary is set to be a bard boundary. A Perfect-matched-boundary (PMB) is placed at the upper arc boundary of air domain because COMSOL Multiphysics needs a finite size domain for calculation. The simulation is performed with a plane wave incident on the diffuser from direction with angular displacement $135°$. The simulation is performed from $100\,\text{Hz}$ to $5000\,\text{Hz}$ with an increment of $10\,\text{Hz}$.

The simulated scattering sound fields for both $500\,\text{Hz}$ and $4000\,\text{Hz}$ are shown below 12. It's easy to find the wavefront patterns of $500\,\text{Hz}$ are more omni-directional than that of $4000\,\text{Hz}$ as waves with longer wavelength are less likely to be influenced by tiny structures.




(a) $500\,\text{Hz}$ (b) $4000\,\text{Hz}$

Figure 12: Simulated scattering sound field of $(7 \times 1)$ diffuser




(a) $500\,\text{Hz}$ (b) $4000\,\text{Hz}$

Figure 13: Radiation Patterns of $(7 \times 1)$ diffuser along arc with radius of $10\,\text{m}$

Figure 13 shows the radiation patterns of the diffuser for both $500\,\mathrm{Hz}$ and $4000\,\mathrm{Hz}$ plotted in sound pressure level (SPL) along an arc with radius $r_0 = 10\,\mathrm{m}$. The reference direction (the $0^\circ$ in polar graph) is $90^\circ$ in simulation coordinate, thus the patterns are exactly for wavefront radiating between $0^\circ$ and $180^\circ$. And they are calculated in exterior field to prevent the influence of near field. From such patterns one can find the fluctuations for $500\,\mathrm{Hz}$ wave is more regular, while that for $4000\,\mathrm{Hz}$ is more violent. Such radiated waves bring more possibility to realize special listening experience as it comes from decorrelation between two ears. To create similar radiation characteristics for waves with longer wavelength, diffusers with larger size is necessary.

The complex pressure response at two listening points labelled in figure 11 are evaluated and are used as transfer functions of the entire scattering system. Two audio signal samples are employed here to test the functionality of this diffuser. The pressure responses at two ears are applied with inverse Fourier transformation so that the audio signals can be processed by the system through convolution. Two audio samples are **'As_long_ago_as_1860_48k.wav'** and **'Drums_short_48k.wav'** with sampling frequency of $48\,\mathrm{kHz}$. Since the highest frequency of interest is $5000\,\mathrm{Hz}$, audio samples are resampled into Nyquist frequency of $10\,\mathrm{kHz}$. The first audio sample is human voice of a man, which covers a broad frequency range and mostly in middle frequency. The second is a drum cut with dominant low frequency sound.

Processing results are quite different for the above 2 audio samples. The resampled audio are regarded as the original source to be processed and compared with new audio signals. The first sample acquires slightly perceptible feeling of being enlarged, as the perceived position of source moves slightly from center of head to both sides of two ears. The second sample doesn't show evident difference under the influence of this diffuser. As states before the second sample is mainly composed of low frequency drum, and considering relative slight effects of such diffuser working on long wavelength than on short wavelength, it's reasonable that human voice reflects more changes than drum.

A natural way to make the effects of diffusers more significant on low frequency sound is to deign a larger diffuser. Meanwhile, larger diffusers are expected to make the virtual size feeling even larger for first sample as the processed signals can be regarded as sound radiated from the diffusers. In order to study the effect of diffuser size on perceived size, many other diffusers with more wells are simulated.

Diffusers with $n = 13, 20, 26, 52, 104$ of wells are designed with the same parameters and simulated with the same procedure as that of $n = 7$. The diagram of such diffusers are shown in figure 14. Under the scattering effects of so many diffusers, the first audio sample experience a gradually increased feeling of source size when the diffusers have more wells. Since the processed signals are binaural, they sounds like two loudspeakers placed by both ears and the angle formed by source and listener becomes larger whem diffuser size increases. Even further, the signals processed by diffusers with $n = 52$ and $104$ even create slight feelings of reverberation. The second sample doesn't shown great differences until $n$ reaches $52$, but it proves that sound sources playing low frequency sound can also be enlarged as long as the dimensions of diffusers are larger enough to influence the travelling of low frequency waves.

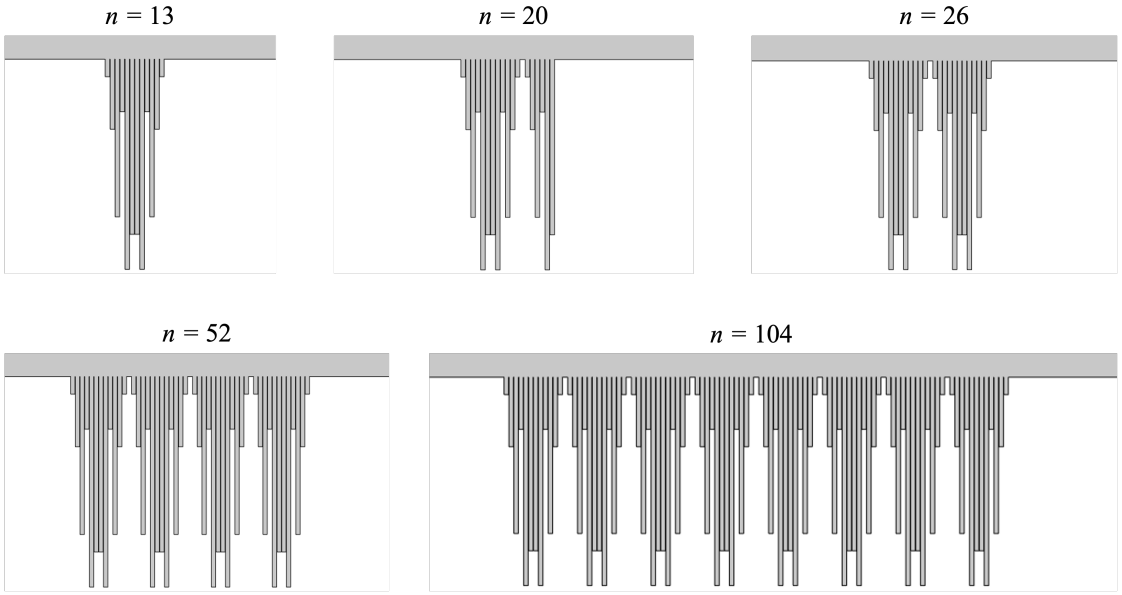
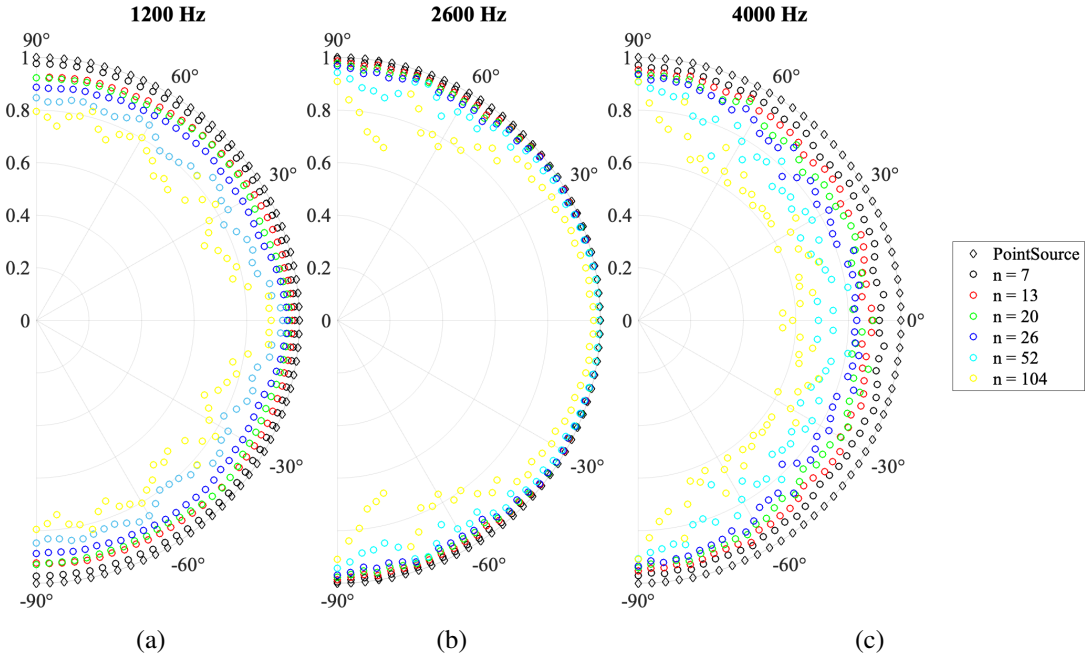



Figure 14: Diffusers with $n = 13, 20, 26, 52, 104$

### 4.1.1 Autocorrelation

Equation 5 shows the argument of autocorrelation function is incident angle. With pressure value achieved from simulations, the autocorrelation of all diffusers above are also calculated to see how diffused their scattering fields are. The intervals between incident waves are $3°$, so there are 61 incident directions in total.



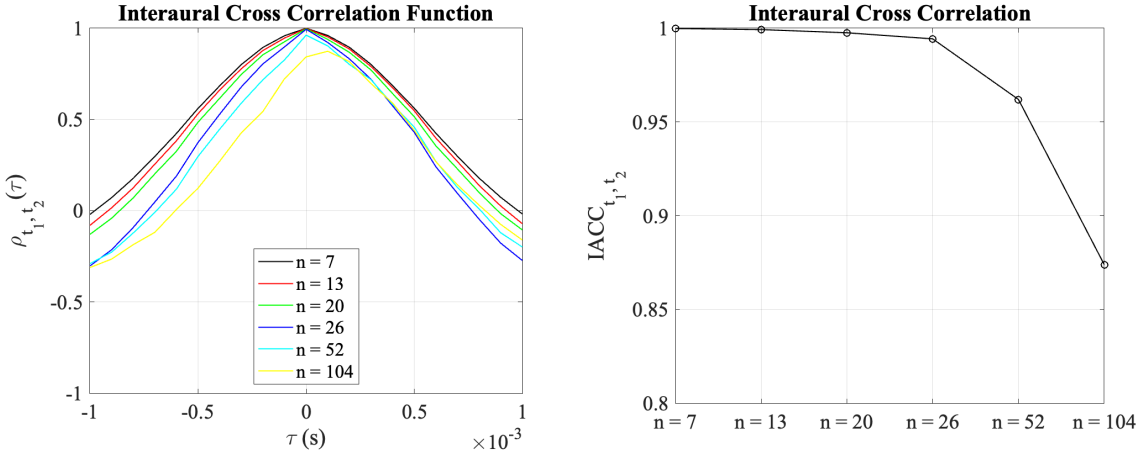

(a) (b) (c)

Figure 15: Autocorrelation of all simulated diffusers and point source under 1200 Hz, 2600 Hz and 4000 Hz.

The autocorrelation under $1200\,\mathrm{Hz}$, $2600\,\mathrm{Hz}$ and $4000\,\mathrm{Hz}$ are plotted in figure 15. The point source has autocorrlation $1$ as it radiates omnidirectionally. In general the diffuser becomes less diffused when $n$ and frequency increase. But $n = 104$ is an exception as it's more diffusive than other smaller diffusers in most cases. It's difficult to explain because it might come from singular scattering by the diffuser.

### 4.1.2 Inter-aural cross correlation

As discussed in section 2.3, the parameter IACC refers to the maximum value of inter-aural cross correlation function. Based on the first audio sample, the inter-aural cross correlation function and IACC values of all the simulated diffusers above are in figure 16. The maximum value of inter-aural cross correlation function is seated at $\tau = 0\,\mathrm{ms}$ as shown in 16a, which means there's no delay between the two channels of binaural signals. In 16b, IACC decreases relative to $n$ and is compatible with the listening results that perceived size of source increases.




(a) (b)

Figure 16: IACC function and IACC value under different $n$ based on $1^{\mathrm{st}}$ audio sample

Figure 16 is the IACC of full length of first audio signal with full band. To analyze the frequency dependence of IACC and see how greatly differently frequency components are affected by diffuser size, the IACC under 1/3 octave bands with central frequency from $125\,\mathrm{Hz}$ and $4000\,\mathrm{Hz}$ are plotted under different $n$ as presented in figure 17. 1/3 octave band filters of order $12$ are applied to signals to extract signals with specific frequency components so that IACC of 1/3 octave bands can be calculated with such extracted signals.

From the figure one can conclude that only diffuser with $n = 52, 104$ brings magnificent decrease in IACC for frequency from $315\,\mathrm{Hz}$ to $1000\,\mathrm{Hz}$. Diffusers with $n = 26, 52$ begin to bring apparent changes when frequency is larger than $1000\,\mathrm{Hz}$. The diffusers with $n = 26, 52$ brings extreme larger deduction of IACC when frequency goes beyond $1587\,\mathrm{Hz}$. In general, the functionality of larger diffuser as an enlarger is more evident.

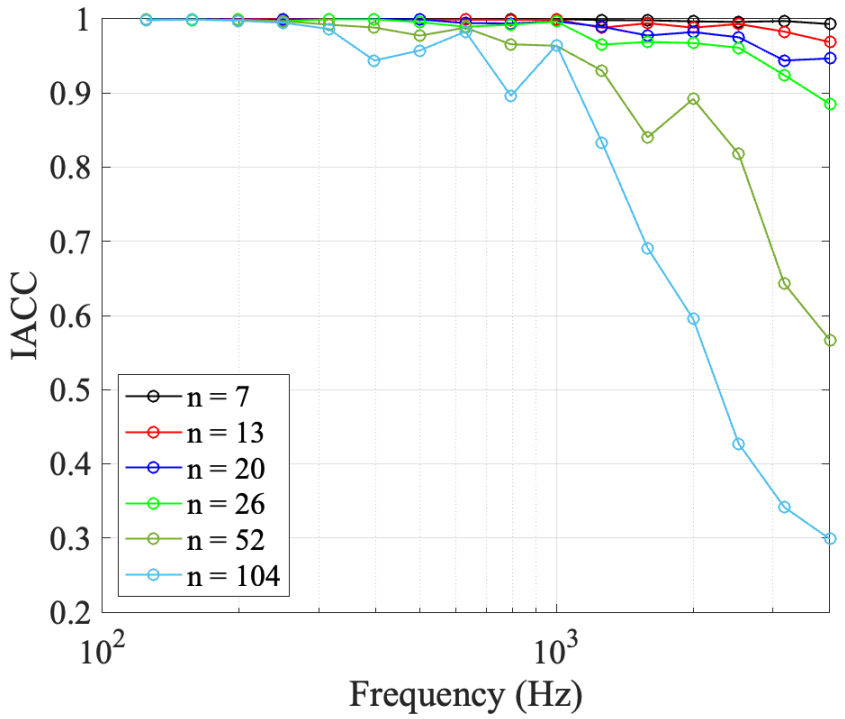



Figure 17: IACC in 1/3 octave band based on $1^{\text{st}}$ audio sample

## 4.2 Circular harmonic expansion

In this section the circular harmonic expansion of sound field is discussed. Previously, to achieve the pressure at specific positions in the sound field, simulation domain has to cover all the positions of interest, which will bring much computational pressure. With circular harmonic expansion, any points in far field can be calculated with one set of circular harmonics coefficients theoretically, which is the reconstruction of sound field. And this enables researcher to choose any positions in sound field they want as listening locations.

The calculation of circular harmonic coefficients asks for a set of sampling points along an circular arc with equal intervals, and the accuracy of reconstructed field depends on the density of points in the arc. So in practice, the reconstructed field usually covers only reasonable area around the arc.

In this thesis, sampling points are placed along a circle with radius of $4\,\mathrm{m}$ and $10°$ angular intervals, which means there will be $36$ points in total. And the largest order of circular harmonics is $17$. Figure 18 is a diagram of sampling points in simulation setup. Since the simulations are performed in physics domain above the hard boundary, only half of the points can be picked here. The other points are obtained through mirroring the points above boundary to beneath.

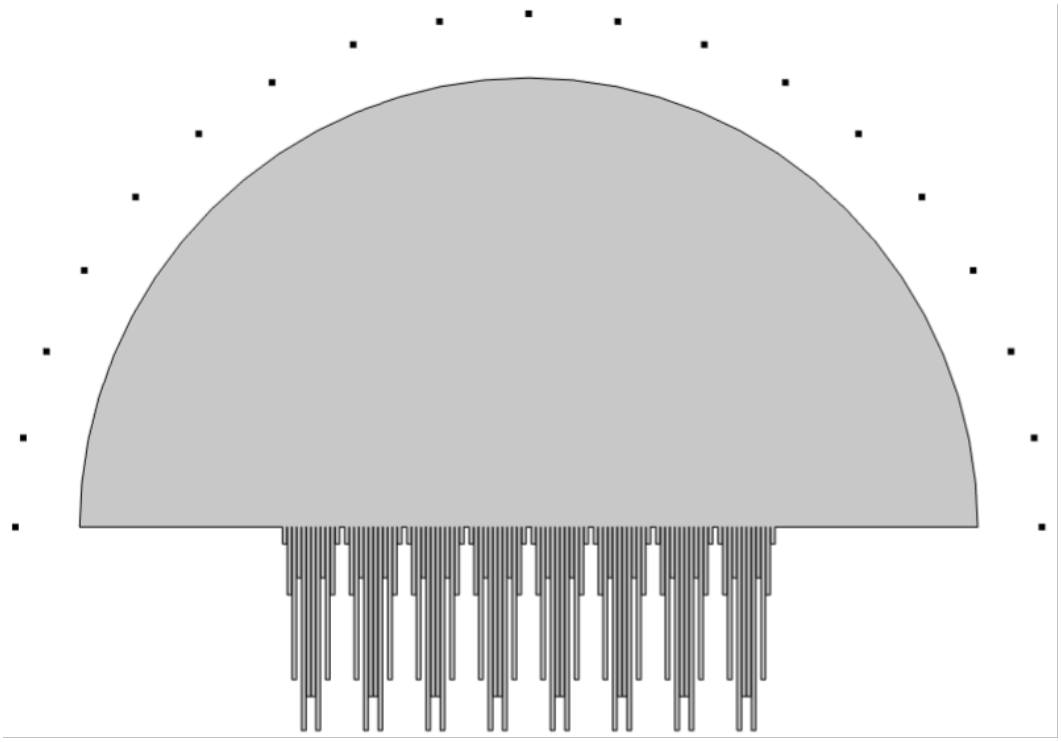


Figure 18: Sampling points along an arc for circular harmonic coefficient calculations

According to equation 43, the circular harmonic coefficients are calculated and plotted in figure 19 with 'shading' off in MATLAB. x-axis is the order of circular harmonics, while y-axis is the frequency distribution of coefficients. Large value in plot means more contribution of circular harmonics to total field. When the diffuser size grows up, the scattering field achieves larger constituents at high order circular harmonics, and low frequency components increase too. This trend is consistent with previous results that larger diffusers are able to influence the low frequency components in sould field to more extent.

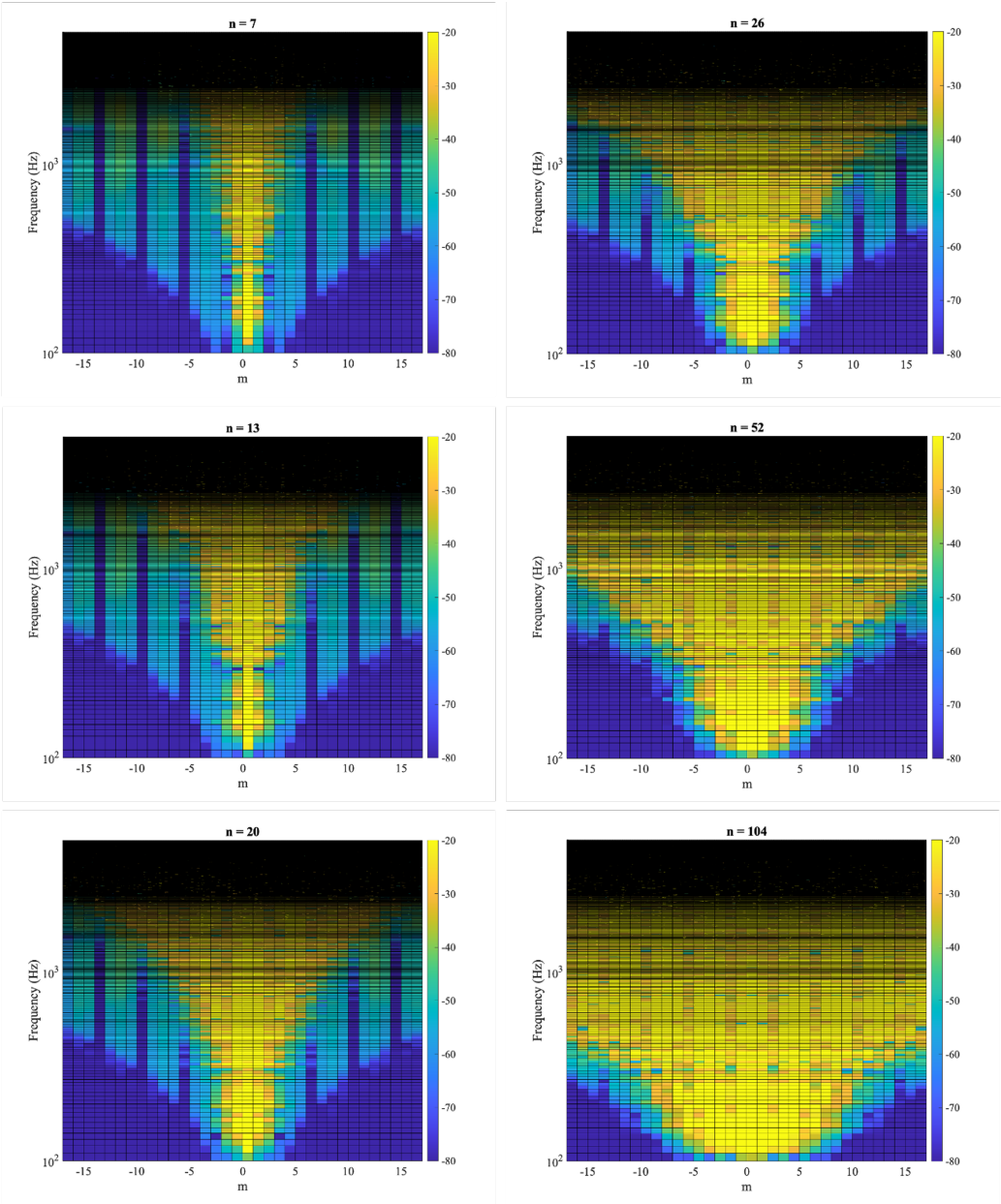



Figure 19: Sampling points along an arc for circular harmonic coefficient calculations

A effective way to prove the reliability of circular harmonic expansion is to compare the original sound field with the reconstructed one. In figure 20 - 25, the simulated sound pressure distribution over azimuth angle and frequency are compared with reconstructed field from circular harmonic expansion for all the $8$ simulated diffusers. The azimuth angle distributions in reconstruction are different from that in simulations. Considering convenience for future selection of listening signals, azimuth angle interval is chosen to let the distance between adjacent points equal to ear distance, which is $18\,\mathrm{cm}$ here. In far-field scenarios, the length of arc approximately equals to direct distance between two points.

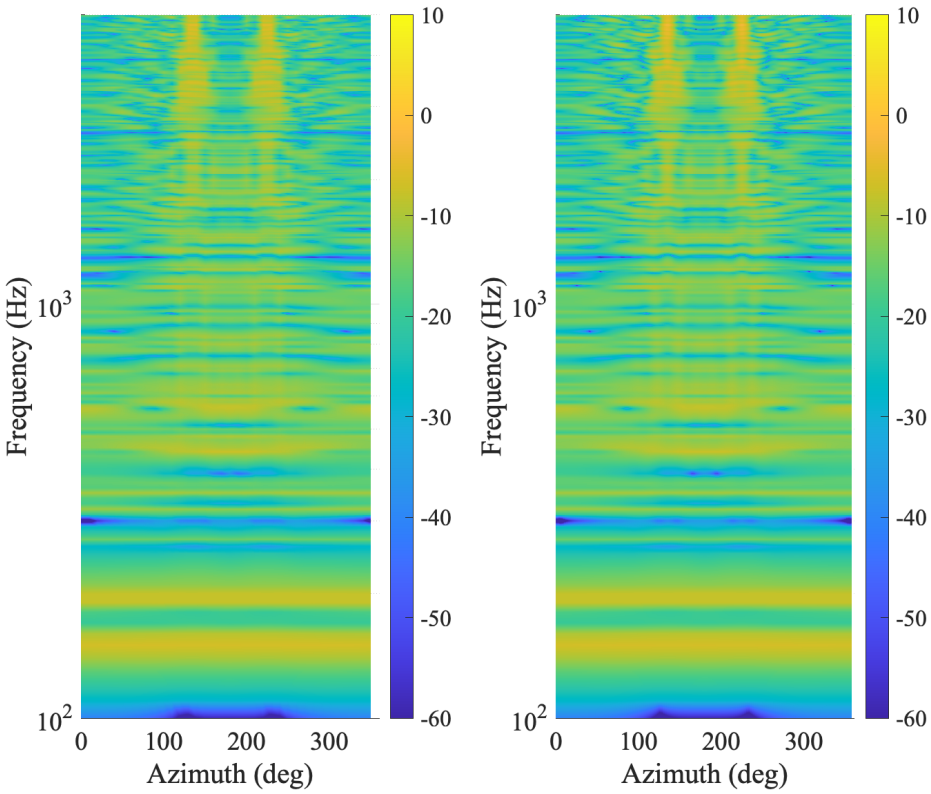



Figure 20: Comparison between simulated and reconstructed field for $n = 7$

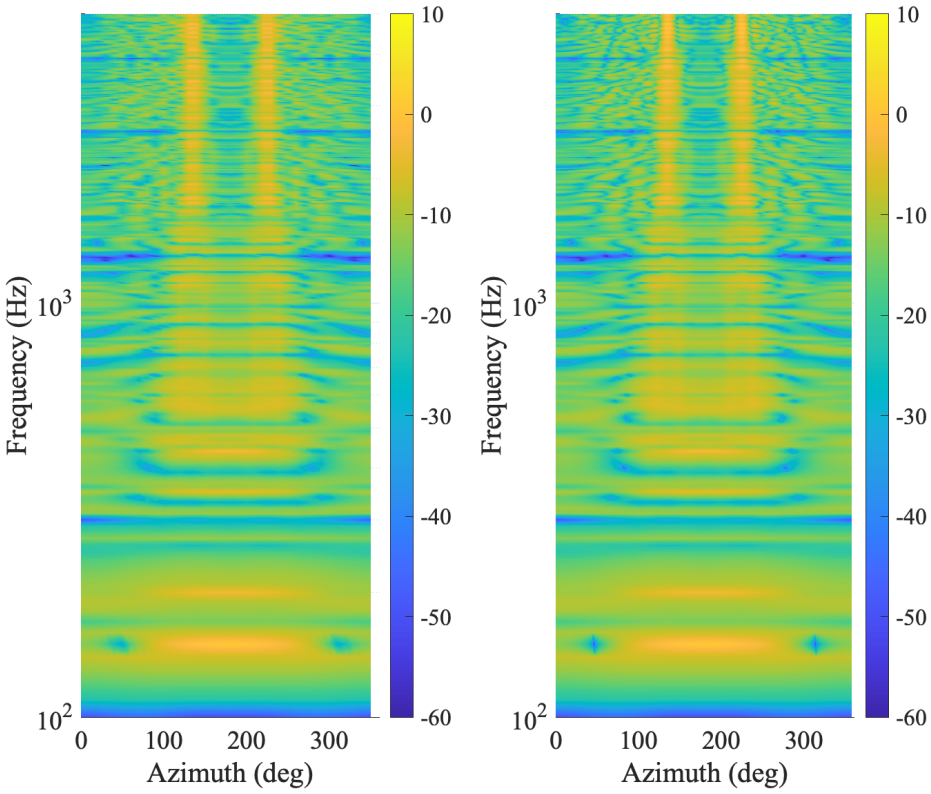



Figure 21: Comparison between simulated and reconstructed field for $n = 13$

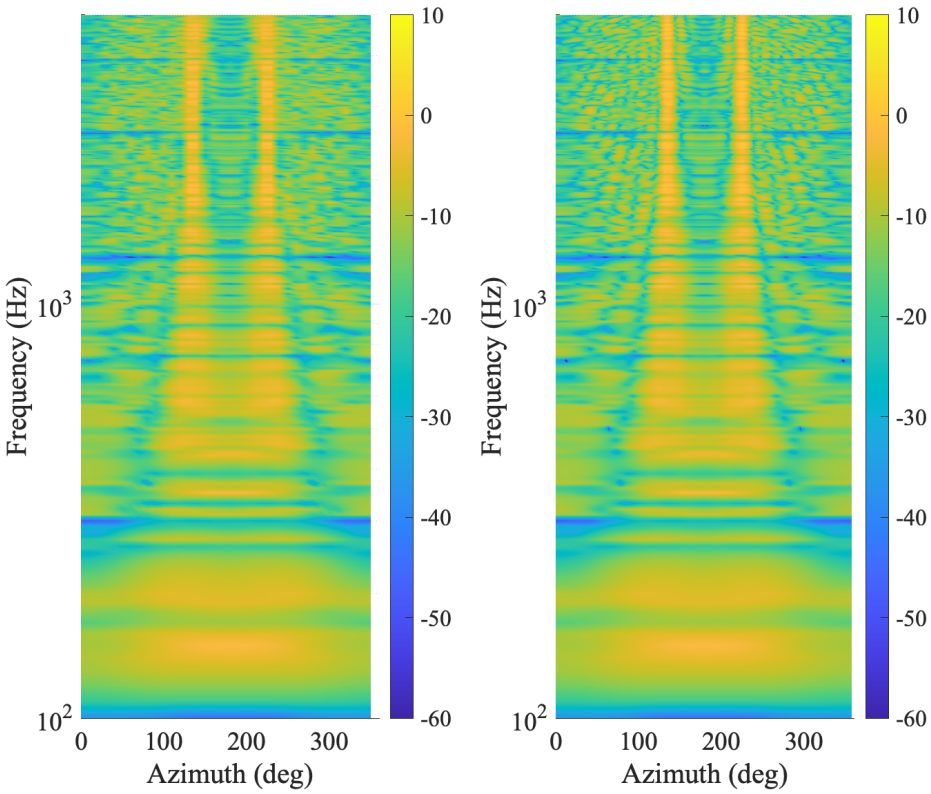



Figure 22: Comparison between simulated and reconstructed field for $n = 20$

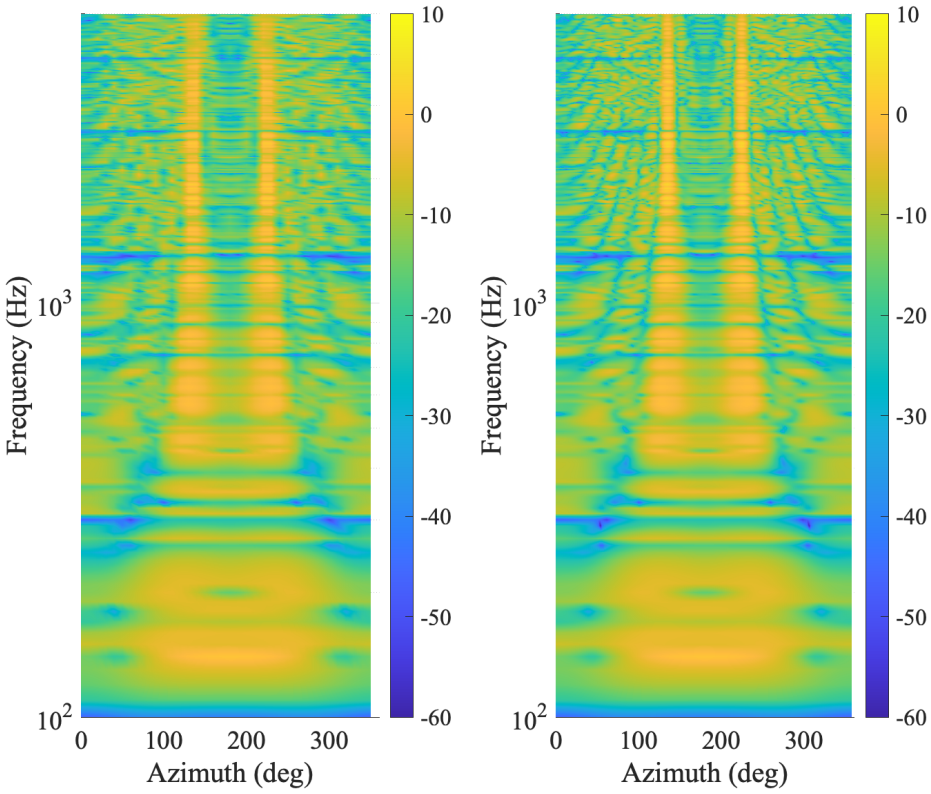



Figure 23: Comparison between simulated and reconstructed field for $n = 26$

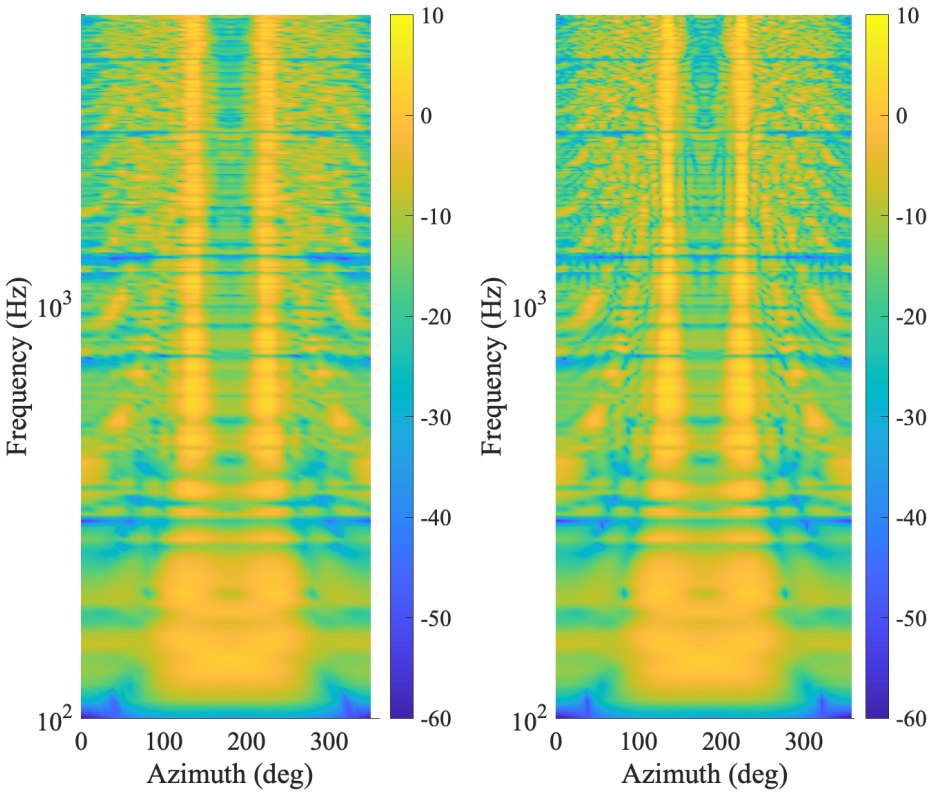



Figure 24: Comparison between simulated and reconstructed field for $n = 52$

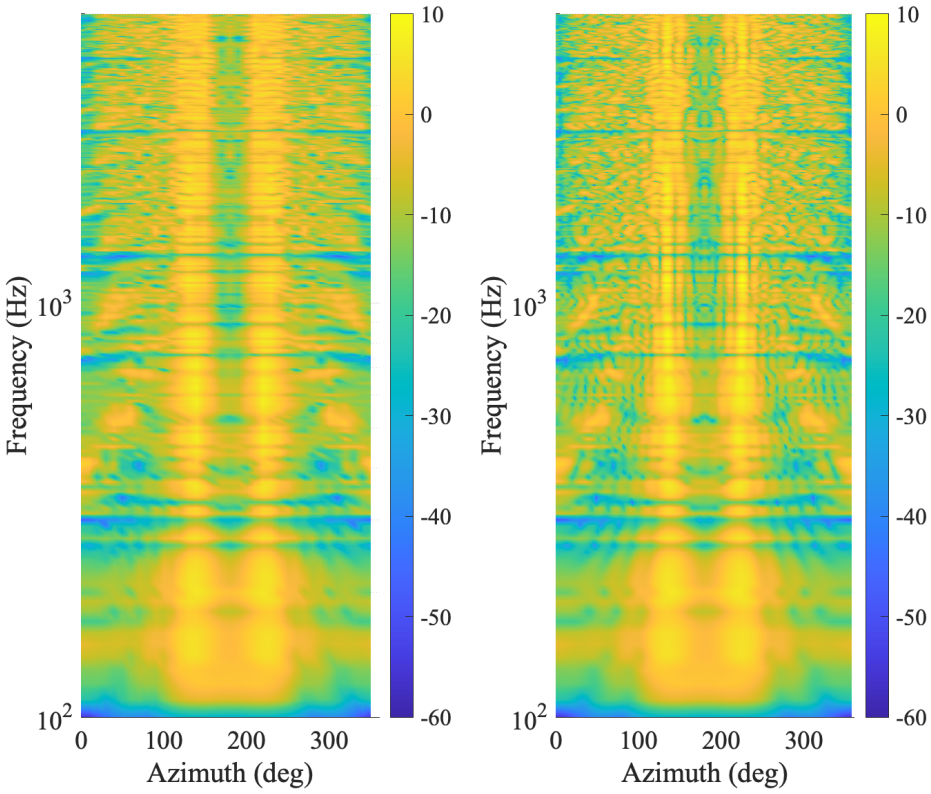



Figure 25: Comparison between simulated and reconstructed field for $n = 104$

The reconstructed field from circular harmonic expansions are almost the same as that of simulated fields under all simulated cases. Only some tiny difference can be detected with cautious observation. But it's totally acceptable as discussed before in theory section that circular harmonics expansion is an approximation of target function. From the figures one can also conclude that the fluctuations of low frequency waves become more and more violent when the diffusers becomes larger, where more decorrelation between two ears happens within low frequency range. Moreover, diffuser with more wells scatters more energy than that with less wells, which can be evidently seen from figures above that the sound pressure levels increase as $n$ increase under same incident conditions.

The inter-aural cross correlation of reproduced listening signals are plotted below in figure 26. Compared with figure 17, The IACC calculated in 1/3 octave band from reconstructed listening signals in figure 26b shows obvious deviations for the diffusers with $n = 52, 104$ with frequency larger than $1000\,\mathrm{Hz}$.

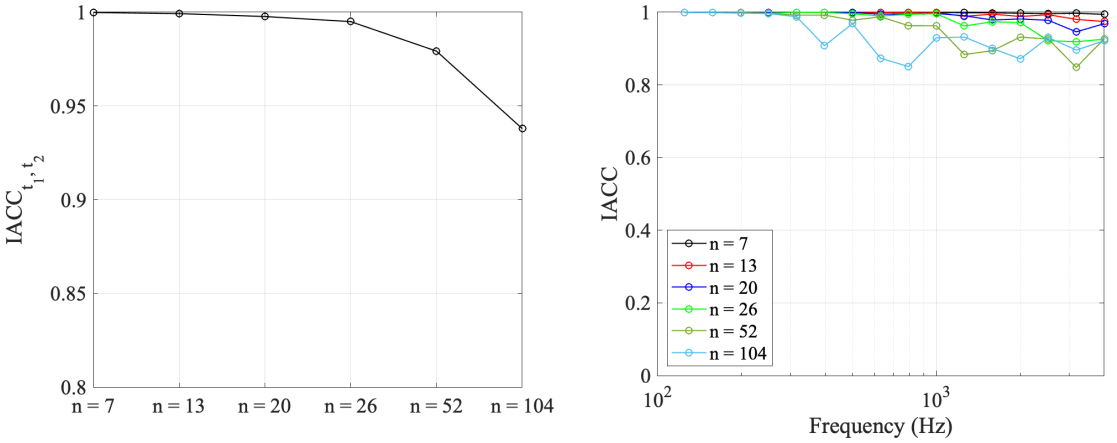



Figure 26: IACC function and IACC value under different $n$ based on reconstructed $1^{\mathrm{st}}$ audio sample

These deviations come from the accuracy of circular harmonics expansions. To acquire more accurate value with lower deviations, the sampling points in simulations should be more dense to achieve more accurate circular harmonic coefficients. In figure 27, the IACC in 1/3 octave band with new sampling grid of $2.5°$ is plotted. The new interval is almost the same as ear distance at $3.9\,\mathrm{m}$, which is $18\,\mathrm{cm}/3.9\,\mathrm{m} = 2.64°$. With more accurate sampling for circular harmonics expansion, the calculated IACC from reconstructed field now becomes close to the simulated results.

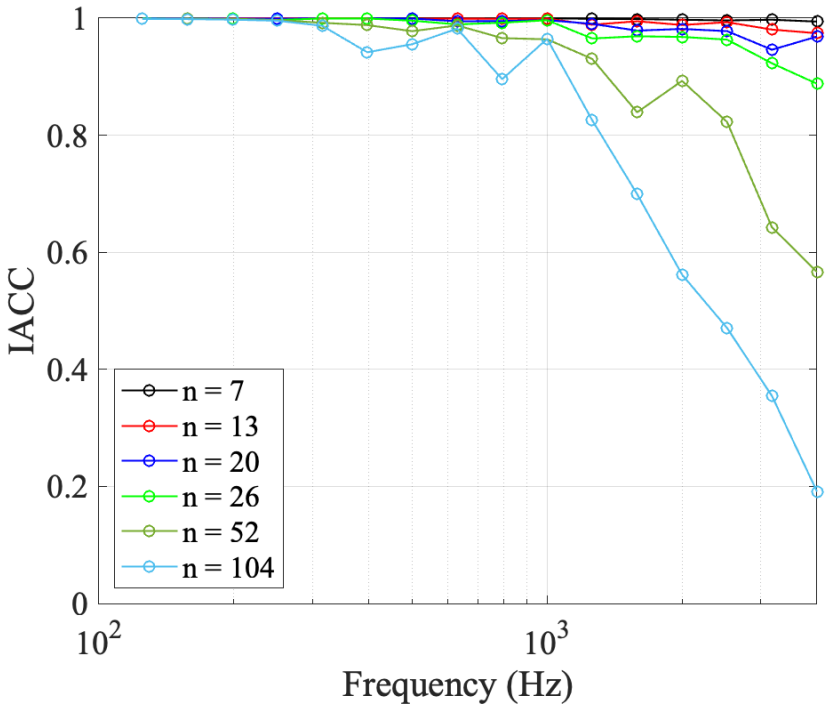



Figure 27: IACC value in 1/3 octave band under different $n$ based on reconstructed $1^{\text{st}}$ audio sample with sampling grid of $2.5°$

## 4.3 Artificial field

Grounded on the results from previous section, the diffuser with $n = 26$ is employed here as a reference target as it's the diffuser supplying best performance with acceptable accuracy under circular harmonics expansions using grid with $10°$ intervals. Two methods of manufacturing sound field artificially are discussed to reduce calculation consumption of simulations. The first method is manufacturing from impulse response, which can be stimulated by a point source. The second method is 'plane wave'. Since the radiated waves from a point source at very far-field can be regarded as plane waves, this method can be interpreted as a derivative method from point source.

### 4.3.1 Impulse response

This method might be deemed as an inverse operation of circular harmonic expansion. The circular harmonic coefficients are described over frequency and order number as one can see in figure 19. The coefficients under a single order would become a function over frequency, thus it's intuitive to consider it as a frequency response, which is equivalent to an impulse response in time domain. This method starts from creating an ideal impulse response in time domain, whose spectrum is a flat line in frequency domain. The created impulse signal is modified so that its spectrum becomes as analogous to that of circular harmonics coefficients

as possible. The created impulse is stimulated at the time when waves travel from origin to listening location. Here the traveling time is $(3.9\,\mathrm{m})/(343\,\mathrm{m/s}) = 11.4\,\mathrm{ms}$.

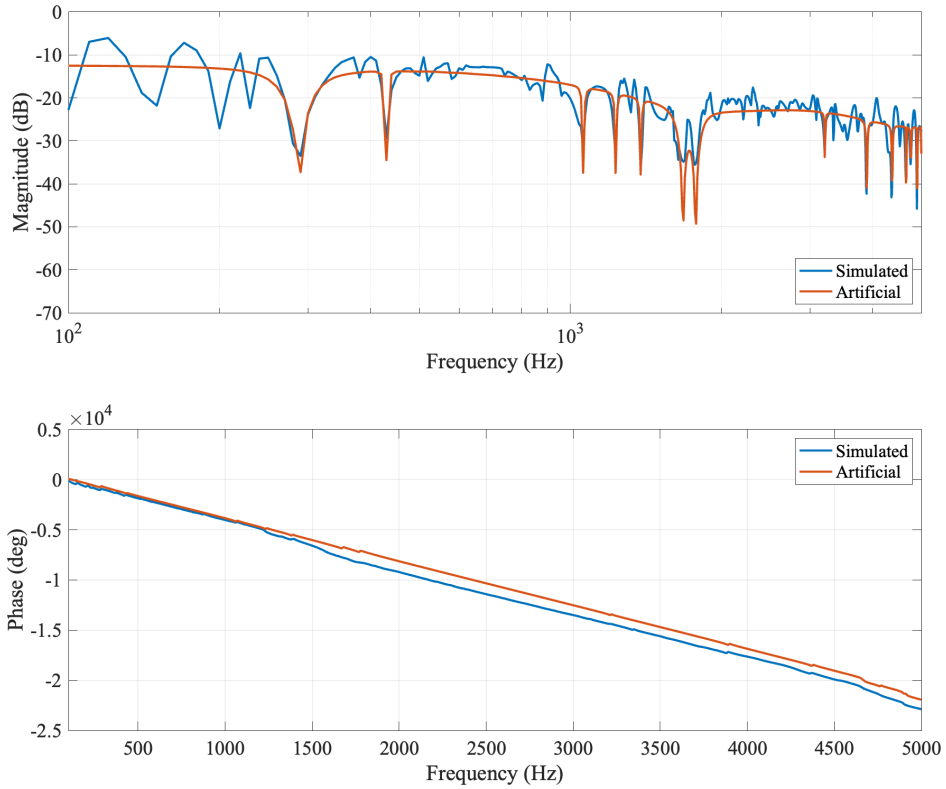



Figure 28: Frequency response of simulated circular harmonics coefficients and artificially manufactured impulse response of $0^{\mathrm{th}}$ order

Figure 28 presents an instance for the above mentioned operations using $0^{\mathrm{th}}$ order circular harmonic coefficients. The original flat line of an impulse response is modified by several filters to fit the simulated response of coefficients. The filters are design by MATLAB and refer to 'butterworth' and 'FIR' filters. Furthermore, the simulated response contains many fluctuations which are difficult to be obtained through filtering, thus some complex Gaussian noise is added to the created response to make it fit target curve better, as shown in following figure 29. The comparison between simulated impulse response of circular harmonic coefficients and artificially generated impulse with filtering and noise are shown in figure 30. The created impulse response is in proper shape and has similar oscillating characteristics with simulated impulse. The amplitude difference comes from accumulative deviations over full band, which is reasonable as impossibility to reproduce a totally identical frequency response with filtering and noise.

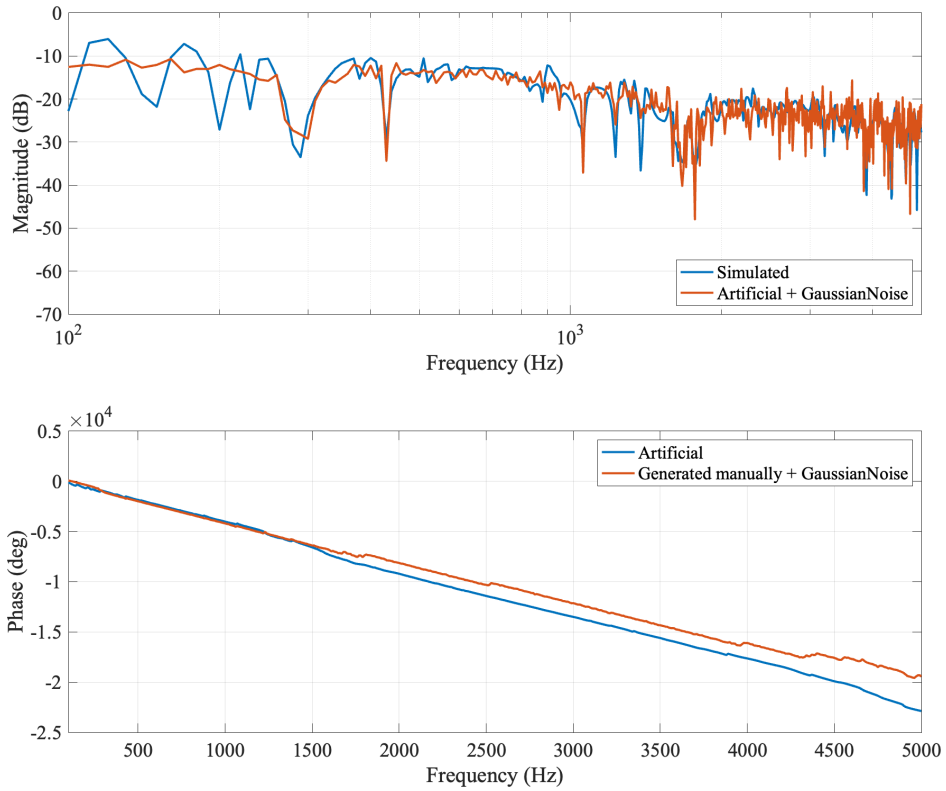



Figure 29: Frequency response of simulated circular harmonics coefficients and artificially manufactured impulse response of $0^{\text{th}}$ order with Gaussian nosie

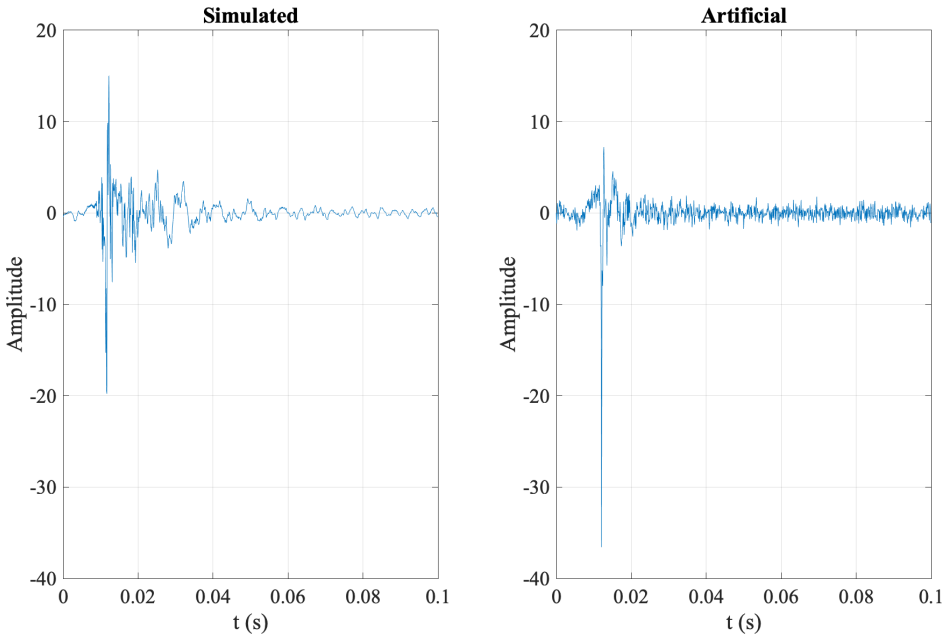



Figure 30: Impulse response of simulated circular harmonic coefficients and artificially created

Afterward, same procedures of generation are applied to higher orders of circular harmonics to create more complete coefficients set. From figure 19 one can find the dominant constituent of circular harmonics are that in low order and circular harmonics in such order contributes to full band of interest. As a result, the coefficients from $0^{\text{th}}$ to $5^{\text{th}}$ are generated and visualized in figure 31.

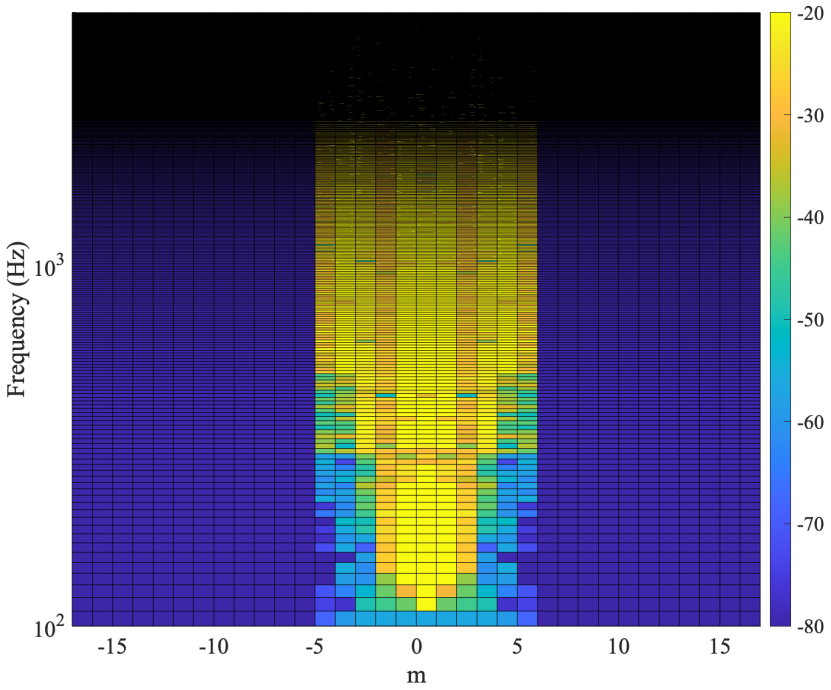



Figure 31: Artificially created circular harmonic coefficients from impulse responses

With newly created coefficients, an artificial sound field can also be generated and the pressure response at same listening positions as before are easily calculated. The audio samples are processed again and compared with simulated ones. The artificially created field still works as an enlarger. Since the high order components are skipped here in this method, losses in high frequency components can be obviously detected, which makes the scattered signals not as bright as that of $n = 26$.

### 4.3.2 Plane wave

A second method of plan wave generation is discussed in this section. It's quite similar to impulse response method, except the sound field is currently a plane wave field.

Circular harmonic coefficients of plane wave can be calculated directly with Bessel function according to equation 37 as MATLAB has a built-in function for Bessel function. However, the listening position in this case has to be small because the arguments of first kind Bessel functions should be small to make the function applicable.

The incident direction of plane wave is set as $-\pi/2$, which means propagating perpendicularly from bottom to top. The listening position now is $(1\,\text{m}, 30°)$ in polar coordinate, so the

distance from source to listener's head center becomes $0.5\,\mathrm{m}$. The travelling time for waves to reach the listener is $(0.5\,\mathrm{m})/(343\,\mathrm{m/s}) = 1.46\,\mathrm{ms}$.

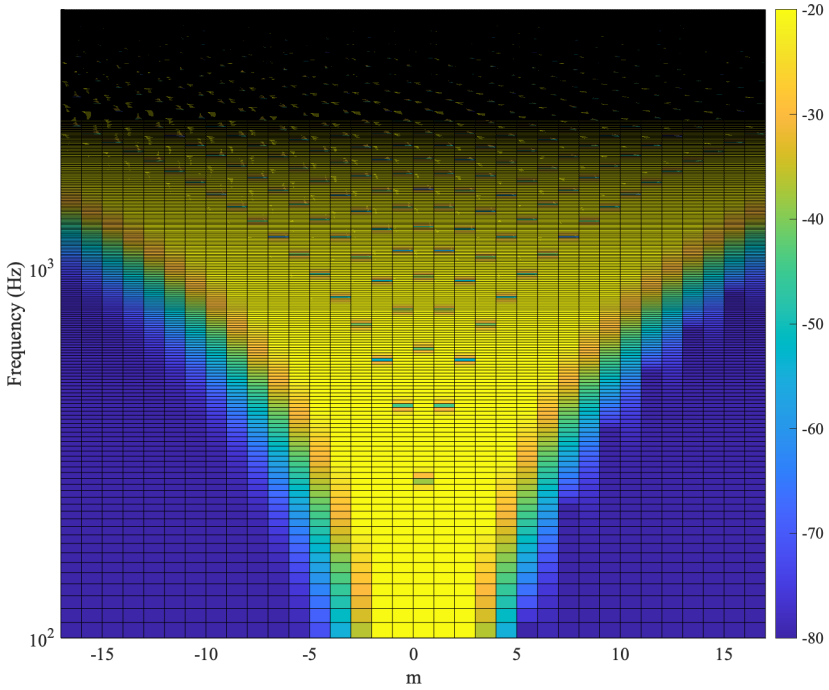



Figure 32: Circular harmonic coefficients of artificially created plan wave

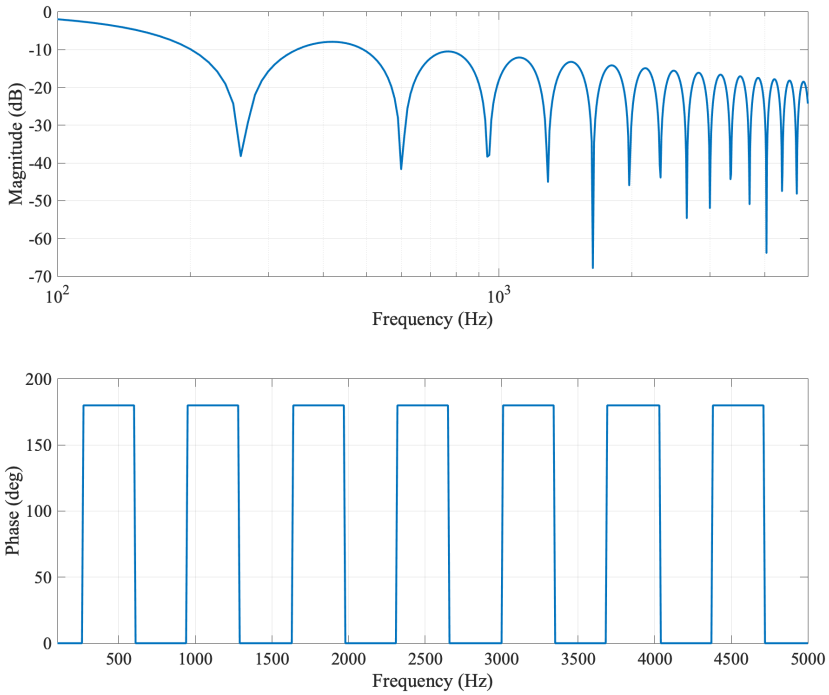



Figure 33: Frequency response of $0^{\mathrm{th}}$ circular harmonic coefficients of artificial plan wave.

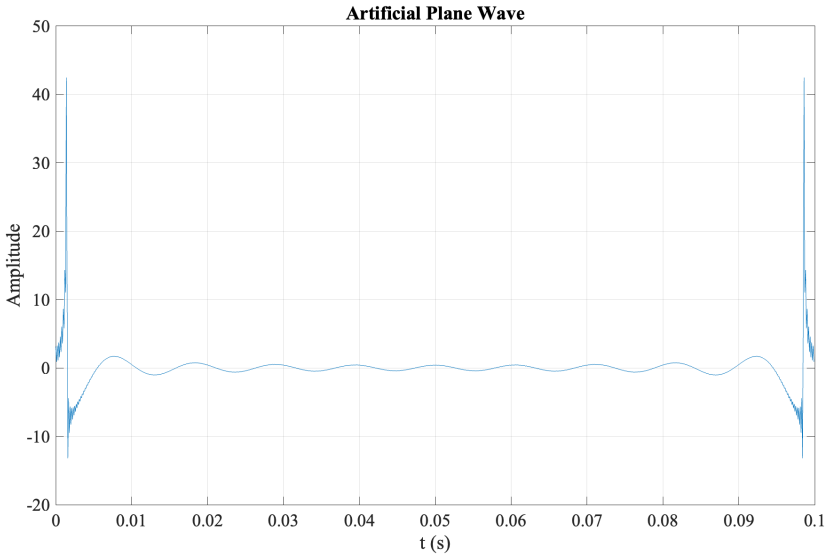



Figure 34: Impulse response of $0^{\text{th}}$ circular harmonic coefficients of artificial plan wave.

The circular harmonics coefficients of this plane wave is plotted in figure 32. The corresponding frequency response and impulse response of $0^{\text{th}}$ order coefficients is in figure 33 and 34. With circular harmonic coefficient, the signals at listener's positions are available. From figure 34 one can see a first impulse at the beginning and another impulse at the end. The first one is the wavefront traveling from the source beneath, while the second is called spatial aliasing caused by high frequency waves. It's inevitable with discrete layout of secondary sources. Subsequently, the created plane waves are also added with extra Gaussian noise to bring more decorrelation between two ears and its coefficients are shown in figure 35. The audio samples at listening positions are tested and the generated plane wave is functional as an enlarger. But listener feels apparent offset of source location as the wavefront arrives at two ears at different time slots. But it's inevitable as the decorrelation of binaural signals is needed. However, this shortcoming can be reduced to some extent through adding random noise to signals.

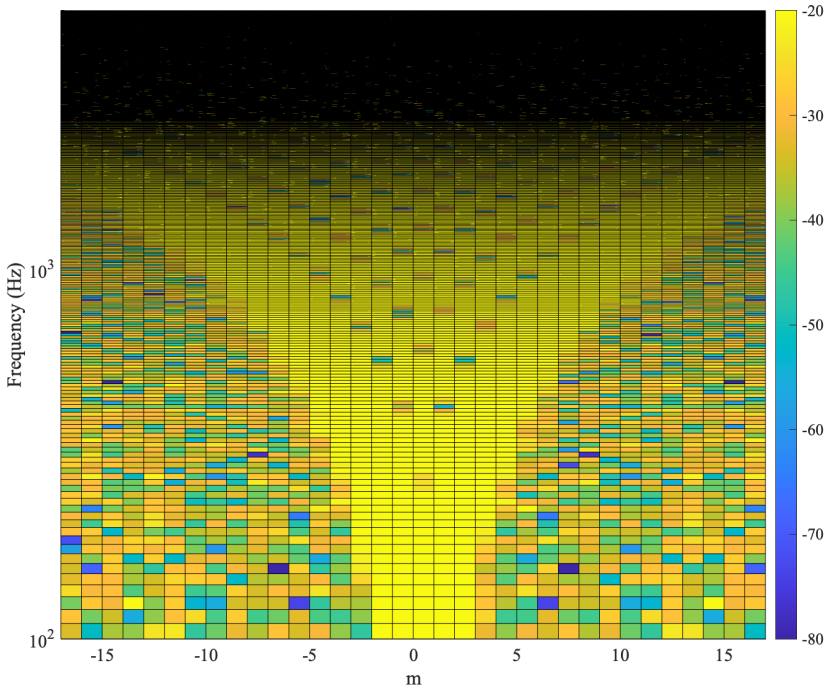



Figure 35: Circular harmonic coefficients of artificially created plan wave with Gaussian Noise

# 5 Discussion

In this thesis, several Schroeder diffusers of different width with same $N$ are simulated and see how they influence the perceived size of sound source. The circular harmonics expansion reveals the increase of high order contribution when diffuser becomes larger. The next step to further apply such properties would be developing a tool which can impose fine adjustments on the source that can gradually tune the perceived size of it.

The binaural signal here only considers the situation of being played in headphones, which is easier to estimate the perceived size with IACC. However, things would be more complicated if the signals were played through loudspeaker pairs, where other factors would influence the prediction of IACC such as cross-talk between left loudspeaker and right ear and vice versa. To realize the functionality of source enlarger with more kinds of hardware set-ups, research with more complete theoretical system and more rigorous verification are necessary.

Circular harmonics expansion is an universal method to decompose a sound field of interest and explore its constituent. In this thesis, sound field formed by the Schroeder diffuser is studied to enlarge the perceived size of the sound source, where circular harmonics contribute greatly to the implementation of later operations. For any other sound field with different characteristics, circular harmonics would also be a powerful tool. One concern is the accuracy of reconstructed sound field, where the sampling grid should be chosen properly so that the calculated circular harmonics coefficients can supply precise results over required frequency range with acceptable deviations.

# 6 Conclusion

Employing Schroeder diffuser as source size enlarger is effective and reliable. The functionality of it are proved with both simulations and artificial methods, creating appreciable virtual listening experience of an enlarged sound sources. Simulation is a reliable method to predict the scattering sound field, but brings heavy computational burden. Artificial methods which are founded on circular harmonic expansion are more economical routes to realise similar effects.

Circular harmonic expansion refers to theory of sound field synthesis, where the reconstruction of sound field count on secondary source array, which might cause unexpected effects such as spatial aliasing due to discrete source layout. Manufacturing artificial field from impulse response shows more solidity on size enlargement, but asks for much work on designing filters. Plane wave method is easier to implement, while the listeners has to bear the slightly offset of source locations.

# References


[1] Käsbach, J., Marschall, M., Epp, B., and Dau, T. The relation between perceived apparent source width and interaural cross-correlation in sound reproduction spaces with low reverberation. In *39th German Annual Conference on Acoustics*. 2013.

[2] Gonzalez, R., Lokki, T., and Politis, A. Functional Modelling of Scattering for Diffusive Geometries. In *Audio Engineering Society Conference: 2022 AES International Conference on Audio for Virtual and Augmented Reality. Audio Engineering Society.* 2022.

[3] Schroeder, M. R. "Binaural dissimilarity and op- timum ceilings for concert halls: More lateral sound diffusion," *J. Acoust. Soc. Am, 65(4), pp. 958–963*. 1979.

[4] Cox, T., and d'Antonio, P. Acoustic absorbers and diffusers: theory, design and application. *CRC press*. 2016.

[5] Gauss, C. F. Disquisitiones arithmeticae. *Yale University Press*. 1966.

[6] Kleiner, M. Acoustics and audio technology. *J. Ross Publishing*. 2011.

[7] Theodoridis, S., and Chellappa, R. Academic press library in signal processing. Vol 4: Image, video processing and analysis, hardware, audio, acoustic and speech processing. *Academic Press*. 2013.

[8] ISO 3382:1997, International Organization for Standardization, Geneva, Switzerland. *Acoustics - measurement of the reverberation time of rooms with reference to other acoustical parameters*. 1997.

[9] Zienkiewicz, O. C., Taylor, R. L., and Zhu, J. Z. The finite element method: its basis and fundamentals. *Elsevier*. 2005.