Doctoral Dissertation

Sound Quality Improvement and Directivity Control for Parametric Array Loudspeaker

March 2022

Doctoral Program in Advanced Information Science and Engineering

Graduate School of Information Science and Engineering

Ritsumeikan University

GENG Yuting

Doctoral Dissertation Reviewed
by Ritsumeikan University

Sound Quality Improvement and Directivity Control for Parametric Array Loudspeaker

(パラメトリックスピーカにおける 音質改善および指向性制御)

March 2022
2022年3月

Doctoral Program in Advanced Information Science and Engineering

Graduate School of Information Science and Engineering
Ritsumeikan University

立命館大学大学情報理工学研究科
情報理工学専攻博士課程後期課程

GENG Yuting
コウ　イクテイ

Supervisor : Professor NISHIURA Takanobu
研究指導教員：西浦 敬信 教授

Dissertation submitted to Graduate School of Information Science and Engineering,
Ritsumeikan University
in partial fulfillment of the requirements for the degree of
DOCTOR of ENGINEERING

Thesis Committee: Takanobu Nishiura
Yoichi Yamashita
Tadahiro Taniguchi
Masato Nakayama

Abstract

With the development of audio technology, the demand for directional sound grows. Parametric array loudspeaker (PAL) utilizes ultrasound to realize a sharp directivity and its practical application for generating directional sound is promising. This thesis focused on the sound quality and directivity control of PAL aiming for practical use.

Ultrasounds are emitted intensely from PAL and the difference frequency component self-demodulates due to the nonlinear interactions in the air. Therefore, the sound reproduced by PAL shows the difficulty in low-frequency reproduction, and it holds more harmonic distortions compared with the sound reproduced by conventional electro-dynamic loudspeakers. It is known that the demodulated sound pressure increases as the emitted power. In the case that the scale of the device is unlimited, enlarging the number of ultrasonic transducers can be a solution to the poor low-frequency reproduction. This thesis gives a method to realize a flat frequency response for PAL by enhancing the low-frequency with the multi-way structure. In this method, the proposed multi-way PAL is composed of Woofer PAL and Tweeter PAL. The scale of these parts is balanced to reach a flatter frequency response. In addition, the hardware arrangement is studied under the consideration of directivity. Moreover, suitable modulation methods are chosen for each band. The harmonic distortion is utilized to enhance low-frequency. Based on the proposed method, a prototype multi-way PAL is built for validation. The experimental results show that the proposed method can achieve a flatter frequency response in a wide band.

In some cases of practical application, the scale of the device may be limited. It is a considerable way to realize a flat frequency response by equalization. Due to the difficulty for PAL in low-frequency reproduction, the total demodulated sound pressure is predicted to decrease after the equalization. Under this condition, it is necessary to over-boost PAL, which means increasing the driving voltage for higher sound pressure. However, it brings an issue that the

nonlinear distortions increase as the driving voltage. Therefore, this thesis gives a sound quality improvement method that involves a consideration of the driving voltage. This method is specialized for harmonic sounds like speech and music. The difference in the harmonic structure between the demodulated sound and the original sound is investigated. The proposed method modifies the spectrum in advance to counteract the nonlinearity so that the original spectrum can be restored in the demodulated sound. Further, the proposed method provides optimized parameters according to the determined driving voltage. The effectiveness has been demonstrated by the experimental results, which show that the proposed method can achieve a higher fidelity of spectral characteristics when reproducing harmonic sounds.

When downscaling the device of PAL, there is another latent problem that the acoustic beam generated by PAL may diffuse. In addition, a sharper beam than ordinary PAL is desired in some cases. Therefore, this thesis also gives a directivity control to form a sharper beam. It is known that the principle of PAL is the parametric effect. If the parametric effect can be weakened spatially, the acoustic beam generated by PAL can be controlled. Based on this assumption, the proposed method utilizes phase inversion to achieve a sharper beamwidth. The characteristics of the modulated wave are considered when designing the phase inversion strategy. The measured results demonstrated that the proposed method can achieve a narrower acoustic beam.

アブストラクト

音響技術の発展と伴い、指向性音源の需要は高まりつつある。パラメトリックスピーカ (Parametric array loudspeaker: PAL) は超音波を利用し、鋭い指向性を実現できる。そのため、近年、PAL を指向性音源としての応用が期待される。この論文は PAL の実用化に向けて、音質の改善と指向性の制御に着目している。

PAL から超音波を大音圧で放射すると、空気の非線形性により、周波数の差の成分が自己復調する。この原理により、PAL は低周波音の再生が困難となる。その上、非線形現象により、従来の動電型スピーカの再生音より、PAL の再生音に高調波歪みが多く含まれる。PAL の復調音は放射した超音波のパワーと正の相関があるため、機材の規模に制限がない場合、PAL を構成する超音波素子の数を増やすことで、低周波音の再生困難という問題を解決できる。この論文では、マルチウェイ構造を用いて、低周波音を強調することで平坦な周波数特性を実現する手法を挙げる。この手法で提案されたマルチウェイ PAL は Woofer PAL と Tweeter PAL で構成される。より平坦な周波数特性を実現するために、両部分の規模と配置及び変調方式に関する検討を行なった。特に、低周波音を強調するために、高調波歪みを有効活用した。提案手法に基づき、試作機を開発した。評価実験の結果、提案手法を用いて、広帯域においてより平坦な周波数特性の実現を確認した。

PAL の実用化場面では、機材の規模が制限される場合もある。この場面において、周波数特性のイコライジングを用いて平坦な周波数特性を実現することが考えられる。しかし、イコライジングを行うと、全体的に再生音の音圧が低下する。そのため、より高い復調音圧を求めるために、印加電圧を上げることが必要となる。これをオーバーブーストと呼ぶ。オーバーブーストにより、非線形歪みの増加という問題があるため、本論文では、印加電圧を考慮した音質改善手法を挙げる。この手法は調波信号に着目し、復調音と原音のスペクトルの差異を調べる。提案手法は原音のスペクトルが復調音に再現するように、あらかじめ原音のスペクトルを歪ませる。その上、予期した印加電圧に最適なパラメータ

を生成する。評価実験の結果、提案手法は復調音に原音のスペクトルをより忠実に再現できることが確認できた。

PAL を小型化する時、PAL で形成された音響ビームが拡散する問題点がある。また、応用場面により、さらに鋭いビームが求められることもある。本論文では、より鋭いビームを形成する指向性制御手法を提案する。PAL の原理はパラメトリック効果である。もしパラメトリック効果を空間上部分的に弱めることが可能であれば、PAL の復調音により形成されたビームを制御可能となる。提案手法では、変調波の一部成分の位相を反転することで、より狭いビーム幅を実現する。位相反転する際、変調方式の特徴も考慮する。評価実験の結果、提案手法を用いることで、より狭い音響ビームの形成が確認できた。

Acknowledgment

First and foremost, I would like to thank my supervisor, Professor Nishiura for his generous support and advice throughout the past four years for my master's program and doctoral program. He created a well-equipped research environment and provided me with opportunities of attending academic conferences and events. His guidance helped me all the time in my researches.

Sincere thanks to Professor Yamashita, for acting as my thesis committee and giving me valuable advice and comments on my research. My sincere thanks also go to Professor Taniguchi, for acting as my thesis committee and providing a valuable chance to participate in research project R-GIRO.

I would like to thank Associate Professor Nakayama from Osaka Sangyo University for many useful suggestions and detailed technical assistance. Without his guidance, I could not carry out my research smoothly. I am also grateful to Lecturer Fukumori, Assistant Professor Iwai, Assistant Wang, for sharing knowledge on researches over the past years. I would like to show my appreciation to Professor Kajikawa from Kansai University, Assistant Professor Yatabe from Waseda University for precious comments on my researches. I would also like to offer my special thanks to Assistant Professor Kong, Specially-appointed Assistant Professor Song for sharing knowledge in researches and carrier paths.

I am grateful to members of my lab members in Acoustics & Signal Processing Lab, past and present, Dr. Wakabayashi, Mr. Oshio, Mr. Zuo, Mr. Ogami, Mr. Nakano, Ms. Otsuka, Mr. Zhang, Mr. Cai, Mr. Fujii, Mr. Mori, Mr. Nguyen, Ms. Sayama, Mr. Shimokata, Mr. Shimada, Mr. Suzuki, Mr. Yoshida, Mr. Asahi, Ms. Harada and Secretary Minai for their warm supports.

Last but not least, I would like to thank my family for their accompany.

Contents

List of Figures

List of Tables

List of Denotations

Abbreviations and Acronyms

AAM	Asymmetrical amplitude modulation
AM	Amplitude modulation
C-SSB	Combined single sideband
DSB	Double sideband
DSB-SC	Double side band with suppressed carrier
FM	Frequency modulation
KZK	Khokhlov-Zabolotskaya-Kuznetsov
LSB	Lower sideband
log-TSP	Logarithmic time-stretched pulse
M-DSB	Multiplexed double sideband
MAM	Modified amplitude modulation
PAL	Parametric array loudspeaker
RMSE	Root mean square error
SRAM	Square-root amplitude modulation
SSB	Single sideband
THD	Total harmonic distortion
TSP	Time-stretched pulse
USB	Upper sideband
VBE	Virtual bass enhancement
W-DSB	Weighted double sideband

Symbols for variants

$\mathbf{A}$	Group of a
a	Amplitude of signal
$b(t)$	Sideband wave
$c(t)$	Carrier wave
$E(t)$	Envelope function
$\mathbf{F}$	Group of f
f	Frequency index
f_0	Fundamental frequency
d	Distance
h	Coefficients of filter or transfer function
k	Bin index
L	Sound pressure level
l	Geometric length
m	Modulation factor
N	Total number of harmonic components in a harmonic sound
n	Index of harmonic components
$\mathbf{P}$	Group of φ
$p(t)$	Demodulated sound
R^2	Coefficient of determination in linear regression
$r(f, v)$	Power response at certain frequency f under driving voltage v
S	Power spectrum
$s(t)$	Signal of sound source
t	Time index
$u(t)$	Signal of ultrasonic wave
v	Driving voltage
w	Beamwidth
$x(t)$	Target signal
η	Compensation for harmonic distortion
θ	Directional angle

λ	Logarithmic spectral slope
μ	Exponential coefficient
σ	Standard deviation
φ	Phase of signal
ϕ	Angle between Woofer PAL and Tweeter PAL

Symbols for operation

$\hat{\cdot}$	Phase shift by $\pi/2$
$\bar{\cdot}$	Average
$*$	Convolution
$\mathcal{A}[\cdot]$	Sinusoidal analysis
$\mathcal{C}_{\frac{1}{2}}[\cdot]$	Frequency compression to half
$\mathcal{D}[\cdot]$	Self-demodulation
$\mathcal{E}[\cdot]$	Envelope calculation
$\mathcal{H}[\cdot]$	Hilbert transform
$\mathcal{M}_{\mathrm{S}}[\cdot]$	Spectral slope control model
$\mathcal{M}_{\mathrm{D}}[\cdot]$	Demodulation model
$\mathcal{N}[\cdot]$	Linear normalization
$\mathcal{S}[\cdot]$	Sinusoidal synthesis

Chapter 1

Introduction

1.1 Background

In the modern society, noise is addressed as a form of pollution, along with air pollution, water pollution, soil pollution, and etc. Noise can be a loud sound, unpleasant sound, or unexpected sound. Different from industrial noises and traffic noises which are generally loud and unpleasant, environmental noises are unnoticed and widespread. Aiming for the next-generation community, it is significant to realize technologies with which the sounds can be carried to target listeners only.

A loudspeaker is a common approach to reproducing sounds. In a typical electro-dynamic loudspeaker, the electro-acoustic transducer driven by an electrical signal pushes and pulls the air to form a sound wave. The sound reproduced by electro-dynamic loudspeaker usually holds a wide directivity, therefore, the spreading sound may reach the non-target listeners and become noise for them. Under the growing demand for directional loudspeakers, parametric array loudspeaker (PAL) [1–3] has been attracting attention in recent years. PAL is an application of parametric array [4] in air, it utilizes the straightness of ultrasound, so the reproduced sound can achieve a sharp directivity. An example of utilizing loudspeakers in an art gallery for the acoustic announcement is shown in Fig. 1.1. The sounds reproduced from electro-dynamic loudspeakers overlap so the visitors hear both of the announcements. In contrast, sounds reproduced from PALs can form narrow acoustic beams so the visitors only hear the desired announcement.

An array of electro-dynamic loudspeakers can be utilized to realize a directional sound,

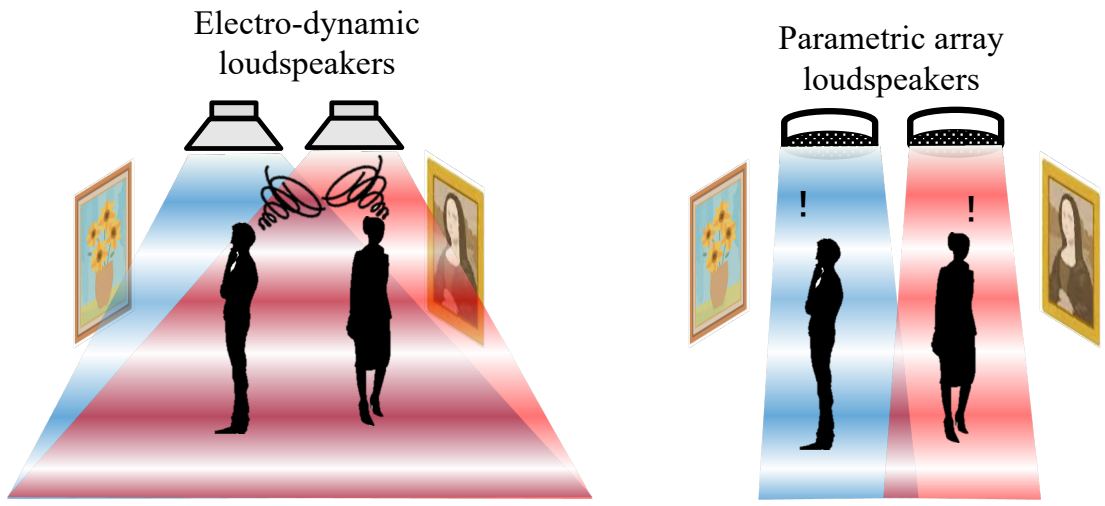


Figure 1.1: Example of loudspeaker utilization in art gallery.

however, a large scale of several meters is necessary to obtain a sharp directivity due to the large wavelength of low-frequency sounds. In contrast, PAL can realize a directional low-frequency sound with a much smaller scale. For example, a commercial PAL (MEE, MSP-50E) holds a diameter of only 20 cm. The small scale of the device brings significant feasibility of private use.

PAL is a promising approach to realizing a directional sound. With the utilization of PAL, people can share a big space without physical partition while enjoying their own acoustic spaces. The noise can be significantly reduced in public areas such as exhibition booths, advertisement billboards, and etc. Under globalization, announcements in different languages can be reproduced simultaneously.

In recent years, trial utilization of PAL has been started in some scenes like announcement and navigation. However, some issues limit the practical use of PAL. For example, a typical PAL is composed of dozens of ultrasonic transducers, which leads to a high fabrication cost. Moreover, it is regarded difficult for PAL to reproduce low-frequency because of low sound pressure in low-frequency. And nonlinear distortions also degrade the sound quality of PAL. In addition, it is necessary for PAL to propagate high-intensity ultrasounds to excite the nonlinear interactions [5], however, it remains a risk of ultrasound exposure [3]. Aiming for the practical use of PAL, performance improvement is still in demand.

1.2 Motivation and Objective

The motivation and objective of this thesis lie on the performance improvement of PAL for practical use. Specifically, sound quality improvement and directivity control are focused on.

The major drawback of PAL is the low sound quality. The reasons for the low sound quality can be concluded in two aspects: difficulty in low-frequency reproduction and nonlinear distortions. It is known theoretically that PAL holds a frequency-dependent response [6], therefore, the reproduced sound shows a low sound pressure in low-frequency. Psychoacoustical approach can be applied to enhance the low-frequency perception, however, physical reproduction of low-frequency is still in demand. It should be noticed that the audible sound reproduced by PAL shows a much lower sound pressure than the emitted ultrasound [3], so equalizing the frequency response will lead to low efficiency. Therefore, it is necessary to investigate an efficient way to improve the sound quality by enhancing the low-frequency.

Another reason for the low sound quality is the nonlinear distortions including harmonic distortions and inter-modulation distortions. The principle of PAL is based on nonlinear interactions in the air. In some sense, the audible sound reproduced by PAL can be regarded as a nonlinear distortion generated from the emitted ultrasound. Therefore, the occurrence of nonlinear distortions is unavoidable. Some researches focused on the elimination of harmonic distortions, however, in some cases the harmonic distortion can be utilized to contribute to the reproduced sound. To realize a higher sound quality, both elimination and utilization should be considered.

The nonlinearity in PAL also brings difficulty in the theoretical analysis for complicated signals. In practical use, speech or music will be reproduced from PAL rather than simple signals like pure tone or white noise. Therefore, it is necessary to develop a signal processing technology to improve the sound quality which is specialized for harmonic sounds.

In terms of directivity control of PAL, the demand for narrow beamforming should be noticed. The scale of the device should be compact enough for private use. However, PAL holds a tendency that the acoustic beam diffuses when the number of ultrasonic transducers is small. This can be explained by that the parametric effect is insufficient and a narrow acoustic beam cannot be formed. Therefore, a feasible narrow-beamforming with a small scale and low cost is of great importance to the popularization of PAL.

1.3 Organization and Major Contributions

The overview of this thesis is shown in Fig. 1.2. Driven by the motivations introduced in Section 1.3, this thesis proposes performance improvement aiming for the practical use of PAL. In specific, sound quality improvement and directivity control are studied under two kinds of conditions. In the first condition, the device scale is not limited, therefore, peak performance is demanded. In the other condition, the device is desired to be compact for private use.

The sound quality improvement for PAL is studied to realize a flat frequency response. Due to the difficulty in low-frequency reproduction, it is considered that the realization of flat frequency response needs an increase of propagation energy. Considering the two conditions, the flat frequency response of PAL can be realized with two aspects. In the first one, the scale of PAL can be enlarged by increasing the number of ultrasonic transducers. An effective enhancement of low-frequency is desired for wide-band sound reproduction. In the other aspects, the driving voltage should be increased for a higher total demodulated sound pressure. The influence brought by high voltage is investigated and tackled. Moreover, the directivity control is studied under the condition that the device scale is limited. The acoustic beam formed by small-scale PAL spreads, therefore, this thesis investigates narrow-edged beamforming.

As shown in Fig. 1.2, the remainder of this thesis is organized as: Chapter 2 gives an introduction about theory, limitation and conventional studies for PAL; Chapter 3 reviews a proposed method for sound quality improvement based on multi-way structure; Chapter 4 gives a proposed method for sound quality improvement based on spectral slope control; Chapter 5 introduces a proposed method for directivity control with phase inversion. Finally, Chapter 6 concludes this thesis.

The Major contributions of this thesis can be highlighted as follows:

1. *Realization of flat response based on multi-way-structured PAL.* A multi-way PAL is proposed, which consists of Woofer PAL and Tweeter PAL. To enhance the low-frequency, Woofer PAL has a larger scale than Tweeter PAL. Suitable modulation methods and hardware arrangements are studied to achieve a flat frequency response. Harmonic distortions are utilized to enhance the low-frequency when reproducing wideband signals. A prototype has been developed based on the theoretical proposal, and the experimental results show that this proposed method realizes a sharp directivity in low-frequency while achieving a sound pressure improvement of 10 dB in

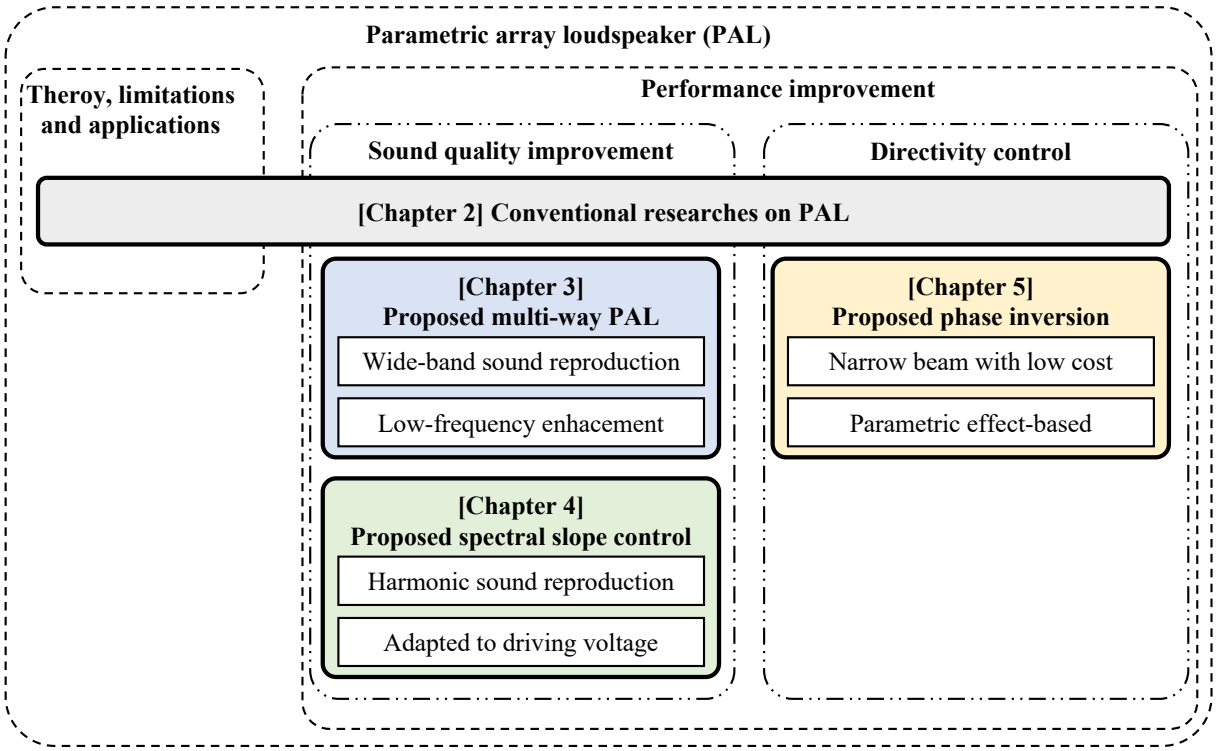


Figure 1.2: Overview of this thesis.

low-frequency.

2. *Equalization of frequency response based on spectral slope control for harmonic sounds.* An equalization method specialized to harmonic sounds is proposed. The difference of spectra between the demodulated sound and the original sound is investigated and models are built with experimental data. The shape of the spectrum is pre-distorted to let the original shape be restored in the demodulated sound of PAL. The influence of driving voltage is first taken into account in this proposed method. In this proposed method, optimal parameters can be generated under any condition of the driving voltage. The nonlinearity of PAL brings a tendency that the distortions get more as the driving voltage increases. This proposed method is of great importance to practical use because it is considerable to increase the volume temporarily.

3. *Narrow beamforming based on phase inversion.* Narrow beamforming is proposed based on phase inversion considering the parametric effect. In this proposed method, PAL is divided into main array and sub-array. Phase-inverse ultrasound is emitted from the sub-array in order to weaken the parametric effect spatially. This method can be applied to small-scale PAL whose acoustic beam tends to diffuse with a low fabrication cost.

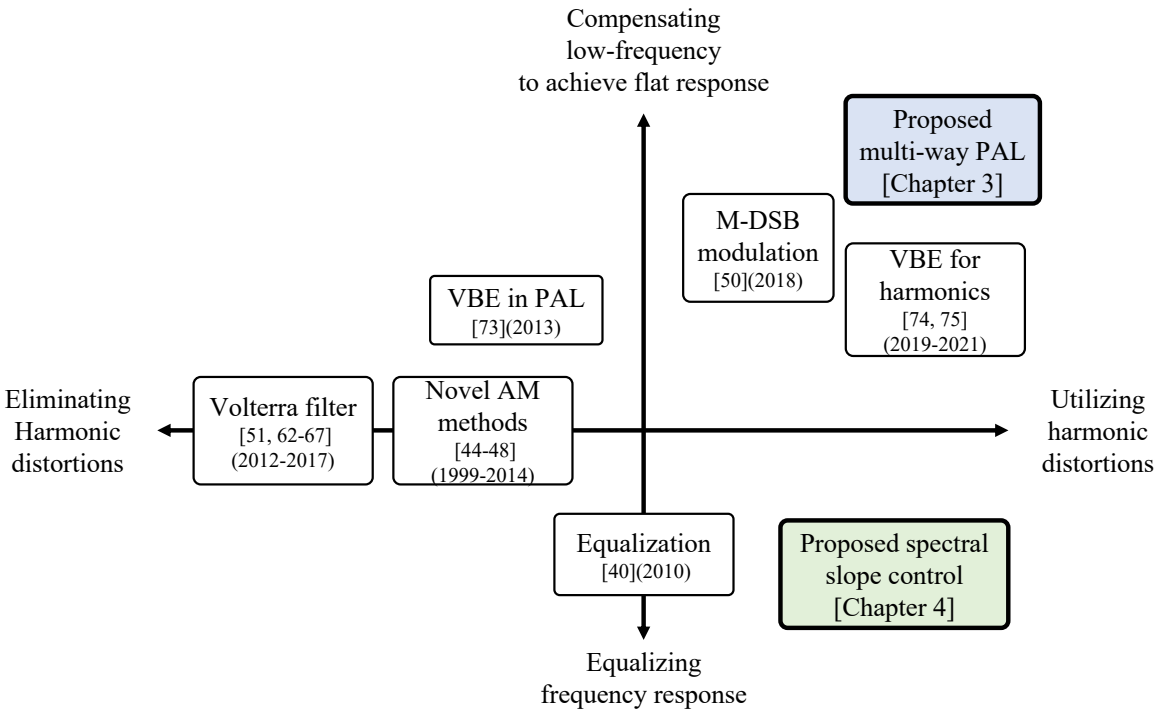


Figure 1.3: Related works on sound quality improvement for PAL.

Related works on sound quality improvement for PAL are summarized in Fig. 1.3. There are two major aspects of sound quality improvement. One is the flattening of frequency response, which can be achieved by compensating the low-frequency or equalizing the frequency response. The other one is the issue of harmonic distortions. It is considerable to eliminate the harmonic distortions to improve the sound quality, while the harmonic distortions can be utilized effectively. This thesis gives two approaches to sound quality improvement. One is the multi-way PAL which will be introduced in Chapter 3. This method realizes a flat response by compensating low-frequency and utilizes the harmonic distortion to further enhance low-frequency. The other proposed method is spectral slope control which will be introduced in Chapter 4. This method focuses on equalizing the frequency response while utilizing the harmonic distortion in the equalization. This method also shows novelty in considering the influence of the driving voltage.

Related works on directivity control for PAL are summarized in Fig. 1.4. The directivity control can be realized by hardware arrangement and array signal processing. An important metric is a cost in computation and implementation. This thesis gives an approach to control the directivity based on phase inversion. This proposed method can be regarded as a hybrid of hardware arrangement and array signal processing. In this method, two signals are generated

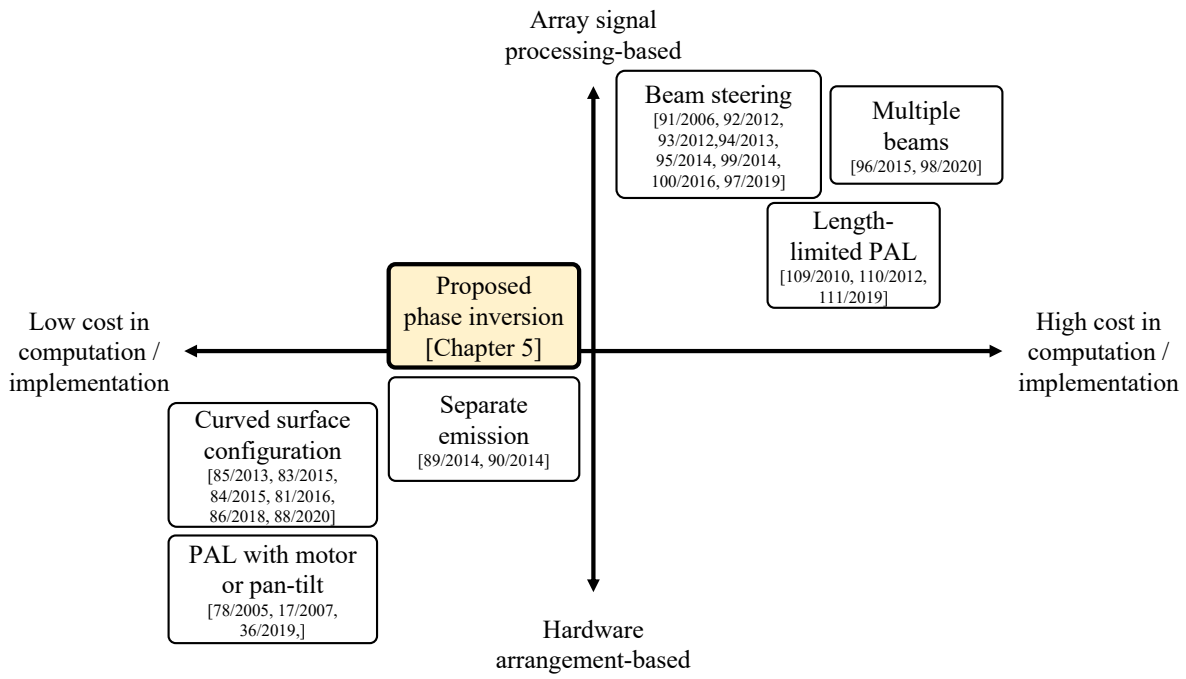


Figure 1.4: Related works on directivity control for PAL.

and emitted so the cost in computation and implementation is relatively lower than approaches that need dozens of signals.

Chapter 2

Conventional Researches on Parametric Array Loudspeaker

2.1 Fundamental Principle of Parametric Array Loudspeaker

It is known that a sound wave distorts as it propagates due to the nonlinear interactions of the medium. If two waves with frequencies are propagated intensely in the air, mixed frequencies and harmonics will be generated during the propagation [5]. The high frequencies tend to be absorbed by the medium easily, however, the difference frequencies can travel along the path of propagation. Based on the scattering of sound by sound [7], Westervelt devised the parametric array [4]. As shown in Fig. 2.1, the generated virtual secondary sources form an end-fire array in the primary beam. Therefore, the difference-frequency wave can achieve a very narrow beam even at low frequencies. The parametric array was first demonstrated underwater [8,9], and then applied as an airborne loudspeaker, named "audio spotlight" as well [10].

The fundamental principle of PAL is based on the parametric effect. When two ultrasonic waves with close frequencies f_a and f_b are emitted intensely in the air, based on the nonlinearity, secondary waves ($f_a - f_b$, $f_a + f_b$, $\cdots$, $n_a f_a \pm n_b f_b$) can be observed on the propagation axis. The ultrasonic components decay rapidly due to the absorption of the air. Despite the possibility that they remain and travel to the position of the listener, they will not be perceived because they are inaudible. In contrast, the difference frequency ($f_a - f_b$) in the audible range decays less, and it is amplified within a distance due to the virtual end-fire array. Based on this property, an

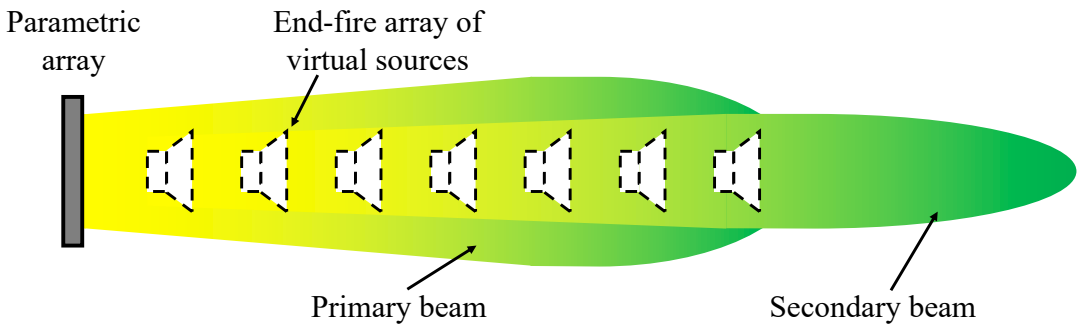


Figure 2.1: Generation of sound beam using parametric array.

audible sound with frequency f_{S} can be reproduced from PAL by emitting a pair of ultrasounds whose difference frequency is exactly f_{S}. The generation of ultrasounds based on the target signal can be also called "modulation" since it is a conversion of audible sound to ultrasound. The reproduction of the target sound can be called "self-demodulation" since the conversion of ultrasound to audible sound occurs in the air without any human intervention.

Double sideband (DSB) modulation [10] is the simplest modulation method to implement by either analog circuit or digital signal processing. A DSB-modulated wave can be generated with an audible target signal $s(t)$ and a carrier wave $c(t)$, and it can be expressed as:

$$\mathrm{DSB}[s(t)] = (1 + m_{\mathrm{DSB}} s(t))\, c(t), \tag{2.1}$$

where m_{DSB} denotes the modulation factor of DSB modulation $(0 < m_{\mathrm{DSB}} \leq 1)$. In most cases, the carrier wave $c(t)$ is a pure tone of ultrasound. Setting the simplest case that the target signal $s(t)$ is a pure tone in audible frequency range, the followings hold

$$c(t) = a_{\mathrm{C}} \cos(2\pi f_{\mathrm{C}} t), \tag{2.2}$$

$$s(t) = a_{\mathrm{S}} \cos(2\pi f_{\mathrm{S}} t), \tag{2.3}$$

$$m_{\mathrm{DSB}} = \frac{a_{\mathrm{S}}}{a_{\mathrm{C}}}, \tag{2.4}$$

where a_{C} and f_{C} denote the amplitude and frequency of carrier wave respectively, a_{S} and f_{S} denote the amplitude and frequency of original signal respectively. Eq. (2.1) can be transformed into the followings:

$$\mathrm{DSB}[a_{\mathrm{S}} \cos(2\pi f_{\mathrm{S}} t)] = \frac{1}{2} a_{\mathrm{S}}^2 \cos(2\pi (f_{\mathrm{C}} + f_{\mathrm{S}}) t) + \frac{1}{2} a_{\mathrm{S}}^2 \cos(2\pi (f_{\mathrm{C}} - f_{\mathrm{S}}) t) + a_{\mathrm{C}} \cos(2\pi f_{\mathrm{C}} t). \tag{2.5}$$

It can be noticed from Eq. (2.5) that the DSB modulated wave consists of three frequency components: upper sideband ($f_{\mathrm{C}} + f_{\mathrm{S}}$), lower sideband ($f_{\mathrm{C}} - f_{\mathrm{S}}$) and carrier wave ($f_{\mathrm{C}}$). The difference frequency between the upper sideband and the carrier wave is f_{S}, and the difference frequency between the lower sideband and the carrier wave is also f_{S}. Therefore, it is known that the target signal with f_{S} can be self-demodulated. However, it can be noticed that the difference frequency between the upper and lower sidebands is $2f_{\mathrm{S}}$. This component will also be self-demodulated due to the nonlinearity, and it is usually regarded as a second-order harmonic distortion because its frequency doubles that of the target signal $s(t)$.

Single sideband (SSB) modulation [1, 11–13] only employs one sideband, so the second-order harmonic distortion will not be self-demodulated theoretically. SSB modulation can be implemented by filtering out the undesired sideband from the DSB-modulated wave or utilizing Hilbert transform. SSB modulation can be categorized into two types: LSB modulation employing the lower sideband and USB modulation employing the upper sideband. LSB-modulated wave and USB-modulated wave can be expressed as:

$$\mathrm{LSB}[s(t)] = (1 + m_{\mathrm{SSB}} s(t)) \cos(2\pi f_{\mathrm{C}} t) + m_{\mathrm{SSB}} \hat{s}(t) \sin(2\pi f_{\mathrm{C}} t), \tag{2.6}$$

$$\mathrm{USB}[s(t)] = (1 + m_{\mathrm{SSB}} s(t)) \cos(2\pi f_{\mathrm{C}} t) - m_{\mathrm{SSB}} \hat{s}(t) \sin(2\pi f_{\mathrm{C}} t), \tag{2.7}$$

where m_{SSB} denotes the modulation factor of SSB modulation ($0 < m_{\mathrm{SSB}} \leq 1$), $\hat{s}(t)$ denotes the signal with a phase shift of $\pi/2$ to $s(t)$. Besides of DSB and SSB modulation, other modulation methods have been proposed for PAL. These modulation methods will be mentioned in Section 2.3.1.

In order to analyze the sound field of the demodulated sound, the most widely-used model is Khokhlov–Zabolotskaya–Kuznetsov (KZK) equation, involving nonlinearity, diffraction, and dissipation [14]. Berktay provided an approximate solution of KZK equation [6], and it is regarded as a theoretical basis for predicting the far-field response of a PAL. Berktay's far-field solution can be simplified as:

$$p(t) \propto \frac{\partial^2 E^2(t)}{\partial t^2}, \tag{2.8}$$

where t denotes the time index, $E(t)$ denotes the envelope function of the primary wave, and $p(t)$ denotes the demodulated sound on the propagation axis. It can be noticed that the demodulated sound can be theoretically predicted if the envelope of the primary wave is known.

Utilizing Berktay's far-field solution, the demodulated sound of DSB modulation and SSB modulation can be predicted. The envelope of DSB-modulated wave can be expressed as:

$$\mathcal{E}[\mathrm{DSB}[s(t)]] = 1 + m_{\mathrm{DSB}} s(t). \tag{2.9}$$

The demodulated sound with DSB modulation can be obtained as:

$$p_{\mathrm{DSB}}(t) \propto m_{\mathrm{DSB}} a_{\mathrm{S}} f_{\mathrm{S}}^2 \cos(2\pi f_{\mathrm{S}} t) + m_{\mathrm{DSB}}^2 a_{\mathrm{S}}^2 f_{\mathrm{S}} \cos(2\pi (2 f_{\mathrm{S}}) t). \tag{2.10}$$

LSB- and USB-modulated wave share a same envelope and can be expressed as:

$$\mathcal{E}[\mathrm{SSB}[s(t)]] = \sqrt{1 + 2 m_{\mathrm{SSB}} s(t) + m_{\mathrm{SSB}}^2 s^2(t) + m_{\mathrm{SSB}}^2 \hat{s}^2(t)}. \tag{2.11}$$

The demodulated sound with SSB modulation can be obtained as:

$$p_{\mathrm{SSB}}(t) \propto m_{\mathrm{SSB}} a_{\mathrm{S}} f_{\mathrm{S}}^2 \cos(2\pi f_{\mathrm{S}} t). \tag{2.12}$$

From Eqs. (2.10) and (2.12), it can be noticed that the demodulated sound of DSB modulation contains harmonic distortion, while the demodulated sound of SSB modulation contains the target signal only in theory.

The most remarkable characteristic of PAL is its ability to form a narrow audible sound beam. Therefore, PAL can be used to create audio spotlights or project the sound to target listeners [10,15–18]. Further, PAL has been applied to surround sound systems with conventional electro-dynamic loudspeakers [19–23]. PAL can also be utilized in active noise control system, instead of conventional electro-dynamic loudspeakers [24–28]. It is reported that bone-borne ultrasound can be perceived by the profoundly deaf [29], therefore, PAL can be utilized as a bone-borne headphone [30]. PAL's sharp directivity also brings unique applications, such as rendering sound images with reflective sounds [31], personal sound zone [32], acoustical space sharing [33], acoustic holography [34–36] and so on.

2.2 Limitations of Parametric Array Loudspeaker

Although PAL can realize a sharper directivity than conventional electro-dynamic loudspeakers, it has two main drawbacks. One is the difficulty in low-frequency reproduction, and the other one is the nonlinear distortions generated in the self-demodulation.

Theoretical analysis based on Berktay's far-field solution shows that the power of demodulated sound is relevant to the frequency of the target signal. From Eqs. (2.10) and (2.12), it can be noticed that in the demodulated sound of DSB and SSB modulation, the component of target frequency f_{S} is proportional to f_{S}^2. Therefore, a simplified conclusion can be drawn approximately that $p \propto f^2$, where p denotes the demodulated sound and f denotes the frequency. This $p \propto f^2$ characteristic suggests that the power of demodulated sound increases 16 times as the frequency doubles, which can be concluded as a +12 dB/octave response. Setting the sound pressure at 2 kHz as a 0 dB reference, it can be derived in theory that the sound pressure at 1 kHz is -12 dB, the sound pressure at 500 Hz is only -24 dB. Obviously, PAL is difficult to reproduce low-frequency due to its physical principle. The difficulty in low-frequency reproduction limits the application of PAL, so the enhancement of low-frequency is an important task for the practical use of PAL.

Equalization can be utilized to obtain a flat frequency response by suppressing the high-frequency and compensating the low-frequency. It can be predicted the total sound pressure will decrease as the equalization is carried out. The sound reproduction of PAL relies on the ultrasonic-audible conversion by nonlinear interactions in the air, and the sound pressure level of demodulated sound is generally 40 dB lower than that of the emitted ultrasound [3]. There is a latent risk of intense ultrasound exposure when the power of emitting ultrasound is increased to reach a higher demodulated sound pressure. It is reported that the burden of PAL is smaller than that of conventional electro-dynamic loudspeakers [37], and it is also known that the exposure threshold can be set less strictly if the exposure duration is short [38]. Nevertheless, the safety issues should be concerned and the outage power of PAL should not be increased unlimitedly to force the reproduction of low-frequency.

Among the nonlinear distortions generated in self-demodulation, harmonic distortion and inter-modulation distortion should be stressed. Assuming that the target signal $s(t)$ is a pure tone at f_{S} as defined in Eq. (2.3), components at $2f_{\mathrm{S}}, 3f_{\mathrm{S}}, \cdots$ can be observed in the demodulated

sound as well as the target signal at f_{S}. These acquired components are harmonic distortion since they hold multiple frequencies to the target signal. In conventional electro-dynamic loudspeakers, harmonic distortion is generated due to the nonlinearity of the analog circuit and the amount of harmonic distortion is usually small. However, in the case of PAL, the amount of harmonic distortion is much more significant. From Eq. (2.10), it can be noticed that a second-order harmonic distortion can be predicted in theory when DSB modulation is utilized. From Eq. (2.12), it can be noticed that SSB modulation generates zero harmonic distortion in theory, however, it is reported that the measured harmonic distortion differs from the prediction [13, 39, 40].

Inter-modulation occurs when the target signal consists of two or more components. Assuming that the target signal $s(t)$ is a dual tones at f_{S1}, f_{S2} and ignoring the phases of each component, the following holds:

$$s(t) = a_{\mathrm{S1}} \cos(2\pi f_{\mathrm{S1}} t) + a_{\mathrm{S2}} \cos(2\pi f_{\mathrm{S2}} t), \tag{2.13}$$

The demodulated sound with DSB modulation can be obtained as:

$$\begin{aligned} p_{\mathrm{DSB}}(t) \propto\ & m_{\mathrm{DSB}} a_{\mathrm{S1}} f_{\mathrm{S1}}^2 \cos(2\pi f_{\mathrm{S1}} t) + m_{\mathrm{DSB}} a_{\mathrm{S2}} f_{\mathrm{S2}}^2 \cos(2\pi f_{\mathrm{S2}} t) \\ & + m_{\mathrm{DSB}} a_{\mathrm{S1}} f_{\mathrm{S1}}^2 \cos(2\pi (2 f_{\mathrm{S1}}) t) + m_{\mathrm{DSB}} a_{\mathrm{S2}} f_{\mathrm{S2}}^2 \cos(2\pi (2 f_{\mathrm{S2}}) t) \\ & + \frac{1}{2} m_{\mathrm{DSB}}^2 a_{\mathrm{S1}} a_{\mathrm{S2}} (f_{\mathrm{S1}} + f_{\mathrm{S2}})^2 \cos\left(2\pi (f_{\mathrm{S1}} + f_{\mathrm{S2}}) t\right) \\ & + \frac{1}{2} m_{\mathrm{DSB}}^2 a_{\mathrm{S1}} a_{\mathrm{S2}} (f_{\mathrm{S1}} - f_{\mathrm{S2}})^2 \cos\left(2\pi (f_{\mathrm{S1}} - f_{\mathrm{S2}}) t\right). \end{aligned} \tag{2.14}$$

And the demodulated sound with SSB modulation can be obtained as:

$$\begin{aligned} p_{\mathrm{SSB}}(t) \propto\ & m_{\mathrm{SSB}} a_{\mathrm{S1}} f_{\mathrm{S1}}^2 \cos(2\pi f_{\mathrm{S1}} t) + m_{\mathrm{SSB}} a_{\mathrm{S2}} f_{\mathrm{S2}}^2 \cos(2\pi f_{\mathrm{S2}} t) \\ & + m_{\mathrm{SSB}}^2 a_{\mathrm{S1}} a_{\mathrm{S2}} (f_{\mathrm{S1}} - f_{\mathrm{S2}})^2 \cos\left(2\pi (f_{\mathrm{S1}} - f_{\mathrm{S2}}) t\right). \end{aligned} \tag{2.15}$$

From Eqs. (2.14) and (2.15), it can be noticed that some mixed frequencies at $f_{\mathrm{S1}} + f_{\mathrm{S2}}$ and $f_{\mathrm{S1}} - f_{\mathrm{S2}}$ can be observed in the demodulated sound of DSB and SSB modulation, besides of target signal components and harmonic distortions. These components are called inter-modulation distortion. For target signals with few components, such as pure tone, the harmonic distortions take the majority of nonlinear distortions. For target signals with many components, the inter-modulation distortion also affect the perception of demodulated sound. In practice, high-order harmonic distortions and inter-modulation distortions can be also observed due to more complicated interactions. The nonlinear distortions degrade the demodulated sound quality, therefore, the reduction of nonlinear distortion is also an important task for the practical use of PAL.

2.3 Related Works

Researches related to PAL are introduced in this section. Modulation methods are summarized in Section 2.3.1. Other studies are categorized based on their purposes. Specifically, approaches aimed to improve the sound quality are introduced in Section 2.3.2. Psychoacoustical approaches to enhance the perception of demodulated sound are introduced in this section. Researches to control the directivity of PAL are introduced in Section 2.3.3.

2.3.1 Modulation Methods

Basic modulation methods for PAL, DSB modulation [10], has been introduced in Section 2.1. The major shortage of DSB modulation is harmonic distortion. It is reported that the second harmonic distortion reaches a similar level to the fundamental signal in the case that the target signal is a pure tone and the value of modulation factor is high [3,10]. Therefore, DSB modulation is generally not recommended except for some cases that the harmonic distortions are utilized on purpose [3].

To reduce the harmonic distortion of DSB modulation, square-root amplitude modulation (SRAM) [41–44] was proposed. The inspiration of SRAM is that the envelope function is squared based on Berktay's far-field solution shown in Eq. (2.8). Therefore, in SRAM, a square-root operation is carried out in advance. It is reported that SRAM generated much less harmonic distortion than DSB modulation [44]. However, the SRAM-modulated wave has an infinite bandwidth due to the square-root operation. Ultrasonic transducers with flat responses in a wide frequency range are necessary to achieve the ideal performance of SRAM. Therefore, SRAM suffers from the limitation in the fabrication of ultrasonic transducers because the typical ultrasonic transducers can achieve sufficient sound pressure in a narrow frequency range.

Another approach to reduce the harmonic distortion of DSB modulation is SSB modulation [1, 11–13], which has also been introduced in Section 2.1. Though SSB modulation has zero harmonic distortion, in theory, experimental results show that the harmonic distortion can be observed in the demodulated sound of SSB modulation [13, 39, 40]. Some variants of SSB modulation have also been proposed. Dynamic SSB modulation was proposed to reduce the power consumption and the ultrasound exposure [12]. Recursive SSB modulation was proposed to approximate the envelope generated by the SRAM [11,45].

Modified amplitude modulation (MAM) combines the advantages of SSB modulation and SRAM [40, 46]. The process of MAM can be viewed as DSB modulation with an orthogonal carrier modulating a pre-distortion term. By adding this pre-distortion term, the distortion can be completely removed if the ultrasonic transducer has infinite bandwidth as required in SRAM. Therefore, a finite approximation using the Taylor series is carried out. It is reported that MAM can achieve low harmonic distortion as SSB and lower inter-modulation distortion than other modulation methods.

Asymmetrical amplitude modulation (AAM) was proposed to extend the frequency range of the demodulated sound [47]. In modulation methods mentioned before, the harmonic distortion of DSB modulation is regarded as a useless by-product. However, in AAM, some selected second harmonic distortions are utilized to extend the high-frequency response, and the other unwanted second harmonic distortions are suppressed by using the SSBAM.

Most amplitude modulation methods need an individual carrier wave to create the difference frequency at the frequency of the target signal. However, the carrier wave can be removed in New DSB modulation [48]. It can also be called DSB with suppressed carrier (DSB-SC) modulation which is widely used in electromagnetic wave communication. DSB and SSB modulation create the difference frequency at f_{S} with sideband $f_{\mathrm{C}} \pm f_{\mathrm{S}}$ and carrier f_{C}. However, in DSB-SC modulation, the difference frequency at f_{S} is created with $f_{\mathrm{C}} + \frac{1}{2} f_{\mathrm{S}}$ and $f_{\mathrm{C}} - \frac{1}{2} f_{\mathrm{S}}$ by compressing the frequency of target signal with 1 octave. In modulation with a carrier, the modulation factor should be limited up to 1, in order to avoid over-modulation. Since DSB-SC modulation employs no carrier, the modulation factor can be removed and there is no risk of over-modulation. Experimental results show that DSB-SC modulation generates less harmonic distortion but brings more inter-modulation distortions into the demodulated sound.

As discussed, the demodulated sound pressure of PAL is low due to the low energy-conversion efficiency in self-demodulation. So there is a trade-off between the sound pressure and the sound quality in practical cases. The sound pressure may take priority in low-frequency because it is difficult for PAL to reproduce low-frequency sound. In contrast, the sound quality may take priority in high-frequency because PAL can achieve an acceptable sound pressure in this frequency range. Weighted double sideband (W-DSB) modulation [49] has been proposed to reduce the harmonic distortions at high-frequency and emphasize the sound pressure at low-frequency. In W-DSB modulation, DSB modulation is employed in low-frequency for higher

sound pressure, and SSB modulation is employed in high-frequency for less harmonic distortions.

Though it is theoretically predicted that PAL is difficult to reproduce low-frequency, modulation methods in order to enhance the low-frequency have been proposed. Multiplexed double sideband (M-DSB) modulation [50] has been proposed to emphasize the low-frequency in the demodulated sound of PAL. In M-DSB, DSB modulation is combined with DSB-SC modulation, and the second-order harmonic distortion is utilized to enhance the low-frequency. It is experimentally validated that M-DSB modulation can achieve a higher sound pressure over traditional modulation methods, however, it brings a half-frequency distortion as it utilizes DSB-SC modulation.

Besides of amplitude modulation (AM), frequency modulation (FM) can also be utilized in PAL [51]. Frequency modulation is favorable for low fabrication cost and high sound pressure, but it suffers from a larger amount and high order nonlinear distortion. Facing the trade-off between sound pressure and sound quality, hybrid methods combining AM and FM have been proposed [52, 53].

Different from FM which carries the information of the target signal on the carrier frequency, AM with time-varying carrier frequency has been proposed [54–57]. It is reported that a constant carrier frequency can cause fatigue error in PAL and lead to spectral peak noise. The carrier frequency varies as time so that the spectral peak noise can be reduced and the lifespan of PAL can be extended.

Besides of changing the carrier frequency, the shape of the carrier wave can also be changed [57, 58]. Carrier waves in traditional modulation methods are the sinusoidal wave, however, in this method, the carrier wave is changed into the rectangular wave to carry more energy. Since the sinusoidal carrier distorts in propagation due to the shock wave theory, this AM with a rectangular carrier does not show a significant sound quality degradation than common AM methods.

2.3.2 Sound Quality Improvement

To improve the demodulated sound quality of PAL, the frequency-dependent response should be compensated, and the harmonic distortions need to be eliminated. Based on Berktay's far field solution [6], PAL shows a +12 dB/octave frequency response. Therefore, a -12 dB/octave

equalization can be utilized to compensate for the frequency response [40]. Further, the effect of ultrasonic transducers used to build the PAL is studied [59].

Volterra filter [60, 61] has been applied to PAL to eliminate the second-order harmonic distortion [51, 62–67]. Volterra filter is a nonlinear model based on Volterra series, which is a linear combination of the nonlinear functions. It can express the second-order harmonic distortion of the system so it has been applied to PAL. In these studies, the system is first identified by Volterra filter so the second-order harmonic distortion can be predicted. Then a phase-inverse harmonic is added to the target signal to remove the second-order harmonic distortion in the self-demodulation. Volterra filter suffers from the computation cost because the number of Volterra kernels increases at an exponential rate as the order increases. Therefore, the order of Volterra filter is usually limited to 2 or 3 considering the computation cost in practical use.

In the studies of modulation methods, it is known that the modulation factor can affect the demodulated sound pressure and sound quality [3, 40]. Therefore, automatic gain control has been applied to PAL in which the modulation factor is determined automatically according to the characteristics of the target signal [68].

The difficulty in the low-frequency reproduction of PAL is physically unsolvable, however, the psychoacoustical approach can be applied to PAL. Missing fundamental is a psychoacoustical phenomenon in pitch perception [69, 70]. It is so-called because a pitch perceived by the human auditory system can be at a frequency for which no actual sound source exists. For example, a complex tone consisting of a harmonic series (frequencies: $3f_0, 4f_0, 5f_0, 6f_0$) is reproduced, and the listener perceives a pitch of f_0 though there is no vibration for f_0. The neural mechanism of missing fundamental is unsolved, but its occurrence has been verified. It is also confirmed that humans can perceive the pitch of a fundamental frequency with a complex tone in which all components have a frequency of over 1000 Hz [71]. This implies that it is possible to give the perception of a low frequency only with high frequencies. Virtual bass enhancement (VBE) [72] is an application of missing fundamental. It generates a harmonic series with a non-linear device to help listeners perceive low-frequency pitches, instead of reproducing bass sound directly. Utilizing missing fundamental phenomenon, harmonic series create the perception of a pitch below the frequency range of loudspeakers.

VBE is then applied to PAL to enhance the low-frequency perception [73]. Experimental results show that this approach improves the low-frequency perception with the cost of a loss

in sound quality. There are also latent problems that it has a risk of causing clipping distortion, and the sound pressure of harmonic sound is insufficient.

To solve the problems, an advanced VBE for PAL has been proposed [74, 75]. As mentioned, the demodulated sound with PAL has a frequency-dependent pressure, so it is effective to enhance the sound by creating the perception of a low-frequency pitch with high-frequency components. The energy conversion efficiency can be improved to enhance the total sound pressure by allocating the energy of low frequencies to high ones. Higher sound pressure can be obtained since the energy of the original signal is concentrated on high-frequency components. Moreover, harmonic distortion control is carried out in which harmonic components are adjusted to equalize harmonic distortion, allowing the components to contribute to the sound pressure. Experimental results validated significant improvement of demodulated sound pressure, but this method remains the problem of the runaway of distortions.

Besides of VBE based on nonlinear function [73] and VBE based on energy re-allocation among harmonic structure [74,75], VBE based on phase vocoder has also been applied for PAL combined with other signal processing techniques to realize the improvement of bass sound quality for PAL [76].

2.3.3 Directivity Control

The directivity control of PAL can be realized by physical approaches and signal processing techniques [77]. Since PAL has a sharp directivity, the target zone, where the demodulated sound can be heard, can be controlled by changing the direction of PAL by rotation stage of pan-tilt unit [17,36,78]. Ultrasound holds straightness due to its high-frequency, therefore, the reflected sound beam can be utilized to carry the sound to desired positions [31, 79–82]. To achieve desired directivity, the ultrasonic transducers of PAL can be configured on a curved surface [81,83,84] or on a spherical surface [85–88]. The self-demodulation occurs when two intense ultrasounds conflict, therefore, the ultrasounds can be emitted separately to generate the demodulated sound at the overlapped position [89,90].

Array signal-processing techniques based on antenna theory have been applied to PAL [91–98]. To encounter the computation cost of array signal-processing, DSP board [91] and FPGA [93] can be utilized for implementation. Recently, the number of the steered beam is no longer

limited to 1 [96,98]. The sideband and carrier wave can also be steered individually [99,100]. In beam steering, spatial aliasing [101] and grating lobe [102] should be concerned. Moreover, the directivity models of PAL have also been studied, such as product directivity principle [91,103] and convolution model [104,105]. Besides of the beam steering of PAL, the beam width control and beam length control are also under-studied. The beamwidth can be controlled as constant to frequency [106–108]. Length-limited PAL has also been developed [108–111].

Chapter 3

Sound Quality Improvement Based on Multi-way Structure for Parametric Array Loudspeaker

3.1 Introduction

Electro-dynamic loudspeakers have been widely popularized as an approach to reproducing audible signals. Due to the wide directivity of electro-dynamic loudspeakers, the spreading sounds may become noise to non-listeners. PAL [1–3, 10, 11, 15] has been drawing attention due to their sharp directivity. It can generate a directional audible sound-beam, and the sound can be carried only to target positions. When emitting two ultrasounds intensely in the air, the difference frequency is generated due to the nonlinearity of the air [5]. Utilizing the self-demodulation phenomenon, it is able to reproduce a target sound by emitting two ultrasounds that the difference frequency is the same as the target sound from PAL. Whilst conventional electro-dynamic loudspeakers emit the target sound directly, a PAL emits an intense ultrasound synthesized by amplitude-modulating an ultrasonic carrier wave with the audible target sound. Due to the non-linear interactions in the air, the emitted sound self-demodulates into audible sounds. The demodulated sound of PAL inherits the straightness of the ultrasounds and achieves sharp directivity.

However, PAL suffers from low sound pressure and difficulty in low-frequency reproduc-

tion. As discussed, the demodulated sound of PAL is generated during the ultrasound-audible conversion due to the nonlinear interactions. The energy-conversion efficiency is low so that the sound pressure level of demodulated sound is generally 40 dB lower than that of the emitted ultrasound [3]. Moreover, the self-demodulation of a PAL is frequency-dependent, so the sound pressure in low-frequency is very low, especially in the bass frequency range (below 250Hz). The difficulty in low-frequency reproduction leads to sound quality degradation and is regarded as an obstacle to the popularization of PAL. Therefore, there remains an essential but problematic task of keeping sharp directivity while reproducing low-frequency with sufficient sound pressure.

Berktay's far field solution [6] indicates that PAL shows a +12 dB/octave frequency response. Based on this theory, the sound pressure at 250 Hz is 36 dB lower than that at 2 kHz, therefore, it is unrealistic to compensate the frequency response of PAL with the equalizer in a wide band. In some studies, the cutoff frequency of PAL is set to 500 Hz and the reproduction of lower frequency is given up [40, 59].

Some researches focused on the psychoacoustical approach to enhance the perception of low-frequency reproduced by PAL [73–75]. The missing fundamental phenomenon is utilized to create a virtual bass perception, however, the sound quality needs to be improved. The virtual bass is based on pitch perception theory [69–72], so the improvement for harmonic signals is significant [74, 75]. However, it is difficult to apply to wideband signals, such as Gaussian noise.

A sound reproduction system combining electro-dynamic sub-woofer and PALs has been proposed [20]. In this system, PALs are employed to reproduce high-frequency, and the electro-dynamic sub-woofer is employed to reproduce low-frequency which is difficult for PAL to reproduce. However, the electro-dynamic sub-woofer owns a very wide directivity, so it remains a problem of the spreading sound in low-frequency.

In order to tackle the task of keeping sharp directivity while reproducing low-frequency with sufficient sound pressure, this thesis proposes a realization of flat frequency response in wideband based on multi-way PAL. Many commercial loudspeaker systems consist of several loudspeaker units: tweeters for high-frequency, woofers for low-frequency, mid-range drivers for mid-frequency, etc. Each reproduces a part of the frequency range, and flat frequency response is achieved with the whole system. Similarly, it is not necessary to force a single PAL to reproduce the whole audible band. The proposed multi-way PAL is composed of Woofer PAL (PAL as a woofer) and Tweeter PAL (PAL as tweeter). A PAL consists of dozens of

ultrasonic transducers, and it is known that the demodulated sound pressure becomes higher as the number of ultrasonic transducers increases [42, 112, 113]. Therefore, Woofer PAL should employ more ultrasonic transducers than Tweeter PAL. Hardware arrangement is studied in order to realize the sound pressure improvement in low-frequency. Moreover, suitable modulation methods are discussed to realize a flat frequency response. Especially, the modulation method employed in Woofer PAL can utilize the harmonic distortion to emphasize the low-frequency reproduction, so that a higher sound pressure in low-frequency can be achieved compared with conventional equalization approaches. A prototype multi-way PAL is developed based on the proposed method, and evaluation experiments are carried out with this prototype to demonstrate the effectiveness.

The remainder of this chapter is organized as follows. In Section 3.2, the proposed PAL with multi-way structure is introduced. In Section 3.3, this thesis describes a developed prototype of multi-way PAL. In Section 3.4, experiments are carried out to evaluate the performance of the proposed method. Finally, this thesis concludes in Section 3.5.

3.2 Proposed Multi-way Parametric Array Loudspeaker

3.2.1 Overview

An overview of the proposed multi-way PAL is shown in Fig. 3.1. It consists of Woofer PAL to reproduce low-frequency sounds and Tweeter PAL to reproduce high-frequency sounds. The target signal $x(t)$ is divided into two frequency bands before the modulation. A low-frequency band signal is to be emitted from Woofer PAL, while a high-frequency band signal is to be emitted from Tweeter PAL. The band division can be expressed as follows.

$$x_{\mathrm{W}}(t) = \frac{\max|x(t)| \cdot x(t) * h_{\mathrm{LPF}_{f_{\mathrm{CO}}}}(t)}{\max|x(t) * h_{\mathrm{LPF}_{f_{\mathrm{CO}}}}(t)|}, \tag{3.1}$$

$$x_{\mathrm{T}}(t) = \frac{\max|x(t)| \cdot x(t) * h_{\mathrm{HPF}_{f_{\mathrm{CO}}}}(t)}{\max|x(t) * h_{\mathrm{HPF}_{f_{\mathrm{CO}}}}(t)|}, \tag{3.2}$$

where $*$ denotes convolution operator, $h_{\mathrm{LPF}_{f_{\mathrm{CO}}}}(t)$ denotes the coefficients of a low-pass filter with a cutoff frequency at f_{CO}, $h_{\mathrm{LPF}_{f_{\mathrm{CO}}}}(t)$ denotes the coefficients of a high-pass filter with a cutoff frequency at f_{CO}. The crossover frequency f_{CO} is set at 1 kHz, therefore, the low-frequency below 1 kHz can be enhanced.

In Eqs. (3.1) and (3.2), low-frequency band signal $x_{\mathrm{W}}(t)$ and high-frequency band signal $x_{\mathrm{T}}(t)$ are normalized to keep a same amplitude as the original target signal $x(t)$. The concept of band division and amplitude normalization is shown in Fig. 3.2. In sweep-sine sound sources

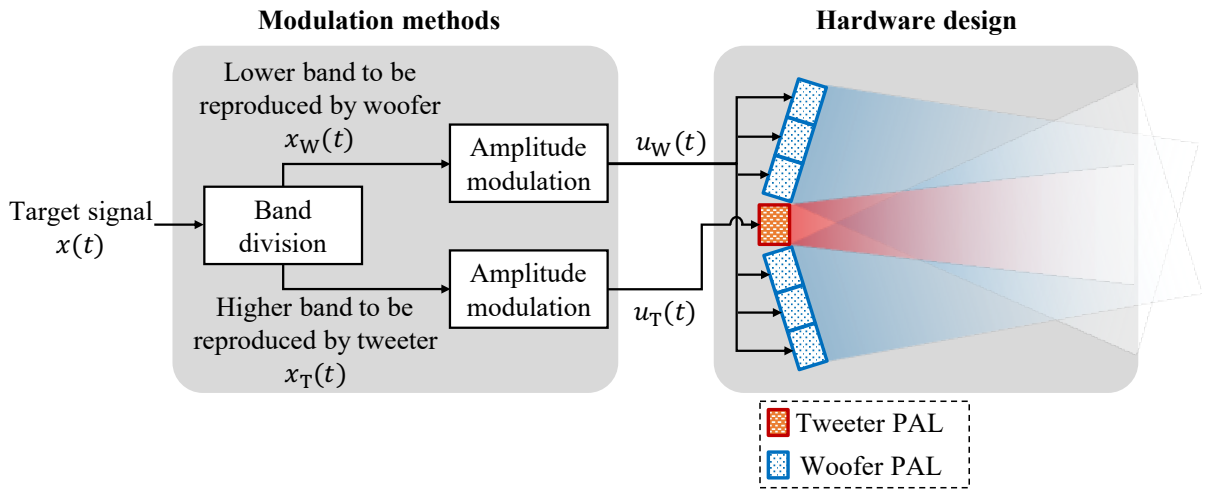


Figure 3.1: Overview of the proposed multi-way parametric array loudspeaker.

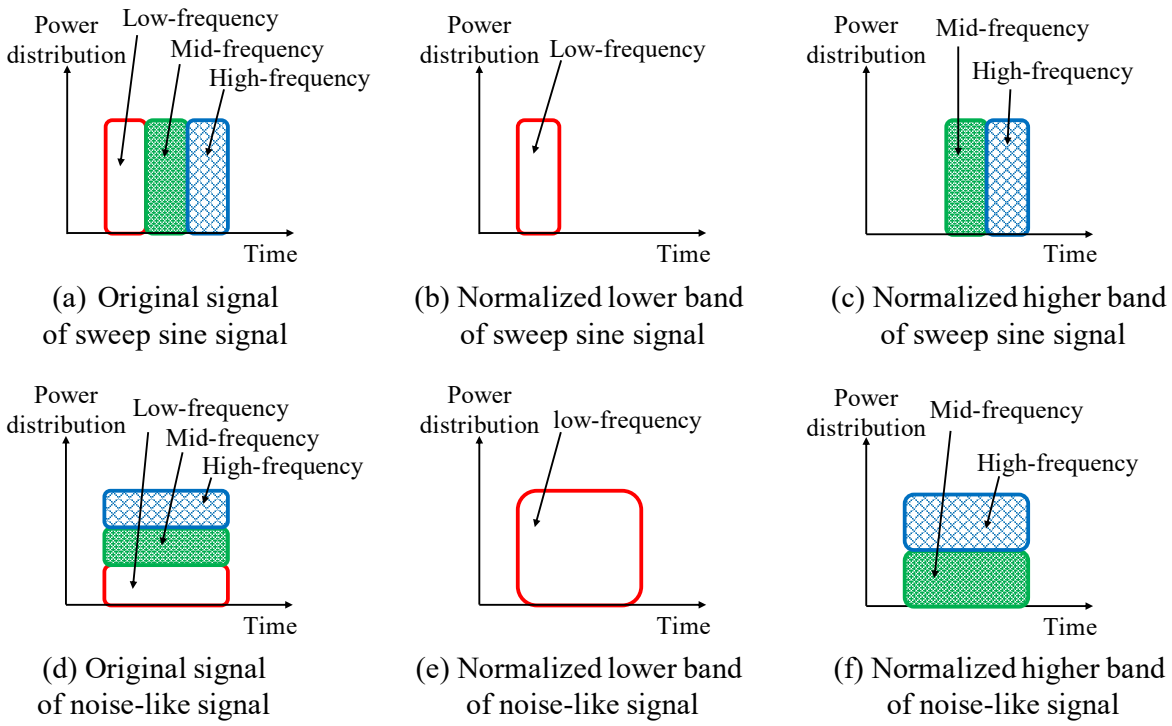


Figure 3.2: Concept of band division and amplitude normalization.

such as time-stretched pulse (TSP) signals, the energy of each frequency band is distributed along the time axis. Therefore, the demodulated sound pressure in low-frequency can be hardly enhanced after the band division and amplitude normalization. However, in wideband signals such as Gaussian noise, the energy of each frequency band is distributed along the frequency axis, in other words, each time interval holds the energy of multiple frequency bands. In this case, the demodulated sound pressure in low-frequency can be significantly enhanced after the band division and amplitude normalization. As discussed, it is difficult to apply the psychoacoustical approaches to wideband signals, so the band division and amplitude normalization are utilized to achieve a significant enhancement in low-frequency.

As shown in Fig. 3.1, the Woofer PAL on the side is configured in a straight line in the proposed method. The aim is to form a planar wave with ultrasonic waves emitted from the Woofer PAL so that the parametric effect can be incited efficiently. Moreover, considering the sharp directivity of PAL, the arrangement of Woofer PAL and Tweeter PAL should be studied to guarantee that the low-frequency sound reproduced by Woofer PAL and the high-frequency sound reproduced by Tweeter PAL can be heard simultaneously. The detail of discussion about

the hardware arrangement will be given in Section 3.2.2.

The low-frequency band signal $x_{\mathrm{W}}(t)$ is amplitude-modulated into ultrasonic signal $u_{\mathrm{W}}(t)$ and then propagated from Woofer PAL. Similarly, the high-frequency band signal $x_{\mathrm{T}}(t)$ is amplitude-modulated into ultrasonic signal $u_{\mathrm{T}}(t)$ and then propagated from Tweeter PAL. In a low-frequency band below 1 kHz, the sound pressure takes priority, so the modulation method which can achieve a higher demodulated sound pressure should be employed. In high-frequency band above 1 kHz, the sound pressure and the sound quality should be balanced. A specific discussion about the modulation methods will be given in Section 3.2.3.

3.2.2 Hardware Arrangement

As shown in Fig. 3.3, the acoustic beam of PAL is approximated to a trapezoid so that the directional angle of PAL can be defined. The calculation of directional angle θ can be expressed as:

$$\theta = \tan^{-1}\left(\frac{w/2 - l/2}{d}\right), \tag{3.3}$$

where d denotes the distance at which the beamwidth is measured, w denotes the measured -6 dB beamwidth at distance d, l denotes the array length or the width of PAL. The directional angle of Woofer PAL θ_{W} can be calculated by replacing w and l in Eq. (3.3) with the measured beamwidth of Woofer PAL w_{W} and the width of Woofer PAL l_{W}. Similarly, the directional angle of Tweeter PAL θ_{T} can be calculated by replacing w and l in Eq. (3.3) with the measured beamwidth of Tweeter PAL w_{T} and the width of Tweeter PAL l_{T}.

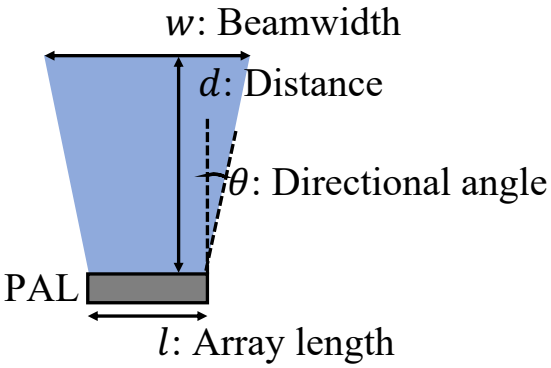


Figure 3.3: Definition of directional angle θ for parametric array loudspeaker.

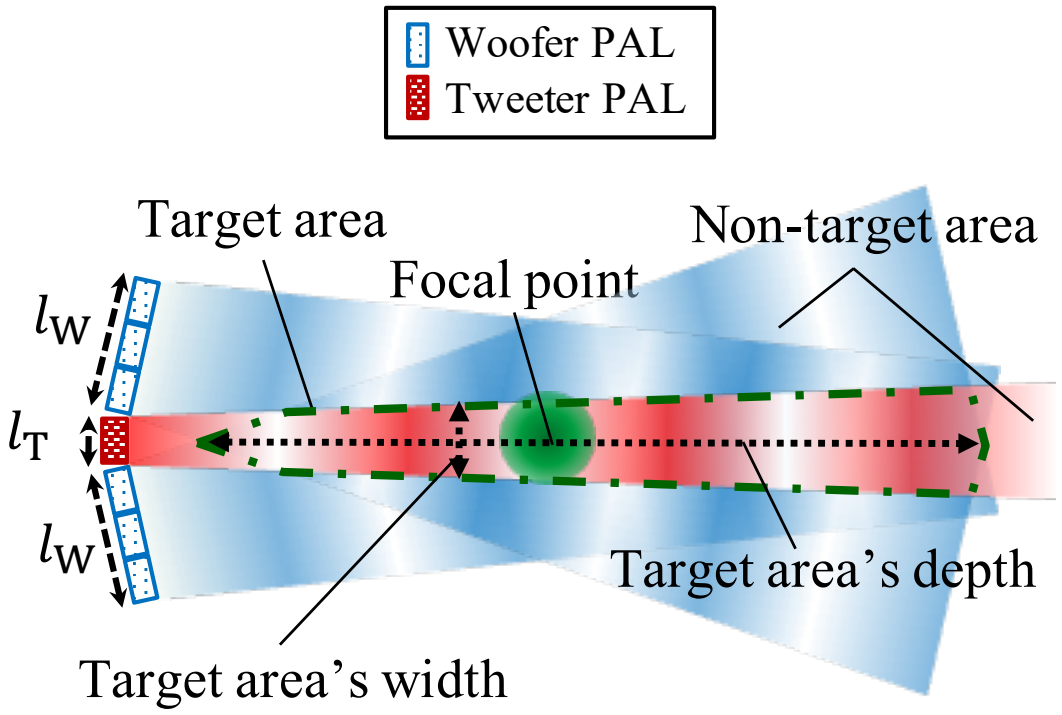


Figure 3.4: Arrangement of Woofer PAL and Tweeter PAL in the proposed method.

Based on the calculated directional angles, the arrangement of Woofer PAL and Tweeter PAL can be determined as shown in Fig. 3.4. As same as shown in Fig. 3.3, the acoustic beams of Woofer PAL and Tweeter PAL are approximated to trapezoids. The focal point in Fig. 3.4 is defined as the intersection point of the propagation axes. It should be mentioned that the focal point is determined based on the directivity only, so there is no necessity to overlap the focal point with the maximum demodulation point of PAL. The area where the acoustic beams of Woofer PAL and Tweeter PAL overlap is defined as the target area because both the low-frequency and the high-frequency can be heard simultaneously. In the target area, flat frequency response is expected to be realized with the multi-way structure. It should be noticed that both the position of the focal point and the target area can be controlled by adjusting the angle between Woofer PAL and Tweeter PAL.

Geometric denotations are shown in Fig. 3.5. Because the Woofer PAL is arranged symmetrically, only one side is shown in this figure. The center point of Tweeter PAL is defined as point O, and the center point of Woofer PAL is defined as point P. The propagation axis of each PAL is perpendicular to the emitter, and the intersection point of the propagation axes is defined

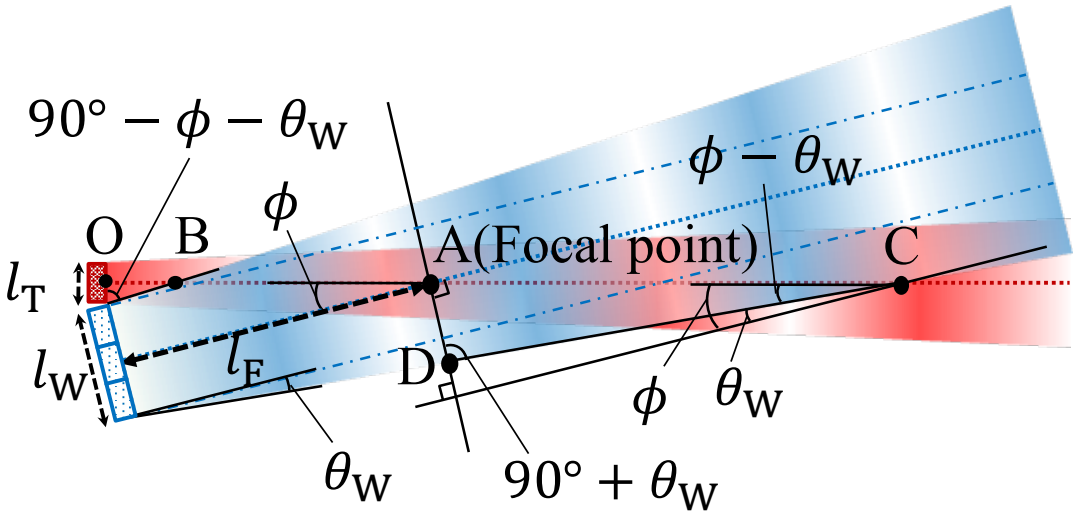


Figure 3.5: Geometric denotations in the arrangement of Woofer PAL and Tweeter PAL.

as the focal point, denoted as point A. The distance from the focal point A to the center point of Woofer PAL is equal to the distance from the focal point A to the center point of Tweeter PAL ($l_{\mathrm{OA}} = l_{\mathrm{PA}} = l_{\mathrm{F}}$). The propagation axis of Tweeter PAL intersects with the boundary of Woofer PAL's acoustic beam at points B and C. It can be noticed that the area between points B and C is the target area on the propagation axis. Line segment AD is perpendicular to line segment PA and point D is on the boundary of Woofer PAL's acoustic beam.

The angle between Woofer PAL and Tweeter PAL can be calculated with the width of the emitters and the distance from the emitters to the focal point, which can be expressed as:

$$\phi = \tan^{-1}\left(\frac{l_{\mathrm{W}}}{2l_{\mathrm{F}}}\right) + \tan^{-1}\left(\frac{l_{\mathrm{T}}}{2l_{\mathrm{F}}}\right). \tag{3.4}$$

The length of line segment OB can be calculated as follows:

$$l_{\mathrm{OB}} = \frac{l_{\mathrm{T}}}{2}\tan(90^\circ - \phi - \theta_{\mathrm{W}}). \tag{3.5}$$

The length of line segment AD can be calculated as follows:

$$l_{\mathrm{AD}} = \frac{l_{\mathrm{W}}}{2} + l_{\mathrm{F}}\tan\left(\theta_{\mathrm{W}}\right). \tag{3.6}$$

The length of line segment AC can be calculated by utilizing the law of sines in triangle ACD, which can be expressed as:

$$l_{\mathrm{AC}} = \left(\frac{l_{\mathrm{W}}}{2} + l_{\mathrm{F}}\tan\theta_{\mathrm{W}}\right)\frac{\sin\left(90^\circ + \theta_{\mathrm{W}}\right)}{\sin\left(\phi - \theta_{\mathrm{W}}\right)}. \tag{3.7}$$

It should be noticed that Eq. (3.7) holds only if $\phi > \theta_{\mathrm{W}}$. In the case that $\phi \leq \theta_{\mathrm{W}}$, point C exists at an infinite distance so that the acoustic beams of Woofer PAL and Tweeter PAL always overlap and no longer separate in theory.

3.2.3 Modulation Methods

Modulation methods for PAL has been introduced in Section 2.3.1, therefore, in this section, the thesis reviews a part of them briefly. DSB modulation is the simplest modulation method to implement. Defining $s(t)$ as a target signal to reproduce, the DSB-modulated wave can be expressed as:

$$\mathrm{DSB}[s(t)] = (1 + m_{\mathrm{DSB}} s(t))c(t). \tag{3.8}$$

As discussed in Chapter 2, DSB modulation causes significant harmonic distortion because it utilizes two sidebands.

M-DSB modulation [50] has been proposed to enhance the low-frequency by utilizing harmonic distortion. It is a combination of DSB modulation and DSB-SC modulation. The modulated wave can be expressed as:

$$\text{M-DSB}[s(t)] = \left(1 + m_{\mathrm{DSB}} s(t) + m_{\mathrm{SC}} \mathcal{C}_{\frac{1}{2}}[s(t)]\right) c(t), \tag{3.9}$$

where $\mathcal{C}_{\frac{1}{2}}[\cdot]$ denotes an operation which compress the frequency of a signal to half. In the previous works [112, 113], M-DSB has been validated effective to enhance the low-frequency when reproducing wideband signals. However, it is reported that the sound quality is degraded when reproducing single tones or harmonic signals due to the effect on pitch perception.

SSB modulation only employs one sideband so the harmonic distortion is expected to be reduced. SSB modulation can be categorized into LSB modulation and USB modulation, according to which sideband it utilizes. The LSB-modulated wave and USB-modulated wave can be expressed as:

$$\mathrm{LSB}[s(t)] = c(t) + m_{\mathrm{SSB}}(s(t)c(t) + \hat{s}(t)\hat{c}(t)), \tag{3.10}$$

$$\mathrm{USB}[s(t)] = c(t) + m_{\mathrm{SSB}}(s(t)c(t) - \hat{s}(t)\hat{c}(t)). \tag{3.11}$$

Experimental results suggest that the modulation method with the highest sound pressure may vary as the frequency of target signal [40, 48, 114]. It is reported that LSB modulation tends to show better performance in low-frequency while USB modulation tends to show better

performance in high-frequency [48,114]. Facing this fact, a considerable approach to achieve a high sound pressure is to combine the modulation methods with the highest sound pressure in each frequency range. To balance the characteristics of LSB modulation and USB modulation for Tweeter PAL, this thesis proposes a novel method which is named combined SSB (C-SSB) modulation. The concept of C-SSB modulation is shown in Fig. 3.6.

The modulated wave with C-SSB modulation can be expressed as:

$$\text{C-SSB}[s(t)] = c(t) + m_{\text{SSB}}\{s_{\text{L}}(t)c(t) + \hat{s}_{\text{L}}(t)\hat{c}(t)\} + m_{\text{SSB}}(s_{\text{H}}(t)c(t) - \hat{s}_{\text{H}}(t)\hat{c}(t)), \quad (3.12)$$

$$s_{\text{L}}(t) = s(t) * h_{\text{LPF}_{f_{\text{C-SSB}}}}(t), \quad (3.13)$$

$$s_{\text{H}}(t) = s(t) * h_{\text{HPF}_{f_{\text{C-SSB}}}}(t), \quad (3.14)$$

where $h_{\text{LPF}_{f_{\text{C-SSB}}}}(t)$ denotes the coefficients of a low-pass filter with cutoff frequency $f_{\text{C-SSB}}$, $h_{\text{HPF}_{f_{\text{C-SSB}}}}(t)$ denotes the coefficients of a high-pass filter with cutoff frequency $f_{\text{C-SSB}}$.

It can be noticed from Eq. (3.12) that LSB modulation is employed in low-frequency while USB modulation is employed in high-frequency in C-SSB modulation. It is ideal to set the crossover frequency $f_{\text{C-SSB}}$ at the frequency that the frequency response of LSB modulation crosses with that of USB modulation. In practice, the demodulated frequency response of each modulation method depends on the characteristic of the ultrasonic transducers. Therefore, preliminary experiments should be carried out to determine the crossover frequency $f_{\text{C-SSB}}$.

Based on the discussion above, C-SSB modulation is employed for Tweeter PAL. The emitting signal from Tweeter PAL can be expressed as:

$$u_{\text{T}}(t) = \text{C-SSB}[x_{\text{T}}(t)]. \quad (3.15)$$

For Woofer PAL which reproduces the low-frequency sound, the choice of modulation method is dependent on the feature of the target signal. When the target signal is a wideband signal such as Gaussian noise, M-DSB modulation is employed. When the target signal is a harmonic signal, LSB modulation is employed. The emitting signal from Woofer PAL can be expressed as:

$$u_{\text{W}}(t) = \begin{cases} \text{LSB}[x_{\text{W}}(t)], \text{if } x_{\text{W}}(t) \text{ is harmonic,} \\ \text{M-DSB}[x_{\text{W}}(t)], \text{if } x_{\text{W}}(t) \text{ is noise-like.} \end{cases} \quad (3.16)$$

The demodulated sounds of Woofer PAL and Tweeter PAL are expected to form a target area where the acoustic beams overlap as shown in Fig. 3.4. The observed demodulated sound in the

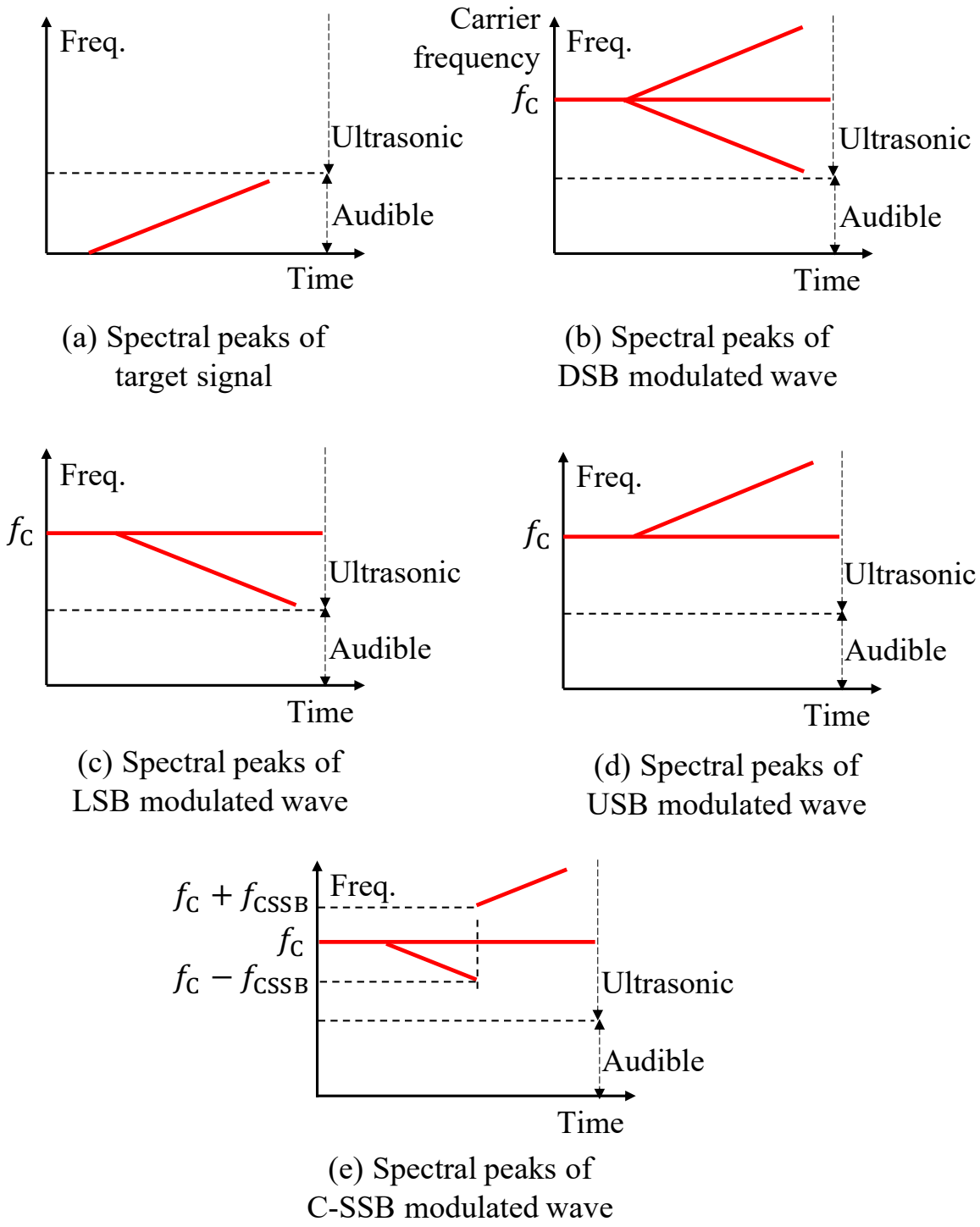


Figure 3.6: Concept of C-SSB modulation.

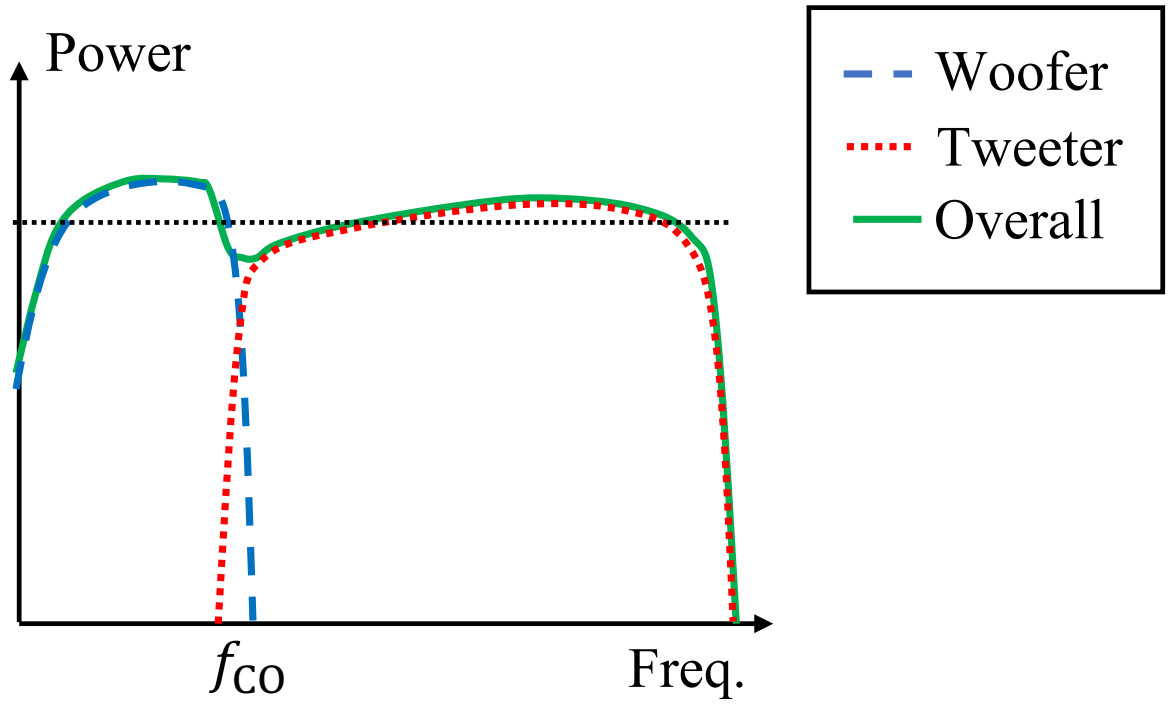


Figure 3.7: Frequency responses of demodulated sound at the focal point.

target area can be expressed as:

$$x_{\mathrm{dem}}(t) = \mathcal{D}_{d_{\mathrm{W}}}[u_{\mathrm{W}}(t)] + \mathcal{D}_{d_{\mathrm{T}}}[u_{\mathrm{T}}(t)]. \tag{3.17}$$

It should be noticed that the self-demodulation is a complicated nonlinear process, so it is denoted as $\mathcal{D}_{\cdot}[\cdot]$ in the Eq. (3.17). The on-axis sound pressure can be approximated as a distance-dependent value [115]. Here, d_{W} denotes the distance from the observation point to Woofer PAL, and d_{T} denotes the distance from the observation point to Tweeter PAL. As discussed, the low-frequency sound is expected to be enhanced by enlarging the Woofer PAL and employing M-DSB modulation which utilizes distortion components. As a result, the flat frequency response can be realized in the target area, especially at the focal point as shown in Fig. 3.7.

3.3 Developed Prototype Multi-way PAL

To demonstrate the effectiveness of the proposed method, a prototype multi-way PAL has been developed. A photograph of the developed prototype is shown in Fig. 3.8. This prototype consists of 7 PAL units: 1 unit in the center for Tweeter PAL, and 6 units on the sides for Woofer PAL. Each PAL unit is composed of 200 ultrasonic transducers (MURATA, MA40S4S). The 200 transducers in each PAL unit are arranged in 20 rows and each row includes 10 transducers. It is reported that a positive correlation can be observed between the demodulated sound pressure and the number of transducers [42, 112]. Based on the results of preliminary experiments, the transducers of Woofer PAL are set 6 times that of Tweeter PAL. In summary, Woofer PAL consists of 1200 ultrasonic transducers, while Tweeter PAL consists of 200 ultrasonic transducers. The sound pressure in high-frequency by Tweeter PAL is promising because it holds a similar scale with a commercial PAL (MEE, MSP-50E), which contains 207 ultrasonic transducers.

The joint between PAL units is flexible, therefore, the angle can be set to any value. As discussed, the Woofer PAL is configured in a straight line, and the angle between Woofer PAL and Tweeter PAL is set to ϕ. The measured width of Tweeter PAL l_{T} is 0.12 m, while the measured width of Woofer PAL l_{W} is 0.36 m. Preliminary experiments have been carried out to measure the directional angles θ_{T} and θ_{W}. The distance d is set as 2.0 m to measure the beamwidth. When reproducing an audible band (0–20 kHz) Gaussian noise, the directional angle of Tweeter PAL is 2.3 deg.. When reproducing a Gaussian noise in 1–20 kHz, the directional angle of Tweeter

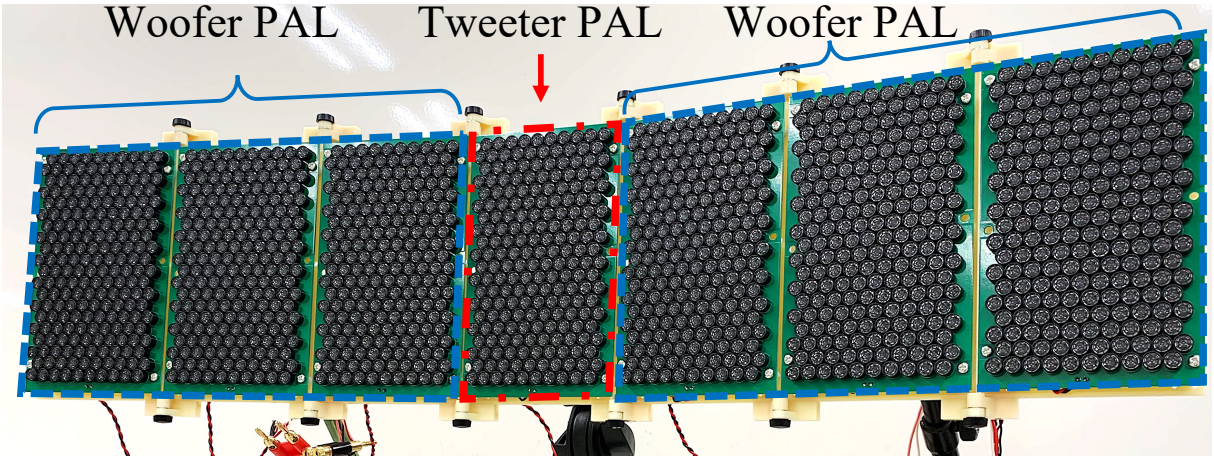


Figure 3.8: Prototype of multi-way parametric array loudspeaker.

PAL is 2.6 deg.. When reproducing an audible band (0–20 kHz) Gaussian noise, the directional angle of Woofer PAL is 1.4 deg.. When reproducing a Gaussian noise in 0–1 kHz, the directional angle of Tweeter PAL is 7.3 deg.. In the proposed method, Woofer PAL is utilized to reproduce low-frequency below 1 kHz, and Tweeter PAL is utilized to reproduce high-frequency above 1 kHz. Therefore, the conditions that $\theta_{\mathrm{T}} = 2.3$ deg. and $\theta_{\mathrm{W}} = 7.3$ deg. will be substituted in the following evaluation experiments.

The distance to focal point l_{F} is set as 2.0 m. Substituting the widths and the directional angles of Woofer PAL and Tweeter PAL to Eq. (3.4), it can be calculated that $\phi = 6.9$ deg.. Moreover, substituting these values to Eqs. (3.5) and (3.6), it can be calculated that $l_{\mathrm{OB}} = 0.25$ m and $l_{\mathrm{AD}} = 0.44$ m. In the area between point O and point B, the acoustic beams do not overlap, so this area is regarded as a non-target area, as shown in Fig. 3.5.

3.4 Evaluation Experiments

Evaluation experiments have been carried out on the prototype introduced in Section 3.3, in order to demonstrate the effectiveness of the proposed method. In Section 3.4.1, the demodulated sounds' directivity characteristics of Woofer PAL and Tweeter PAL are evaluated. In Section 3.4.2, the frequency responses of demodulated sounds for Woofer PAL and Tweeter PAL under each modulation method are evaluated. In Section 3.4.3, the Woofer PAL and Tweeter PAL are combined into the proposed multi-way PAL, and then the frequency responses of demodulated sounds in the target area and non-target area are evaluated.

3.4.1 Experiment on the Directivity Characteristics of Woofer and Tweeter

In this Section, the directivity characteristics of demodulated sound are evaluated under the conditions that Woofer PAL or Tweeter PAL is driven individually or simultaneously. The experimental conditions for this experiment are listed in Table 3.1. The experimental equipment for this experiment is listed in Table 3.2.

White Gaussian noise in the audible band (0–20 kHz) is utilized as a sound source. To demonstrate the necessity of amplitude normalization, 2 comparison patterns have been prepared: the conventional method means that the sound source is processed by band division without amplitude normalization, and the proposed method means that the sound source is processed by both the band division and the amplitude normalization. In the case of measuring the directivity characteristics of Woofer PAL's demodulated sound, the sound source is processed by a low-pass filter as described in Section 3.2. The sound source with/without amplitude normalization is modulated with LSB modulation as discussed in Section 3.2.3, and then the modulated wave is emitted from Woofer PAL. In the case of measuring the directivity characteristics of Tweeter PAL's demodulated sound, the sound source is processed by a high-pass filter as described in Section 3.2. The sound source with/without amplitude normalization is modulated with C-SSB modulation as discussed in Section 3.2.3, and then the modulated wave is emitted from Tweeter PAL. It should be mentioned here that the AM is implemented with digital signal processing. The modulated wave is converted into an analog signal with a D/A converter, and amplified with a power amplifier before being emitted from PAL.

The experimental arrangement is shown in Fig. 3.9. Microphones are set at 17 positions on

Table 3.1: Experimental conditions for directivity characteristics.

Ambient noise level	L_{A} = 27.3 dB
Sampling frequency	192 kHz
Quantization	32 bits
Environment	Office room (T_{60} = 0.65 s)
Sound source	White Gaussian noise (0–1, 1–20 kHz)
Driving voltage	8 V (rms)
Modulation methods	LSB (Woofer PAL) C-SSB (Tweeter PAL / $f_{\mathrm{C\text{-}SSB}}$ = 5 kHz)
Carrier frequency	f_{C} = 40 kHz
Modulation factor	m_{SSB} = 1.0

Table 3.2: Experimental equipment for directivity characteristics.

Ultrasonic transducer	MURATA, MA40S4S (1400 pcs.)
Power amplifier	VICTOR, PS-A2002
A/D, D/A converter	RME, FIREFACE UFX
Microphone amplifier	THINKNET, MA2016C
Microphone	SONY, ECM-88B

the horizontal axis with an interval of 0.1m, and they are set at 7 positions on the vertical axis with an interval of 0.5 m. The center of Tweeter PAL, point O in Fig. 3.5, is set at (0 m, 0 m) in Fig. 3.9. Woofer PAL and Tweeter PAL are arranged as shown in Fig. 3.4. It can be noticed that the microphones are set at y = 0.25 m instead of y = 0 m due to the size of the prototype multi-way PAL.

The recorded sounds observed at observation points are processed with a low-pass filter to filter out the ultrasonic components so that the demodulated sound pressure can be calculated with an acoustic calibrator. The calculation of demodulated sound pressure can be expressed as:

$$L_{\mathrm{ob}} = L_{\mathrm{cali}} - 10\log_{10}\left(\frac{\sigma_{\mathrm{ob}}^2}{\sigma_{\mathrm{cali}}^2}\right), \tag{3.18}$$

where L_{ob} denotes the observed sound pressure level of demodulated sound, L_{cali} denotes the sound pressure level of the acoustic calibrator, σ_{ob} denotes the standard deviation of the measured

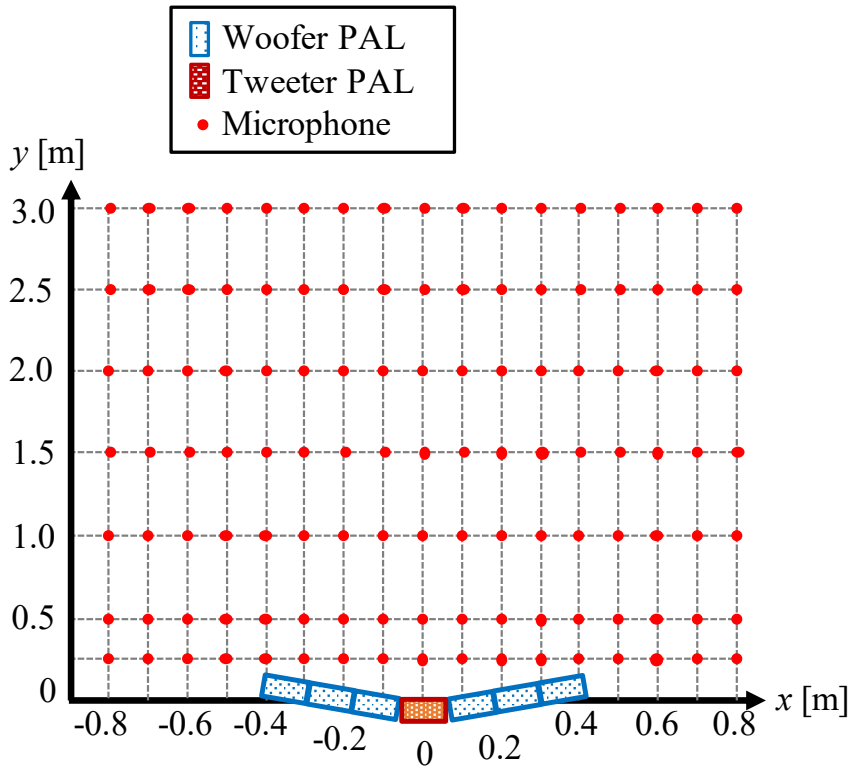
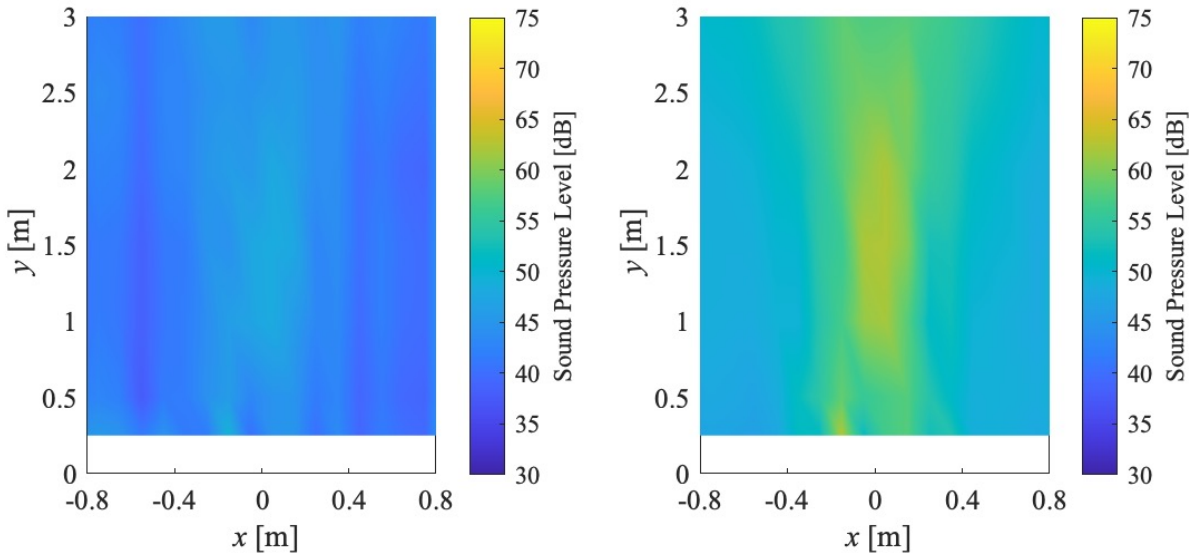


Figure 3.9: Experimental arrangement for measurement of sound pressure distribution.

signal of demodulated sound, σ_{cali} denotes the standard deviation of the measured signal of acoustic calibrator. Since the sound pressure level of acoustic calibrator L_{cali} is known, the sound pressure level of demodulated sound can be calculated.

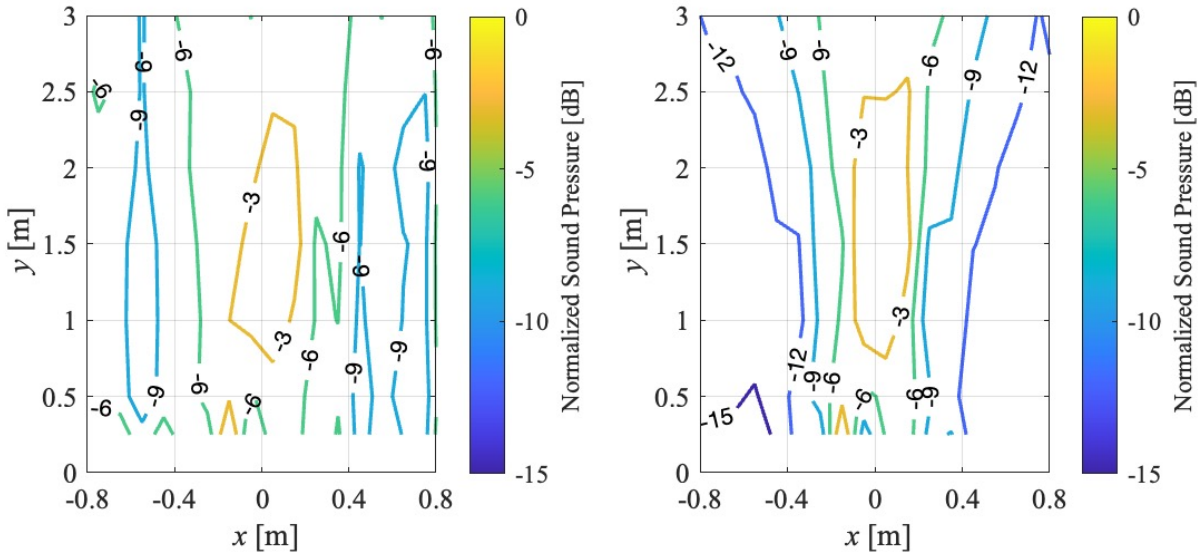
The sound pressure distribution of Woofer PAL is shown in Fig. 3.10, where (a) shows the result under the condition that the amplitude normalization was not included, and (b) shows the result under the condition that the amplitude normalization was included. It can be confirmed that the demodulated sound pressure increases by over 10 dB when the amplitude normalization was included. Therefore, the effectiveness of amplitude normalization discussed in Section 3.2 can be validated. The contour lines of the sound pressure distribution for Woofer PAL are shown in Fig. 3.11. The highest sound pressure level observed in each condition is normalized to 0 dB, and the areas within -3, -6, -9, -12, -15 dB are shown. It can be confirmed from Fig. 3.11 (a) that the contour lines show a sparse and disordered tendency because the demodulated sound pressure is low. In contrast, it can be confirmed from Fig. 3.11 (b) that the acoustic beam is generated narrow while the frequency is low. This can be explained by that the demodulated sound pressure significantly increased due to the amplitude normalization.



(a) Conventional method (w/o amplitude normalization)

(b) Proposed method (w/ amplitude normalization)

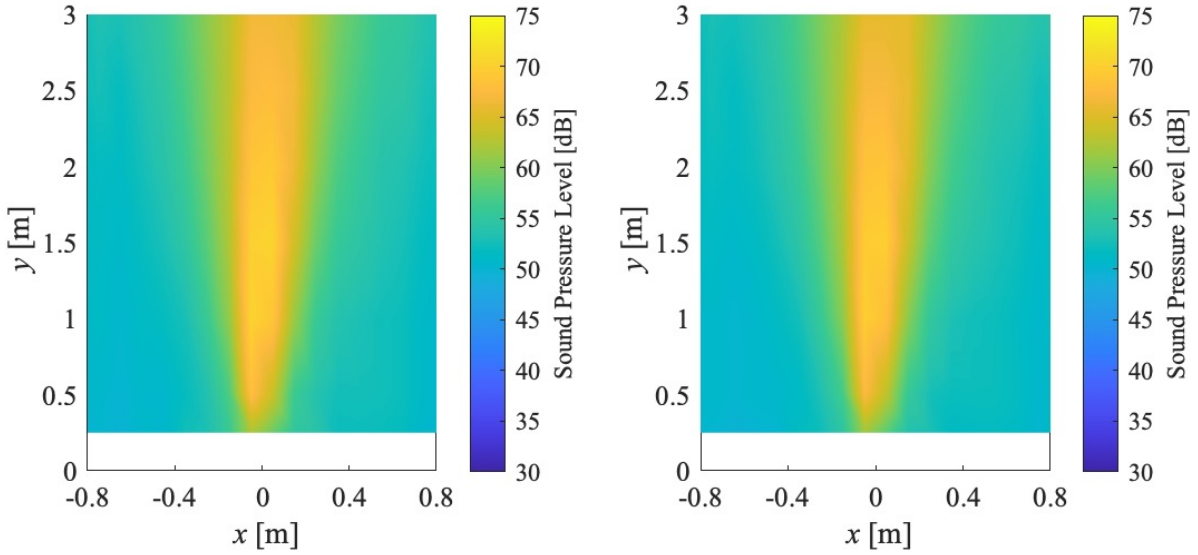
Figure 3.10: Sound pressure distribution for Woofer PAL (0–1 kHz).



(a) Conventional method (w/o amplitude normalization)

(b) Proposed method (w/ amplitude normalization)

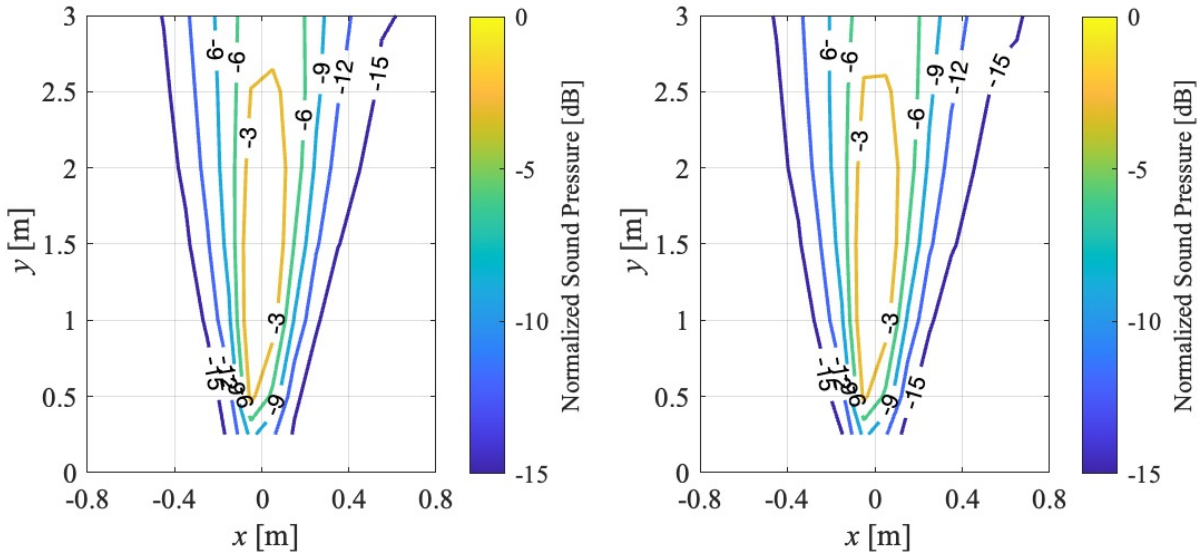
Figure 3.11: Contour lines of sound pressure distribution for Woofer PAL (0–1 kHz).



(a) Conventional method (w/o amplitude normalization)

(b) Proposed method (w/ amplitude normalization)

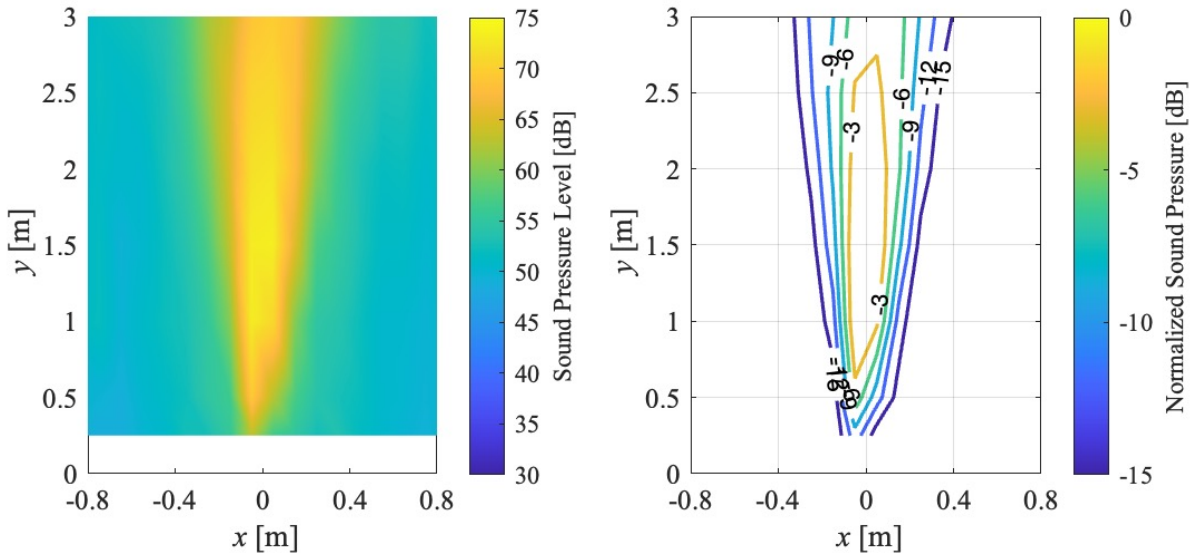
Figure 3.12: Sound pressure distribution for Tweeter PAL (1–20 kHz).



(a) Conventional method (w/o amplitude normalization)

(b) Proposed method (w/ amplitude normalization)

Figure 3.13: Contour lines of sound pressure distribution for Tweeter PAL (1–20 kHz).



(a) Sound pressure distribution

(b) Contour lines of sound pressure distribution

Figure 3.14: Directivity characteristics for the proposed method (driving Woofer PAL and Tweeter PAL simultaneously).

The sound pressure distribution of Tweeter PAL is shown in Fig. 3.12, and the contour lines of the sound pressure distribution for Tweeter PAL are shown in Fig. 3.13. Since the most of energy in demodulated sound distributes in the high-frequency, there is little difference between the results in Fig. 3.12 (a) and Fig. 3.12 (b). In both Fig. 3.12 and Fig. 3.13, typical narrow acoustic beams can be observed.

Moreover, directivity characteristics for the proposed method, in which Woofer PAL and Tweeter PAL are driven simultaneously, are shown in Fig. 3.14. The sound pressure distribution is shown in Fig. 3.14 (a), and its contour lines are shown in Fig. 3.14 (b). It can be confirmed from Fig. 3.14 (a) that the on-axis demodulated sound pressure of the proposed multi-way PAL is higher than the individual Tweeter PAL. It can also be confirmed from Fig. 3.14 (b) that the proposed multi-way PAL shows a narrower acoustic beam over individual Tweeter PAL. These phenomena can be explained by that the carrier wave emitted from Woofer PAL contributed to the demodulated sound of Tweeter PAL in the area where the acoustic beams overlap. Similarly, it is considerable that the carrier wave emitted from Tweeter PAL contributed to the demodulated sound of Woofer PAL. Therefore, the demodulated sound of the proposed multi-way PAL is

expected to achieve a higher sound pressure and a sharper directivity characteristic than that of individual Woofer PAL or Tweeter PAL.

From the results above, it is validated that Woofer PAL and Tweeter PAL can form a narrow acoustic beam in low-frequency and high-frequency respectively. In addition, it is demonstrated that the sound pressure level in low-frequency can be emphasized by the amplitude normalization in the proposed multi-way PAL. Specifically, a narrow acoustic beam can be formed with a sound pressure level at 60 dB in the center and a -6 dB beamwidth at 0.4 m when reproducing a white Gaussian noise below 1 kHz, which shows a remarkable significance in practical use.

3.4.2 Experiment on the Frequency Response of Woofer and Tweeter under Modulation Methods

In this Section, the frequency responses of demodulated sound are evaluated under each modulation method for Woofer PAL and Tweeter PAL individually. The experimental conditions for this experiment are listed in Table 3.3. The experimental equipment for this experiment is listed in Table 3.4.

The frequency response of demodulated sound can be measured from the impulse response. Time-stretched pulse (TSP) [116] signal has been proposed to calculate the impulse response by convolution on the observed signal and a dry source of reverse TSP signal. In the impulse response calculated by the TSP signal, the direct sound and reflected sound can be separated in the time domain, however, the harmonic distortion cannot be separated. Logarithmic TSP (log-TSP) [117, 118] signal has been proposed to measure the impulse response without being affected by harmonic distortions. In the impulse response calculated by log-TSP signal, the fundamental frequency and harmonic distortions can be separated in the time domain. Therefore, a log-TSP signal in the audible band (0–20 kHz) is utilized to measure the frequency responses of demodulated sound excluding the harmonic distortions. The result can be used to evaluate the demodulated sound when reproducing pure tones or harmonic sounds. In addition, white Gaussian noise in the audible band (0–20 kHz) is utilized to measure the frequency responses of demodulated sound including the harmonic distortions. The result can be used to evaluate the demodulated sound when reproducing wideband signals.

For Woofer PAL, the band of the sound source is limited to 0–1 kHz with a low-pass filter.

Table 3.3: Experimental conditions for frequency responses of modulation methods.

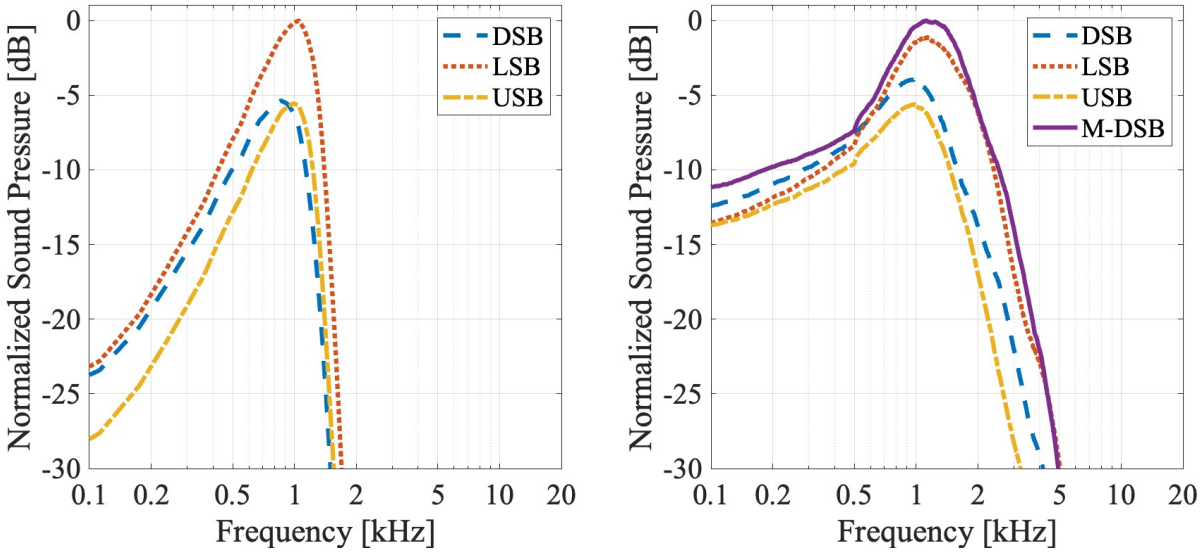
Ambient noise level	L_{A} = 27.3 dB
Sampling frequency	192 kHz
Quantization	32 bits
Environment	Office room (T_{60} = 0.65 s)
Driving voltage	8 V (rms)
Sound source	Log-TSP signal (0–20 kHz) White noise (0–20 kHz)
Modulation methods	DSB, LSB, USB, C-SSB, M-DSB
Carrier frequency	f_{C} = 40 kHz
Cutoff frequency	$f_{\text{C-SSB}}$ = 5 kHz (for C-SSB)
Modulation factor	m_{SSB} = 1.0, m_{DSB} =1.0, m_{SC} = 1. 0

Table 3.4: Experimental equipment for frequency responses of modulation methods.

Ultrasonic transducer	MURATA, MA40S4S (1400 pcs.)
Power amplifier	VICTOR, PS-A2002
A/D, D/A converter	RME, FIREFACE UFX
Microphone amplifier	THINKNET, MA2016C
Microphone	SENNHEISER, MKH-8020

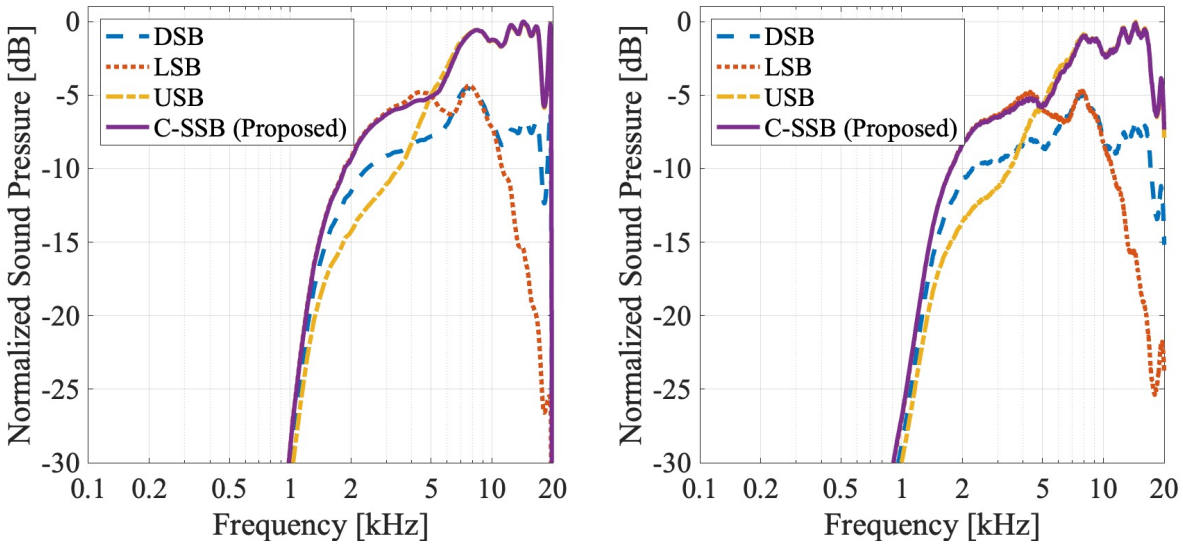
DSB modulation, LSB modulation, and USB modulation are utilized when the sound source is a log-TSP signal. In the case that the sound source is white Gaussian noise, M-DSB modulation is also utilized, as discussed in Section 3.2.3. For Tweeter PAL, the band of the sound source is limited to 1–20 kHz with a high-pass filter. DSB modulation, LSB modulation, USB modulation, and C-SSB modulation are utilized for both log-TSP signal and white Gaussian noise. It is known from preliminary experiments that the frequency response of LSB modulation crosses with that of USB modulation at about 5 kHz, therefore, the crossover frequency of C-SSB modulation $f_{\text{C-SSB}}$ is set at 5 kHz. The demodulated sounds of Woofer PAL and Tweeter PAL under each condition of modulation method are recorded at (0 m, 2.0 m) in Fig. 3.9, which is the same position as focal point A in Fig. 3.5.

The measure frequency responses of Woofer PAL and Tweeter PAL under each condition of



(a) Log-TSP signal (b) White noise

Figure 3.15: Measured frequency responses of Woofer PAL.



(a) Log-TSP signal (b) White noise

Figure 3.16: Measured frequency responses of Tweeter PAL.

modulation method are shown in Figs. 3.15 and 3.16. The results measured with log-TSP signal are shown in Fig. 3.15 (a) and Fig. 3.16 (a), while the results measured with white Gaussian noise are shown in Fig. 3.15 (b) and Fig. 3.16 (b). Since the frequency range for Woofer PAL is narrow, the horizontal axes are set in a logarithmic scale. It can be noticed from Fig. 3.15 (a) that LSB modulation shows the highest sound pressure, and the sound pressure of DSB modulation is between that of LSB modulation and that of USB modulation. It can be noticed from Fig. 3.15 (b) that M-DSB modulation shows higher sound pressure over other modulation methods in low-frequency. It should be noticed that the frequency responses shown in Fig. 3.15 (a) include no harmonic distortion. However, the frequency responses shown in Fig. 3.15 (b) include all harmonic distortions. M-DSB modulation utilizes harmonic distortion to enhance the low-frequency as discussed, and a corresponding conclusion can be drawn from Fig. 3.16 (b). DSB modulation tends to show much more harmonic distortion than LSB modulation and USB modulation, which can explain why the sound pressure of DSB modulation is higher in the frequency band of 0–0.5 kHz in Fig. 3.16 (b). The same tendency can be confirmed in Fig. 3.16 (a) and Fig. 3.16 (b). The frequency response of C-SSB modulation almost overlaps that of LSB modulation in 1–5 kHz and it almost overlaps that of USB modulation in 5–20 kHz. Therefore, it is demonstrated that the proposed C-SSB modulation can balance the advantages of LSB modulation and USB modulation. With the result above, the discussion of modulation methods in Section 3.2.3 has been validated.

3.4.3 Experiment on the Frequency Response of the Proposed Multi-way Parametric Array Loudspeaker

In this section, Woofer PAL and Tweeter PAL are combined as the proposed multi-way PAL, and the frequency response of the whole proposed method is evaluated. The experimental conditions for this experiment are listed in Table 3.5. The experimental equipment for this experiment is listed in Table 3.6.

As described in Section 3.2 and Section 3.3, Woofer PAL contains 1200 ultrasonic transducers, and the low-frequency band of the sound source is emitted after being modulated with LSB modulation or M-DSB modulation. Tweeter PAL contains 200 ultrasonic transducers, and the high-frequency band of the sound source is emitted after being modulated with C-SSB modula-

Table 3.5: Experimental conditions for frequency responses of the proposed method.

Ambient noise level	L_{A} = 27.3 dB
Sampling frequency	192 kHz
Quantization	32 bits
Environment	Office room (T_{60} = 0.65 s)
Driving voltage	8 V (rms)
Sound source	Log-TSP signal (0–20 kHz) White noise (0–20 kHz)
Modulation methods	LSB, USB, C-SSB, M-DSB
Carrier frequency	f_{C} = 40 kHz
Cutoff frequency	$f_{\text{C-SSB}}$ = 5 kHz (for C-SSB)
Modulation factor	m_{SSB} = 1.0, m_{DSB} =1.0, m_{SC} = 1. 0

Table 3.6: Experimental equipment for frequency responses of proposed method.

Ultrasonic transducer	MURATA, MA40S4S (1400 pcs.)
Power amplifier	VICTOR, PS-A2002
A/D, D/A converter	RME, FIREFACE UFX
Microphone amplifier	THINKNET, MA2016C
Microphone	SENNHEISER, MKH-8020

tion. The proposed method is compared with 4 patterns of conventional full band reproduction: Conv. (LSB, 200), Conv. (USB, 200), Conv. (LSB, 1400), and Conv. (USB, 1400). Here, “LSB” or “USB” suggests the modulation method, and “200” or “1400” indicates the number of employed ultrasonic transducers. Based on the experimental results in Section 3.4.2, the frequency responses of DSB modulation and C-SSB modulation can be predicted from the frequency responses of LSB modulation and USB modulation. Therefore, only LSB modulation and USB modulation are utilized in this experiment. It should be noticed that in the case of 200 ultrasonic transducers, the PAL unit for Tweeter PAL is utilized; in the case of 1400 ultrasonic transducers, all PAL units are utilized. Though the reproduction frequency range is different from the proposed method, the arrangement is the same as shown in Fig. 3.9.

The demodulated sounds are observed at 6 positions in Fig. 3.9: (0 m, 1.0 m) as the focal

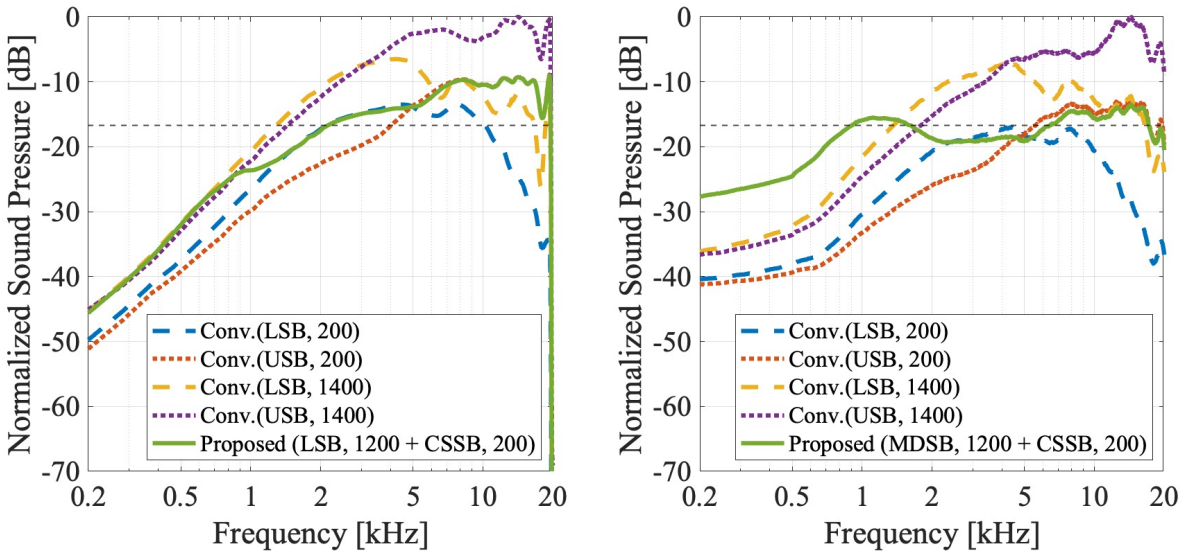
point, (0 m, 1.0 m) and (0 m, 3.0 m) in the target area, (0 m, 0.25 m) on the boundary of target area, (-0.5 m, 1.0 m) and (-0.5 m, 2.0 m) in the non-target area. The measured frequency responses at each position are shown in Figs. 3.17 – 3.22. The frequency responses measured with log-TSP signal are shown in Figs. 3.17 (a) – 3.22 (a), and the frequency responses measured with white Gaussian noise are shown in Figs. 3.17 (b) – 3.22 (b).

As discussed in Section 3.4.2, the frequency responses measured with log-TSP signal include no harmonic distortion, while the frequency responses measured with white Gaussian noise include all harmonic distortions. Therefore, it can be confirmed from Figs. 3.17 (a) – 3.22 (a) that the proposed method shows similar tendency with Conv. (LSB, 1400) and Conv. (USB, 1400) in the low-frequency. However, in Figs. 3.17 (b) – 3.22 (b), the proposed method shows a significant improvement over Conv. (LSB, 1400) and Conv. (USB, 1400) in the low-frequency. One reason is that the amplitude normalization contributes to the demodulated sound pressure of wideband signal as shown in Fig. 3.2. Another reason is that the harmonic distortions are utilized in M-DSB modulation to enhance the low-frequency.

In the target area, as shown in Figs. 3.17 – 3.19, the proposed method shows a sound pressure equal to or higher than that of conventional full band reproduction employing 1400 ultrasonic transducers in low-frequency. Especially in the case of reproducing white Gaussian noise, the proposed method increases the sound pressure below 1 kHz by about 10 dB. In high-frequency above 1 kHz, the proposed method achieves an equal level with conventional full band reproduction employing 200 ultrasonic transducers. Therefore, it is demonstrated that in the target area, the proposed method can balance the sound pressure in low-frequency and high-frequency and achieve a flatter frequency response in wideband.

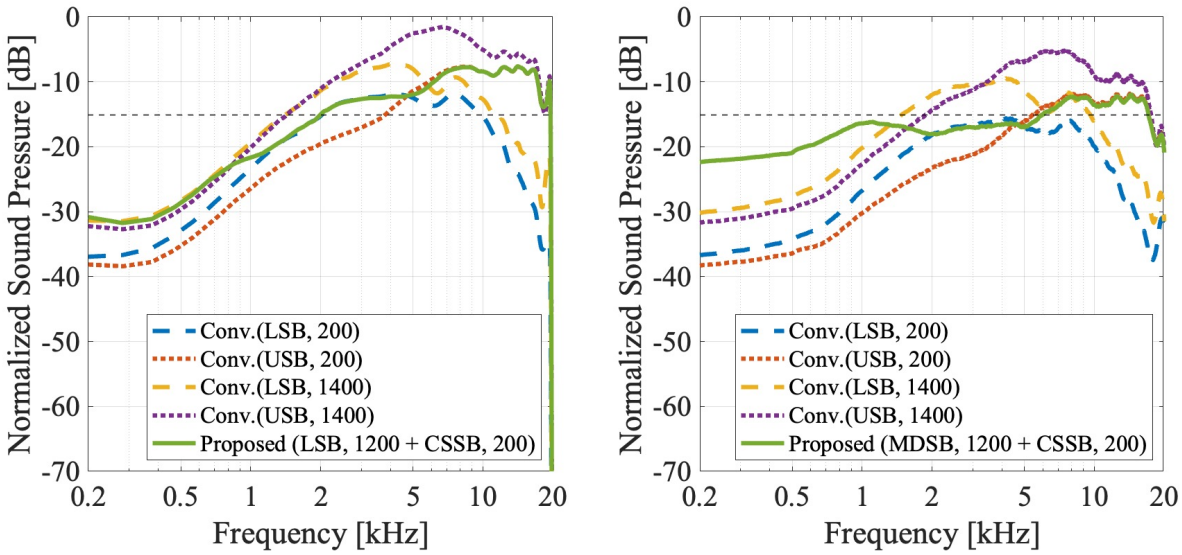
On the boundary of the target area, as shown in Fig. 3.20, the proposed method can achieve a flatter frequency response than conventional full band reproduction. In the non-target area, as shown in Figs. 3.21 and 3.22, the demodulated sound pressure decreases by about 10 dB in low-frequency and about 20 dB in high-frequency, compared with the those observed in the target area. This result corresponds to the results shown in Figs. 3.11, 3.13 and 3.14 (b).

To evaluate if the flat frequency response is realized by the proposed method, the flatness of frequency responses measured in the target area are shown numerically. Root mean square error (RMSE) between measured frequency response and flat frequency response is calculated for evaluation. The results of RMSE between measured and desired flat frequency responses



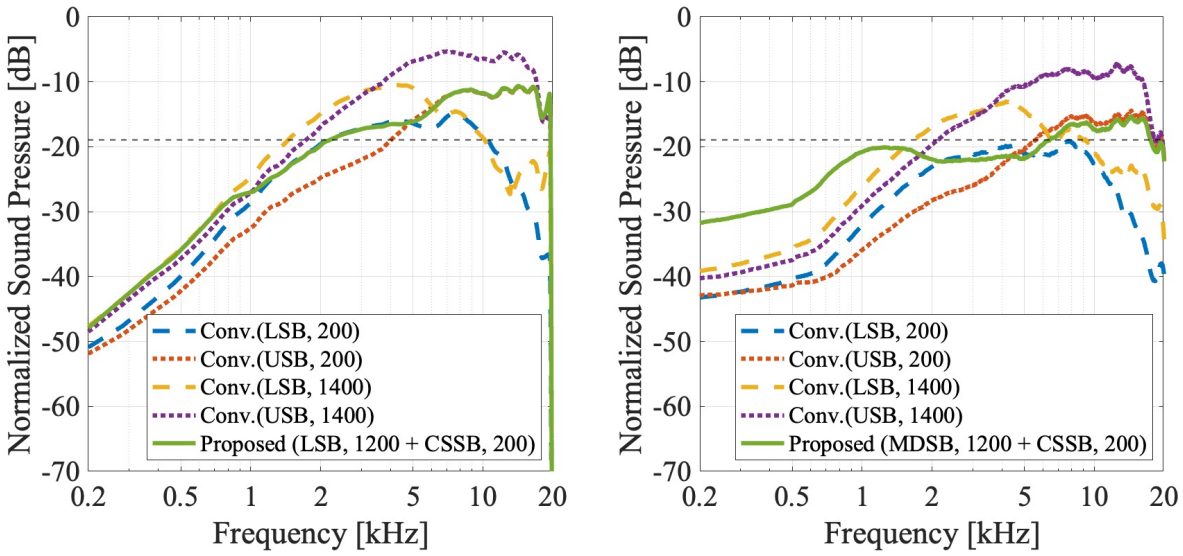
(a) Log-TSP signal (b) White noise

Figure 3.17: Measured frequency responses at the focal point (0 m, 2.0 m).



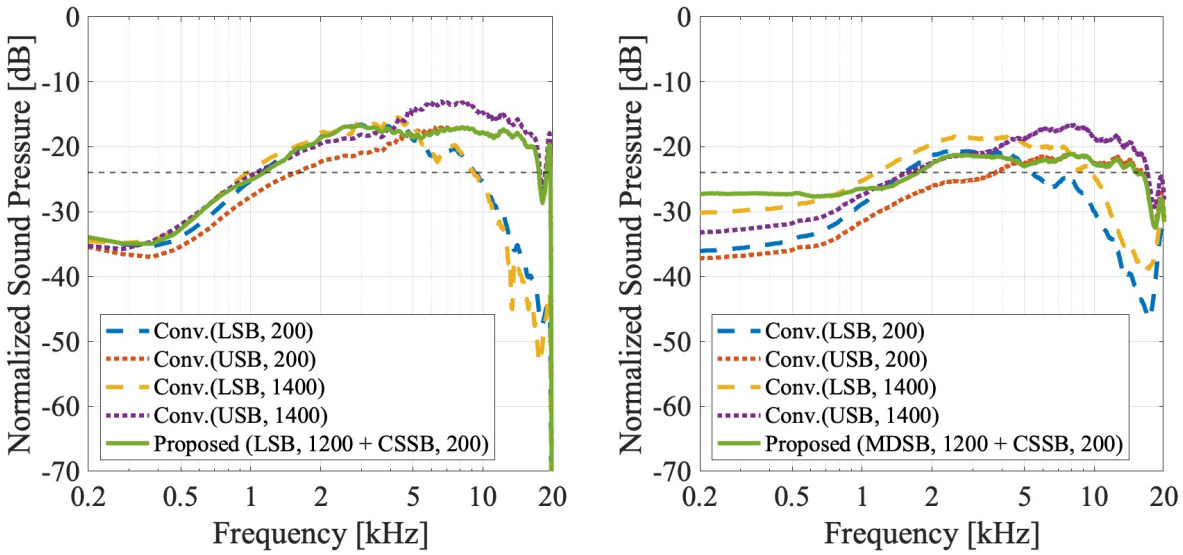
(a) Log-TSP signal (b) White noise

Figure 3.18: Measured frequency responses in the target area (0 m, 1.0 m).



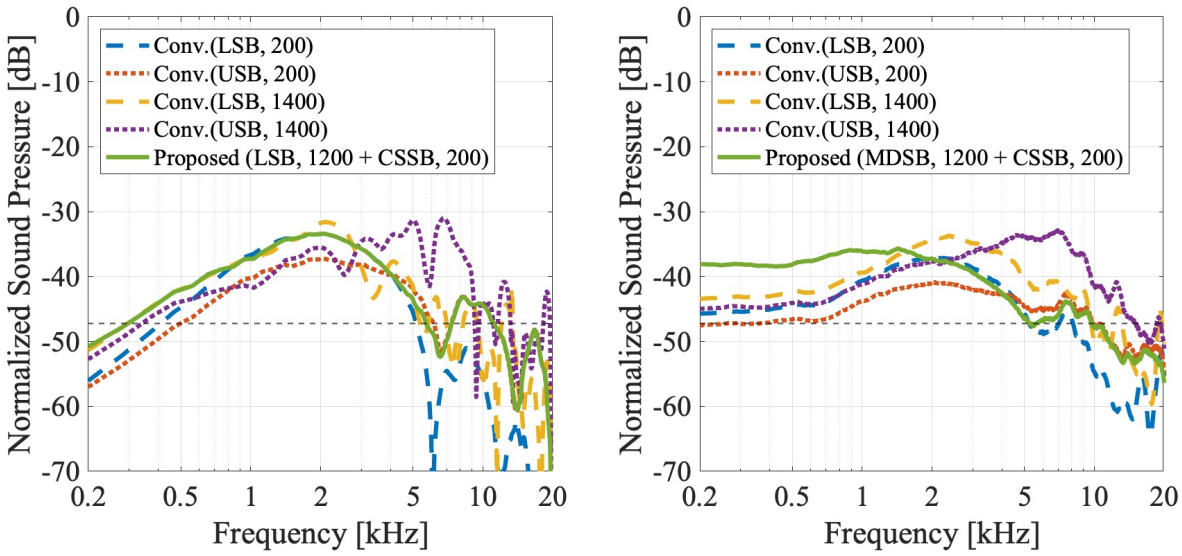
(a) Log-TSP signal (b) White noise

Figure 3.19: Measured frequency responses in the target area (0 m, 3.0 m).



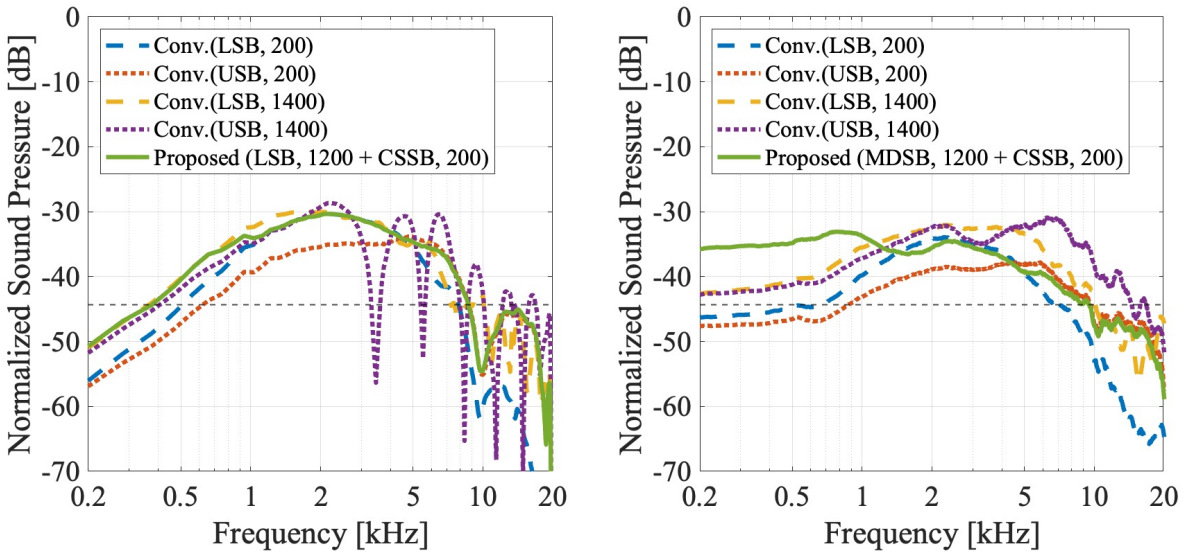
(a) Log-TSP signal (b) White noise

Figure 3.20: Measured frequency responses on the boundary of the target area (0 m, 0.25 m).



(a) Log-TSP signal (b) White noise

Figure 3.21: Measured frequency responses in the non-target area (-0.5 m, 1.0 m).



(a) Log-TSP signal (b) White noise

Figure 3.22: Measured frequency responses in the non-target area (-0.5 m, 2.0 m).

Table 3.7: RMSE between measured and desired flat frequency responses with log-TSP signal [dB].

Position	Conv. * (LSB, 200)	Conv. (USB, 200)	Conv. (LSB, 1400)	Conv. (USB,1400)	Prop.
(0 m, 1.0 m)	8.45	7.43	7.08	6.76	**5.36**
(0 m, 2.0 m)	8.70	8.69	7.41	8.89	**7.14**
(0 m, 3.0 m)	8.90	9.03	7.91	9.07	**7.35**
Average	8.68	8.38	7.47	8.24	**6.62**

Table 3.8: RMSE between measured and desired flat frequency responses with white Gaussian noise [dB].

Position	Conv. * (LSB, 200)	Conv. (USB, 200)	Conv. (LSB, 1400)	Conv. (USB,1400)	Prop.
(0 m, 1.0 m)	7.29	5.91	6.75	5.74	**2.59**
(0 m, 2.0 m)	7.25	6.39	5.86	7.36	**2.63**
(0 m, 3.0 m)	7.39	6.55	5.70	7.12	**3.14**
Average	7.31	6.28	6.10	6.74	**2.79**

with a log-TSP signal are shown in Table 3.7, and the results of RMSE between measured and desired flat frequency responses with white Gaussian noise are shown in Table 3.8. The results indicate that the proposed method achieves a flatter frequency response in the target area than conventional full band reproduction.

The difficulty in low-frequency reproduction leads to the fact that though a flat frequency response can be achieved by equalization, the total sound pressure decreases because the flat frequency response is realized by suppressing the high-frequency. Therefore, forcing a flat frequency response in large-scale PAL such as the condition of Conv. (USB, 1400) will lead to a significant decrease of demodulated sound pressure. However, the proposed method realizes a flat frequency response by enhancing the sound pressure in low-frequency. The multi-way structure brings specialized roles to Woofer PAL and Tweeter PAL so that the low-frequency can be reproduced with higher efficiency. Based on all experimental results and discussions above, the effectiveness of the proposed method has been demonstrated.

3.5 Conclusion

It is known that PAL can realize a sharp directivity but it has a low sound pressure when reproducing low-frequency. The difficulty of low-frequency reproduction has not been solved only by conventional modulation methods. To tackle this task, a multi-way PAL which can achieve a flat frequency response in a wide band is proposed. In the proposed method, the low-frequency and high-frequency are reproduced from Woofer PAL and Tweeter PAL, respectively. The scale of the Woofer PAL is enlarged to gain a higher demodulated sound pressure in low-frequency. And the hardware arrangement is studied considering the directional characteristics because the low-frequency reproduced from Woofer PAL and the high-frequency reproduced from Tweeter PAL should be heard simultaneously. Moreover, suitable modulation methods are also discussed, in order to reach a flat frequency response in a wider band.

The results of evaluation experiments on the prototype multi-way PAL have demonstrated the effectiveness of the proposed method. Especially, in low-frequency band below 1 kHz, the sound pressure is enhanced by about 10 dB over conventional full band reproduction, which is achieved by the studies on the multi-way structure, the band division, and modulation method. Moreover, a sharp directivity is formed in low-frequency band, in which the sound tends to spread with a wide directivity.

It can be noticed that the reproduction of wideband signals can be improved with the proposed method. This suggests a better performance in practical use when reproducing percussion sounds. The performance for harmonic sound reproduction still needs to be improved. Therefore, the psychoacoustical approaches based on missing fundamental can be combined into the proposed method in the future. Moreover, the phase characteristic, wavefront synthesis, and down-scaling the size of the device should also be considered in future works.

The work presented in this chapter is published in IEICE Transactions on Information and Systems (Japanese Edition) [119].

Chapter 4

Sound Quality Improvement Based on Spectral Slope Control for Over-boosted Parametric Array Loudspeaker

4.1 Introduction

PAL [1,3,4] has attracted attention because it can form a narrow acoustic beam and carry sound only to target positions by exploiting the straight projection of ultrasound. When emitting two high-intensity ultrasound signals, a difference-frequency is generated in the air due to nonlinearity [5]. Hence, this frequency is able to reproduce a tone f_{S} by emitting primary waves: carrier wave f_{C} and sideband wave $f_{\mathrm{C}} \pm f_{\mathrm{S}}$. Due to the narrow directivity, PAL is useful in several applications, such as audio spotlights [10,90], 3-D sound field reproduction [20,36], and active noise control [25,27].

However, PAL suffers from low sound quality due to its frequency-dependent response and harmonic distortions. Berktay's far-field solution [6] has given a theoretical analysis of the demodulated sound of PAL under DSB modulation. And other modulation methods have been developed to improve the demodulated sound quality of PAL, which are introduced in Section 2.3.1. A filter with a gain of -12 dB/octave can be used for equalization against PAL's frequency-dependent response [40,43]. Moreover, approaches using the Volterra filter have been proposed to reduce the harmonic distortions as well [64,66].

PAL also faces a low reproduced sound pressure because the demodulated sound is generated among nonlinear interactions in the air. In practical use, it is necessary to drive a higher voltage than the standard voltage on PAL to obtain a higher demodulated sound pressure, which can be called "over-boosting" PAL. It has been demonstrated that the primary sound pressure affects the amount of harmonic distortion [40]. Therefore, it can be predicted that over-boosted PAL shows poorer sound quality. Although there is a risk of degrading sound quality, sufficient sound pressure takes priority over sound quality, so it is necessary to over-boost PAL for time intervals with insufficient levels. The goal is to reduce the distortion of PAL under the over-boosting. Therefore, a high demodulated sound pressure can be achieved by over-boosting PAL, while the sound quality can be improved with the propose process.

If the optimal parameters can be given to improve the sound quality for each driving voltage, it is able to counteract the influence of over-boosting. To improve the sound quality, the focus is put on the spectral slope [120], a metric representing the shape and tendency of the spectrum. A typical harmonic sound owns a harmonic structure, consisting of a fundamental frequency f_0, and harmonic components with multiple frequencies of f_0 ($2f_0$, $3f_0$, $\cdots$,Nf_0). Clearly, the slope of the demodulated sounds varies from that of the sound source due to PAL's frequency-dependent response and nonlinear distortions. If the demodulated sound's slope can be predicted, it is able to restore the sound source's slope in the demodulated sound by adjusting the slope in advance. It can also noticed that part of the harmonic distortion has the same frequency as some high-frequency natural components. This reveals the possibility of making use of harmonic distortions, which is in contrast to conventional approaches.

This thesis reviews a sound quality improvement method for harmonic sounds reproduced by an over-boosted PAL. A voltage adaption is involved in the proposed method. Once the driving voltage is determined, the sound source will be processed with optimized parameters to the driving voltage. The process of the proposed method consists of an equalization based on the spectral slope, and a harmonic control to modify the sound source's spectrum in detail. Models were constructed for adjusting the spectral slope and using the harmonic distortions. With the proposed method, the sound source's spectrum can be restored in the demodulated sound with less harmonic distortion under over-boost voltage.

The remainder of this chapter is organized as follows. In Section 4.2, theoretical analysis is given and preliminary experiments are carried out. In Section 4.3, the proposed sound quality

improvement based on spectral slope control is described. In Section 4.4, experiments are carried out to evaluate the performance of the proposed method. Finally, this thesis concludes in Section 4.5.

4.2 Theoretical Analysis and Preliminary Experiments

4.2.1 Theoretical Analysis of Harmonic Sound Reproduction by Parametric Array Loudspeaker

The basic theory of the PAL was originally introduced by Westervelt [4], and a further explanation was given when Berktay provided the far-field solution [6], which defines the demodulated sound as:

$$p(t) \propto \frac{\partial^2 E^2(t)}{\partial t^2}, \tag{4.1}$$

where t denotes the time index, $E(t)$ denotes the envelope function of the primary wave, and $p(t)$ denotes the demodulated sound. Using Eq. (4.1), it is able to predict the demodulated sound by analyzing the envelope of the primary wave. When $s(t)$ denotes the sound source, the DSB-modulated wave can be expressed as:

$$\mathrm{DSB}[s(t)] = (1 + m_{\mathrm{DSB}} s(t)) \cos(2\pi f_{\mathrm{C}} t), \tag{4.2}$$

where m_{DSB} denotes the modulation factor ($0 < m_{\mathrm{DSB}} \leq 1$) and f_{C} denotes the carrier frequency. Using Hilbert transform, the analytic signal of $s(t)$ can be obtained, which can be expressed as:

$$\mathcal{H}[s(t)] = s(t) + i\hat{s}(t), \tag{4.3}$$

where $\hat{s}(t)$ denotes the signal with a phase shift of 90° to $s(t)$. Then the envelope of the signal can be obtained as:

$$\mathcal{E}[s(t)] = \sqrt{\mathcal{H}[s(t)](\mathcal{H}[s(t)])^*}, \tag{4.4}$$

where * denotes the complex conjugate. Based on Eqs. (4.2), (4.3), and 4.4,the envelope of DSB-modulated wave can be calculated as:

$$\mathcal{E}[\mathrm{DSB}[s(t)]] = 1 + m_{\mathrm{DSB}} s(t). \tag{4.5}$$

Assuming that the sound source $s(t)$ can be decomposed into simple sinusoidal components, the following holds that

$$s(t) = \sum_{n=1}^{N} a_n \cos(2\pi f_n t). \tag{4.6}$$

The demodulated sound with DSB modulation can be obtained from Eqs. (4.1), (4.5) and (4.6):

$$\begin{aligned} p_{\mathrm{DSB}}(t) \propto & \sum_{n=1}^{N} m_{\mathrm{DSB}} a_n f_n^2 \cos(2\pi f_n t) \\ & + \sum_{n_1=1}^{N}\sum_{n_2=1}^{N} \frac{m_{\mathrm{DSB}}^2}{2} a_{n_1} a_{n_2} (f_{n_1}+f_{n_2})^2 \cos\left(2\pi (f_{n_1}+f_{n_2})t\right) \\ & + \sum_{n_1=1}^{N}\sum_{n_2=1}^{N} \frac{m_{\mathrm{DSB}}^2}{2} a_{n_1} a_{n_2} (f_{n_1}-f_{n_2})^2 \cos\left(2\pi (f_{n_1}-f_{n_2})t\right). \end{aligned} \tag{4.7}$$

Similarly, for SSB modulation, it is able to predict the demodulated sound by analyzing the envelope as follows:

$$\mathrm{SSB}[s(t)] = (1 + m_{\mathrm{SSB}} s(t)) \cos(2\pi f_{\mathrm{C}} t) \pm m_{\mathrm{SSB}} \hat{s}(t) \sin(2\pi f_{\mathrm{C}} t), \tag{4.8}$$

$$\mathcal{E}[\mathrm{SSB}[s(t)]] = \sqrt{1 + 2 m_{\mathrm{SSB}} s(t) + m_{\mathrm{SSB}}^2 [s^2(t) + \hat{s}^2(t)]}, \tag{4.9}$$

$$\begin{aligned} p_{\mathrm{SSB}}(t) \propto & \sum_{n=1}^{N} m_{\mathrm{SSB}} a_n f_n^2 \cos(2\pi f_n t) \\ & + \sum_{n_1=1}^{N}\sum_{n_2=1}^{N} m_{\mathrm{SSB}}^2 a_{n_1} a_{n_2} (f_{n_1}-f_{n_2})^2 \cos\left(2\pi (f_{n_1}-f_{n_2})t\right). \end{aligned} \tag{4.10}$$

It can be noticed from Eq. (4.7) that the demodulated sound $p_{\mathrm{DSB}}(t)$ consists of 3 parts: the natural components f_n, which exists in the sound source; the sum-frequencies components $f_{n_1} + f_{n_2}$; and the difference-frequencies components $|f_{n_1} - f_{n_2}|$. In the case of SSB modulation, as shown in Eq. (4.10), the demodulated sound $p_{\mathrm{SSB}}(t)$ contains 2 parts: the natural components f_n and the difference-frequencies components $|f_{n_1} - f_{n_2}|$.

In the most simple scene, $s(t)$ is a pure tone with $N = 1$. The natural component's amplitude p is weighted by the square of frequency f, as $p \propto f^2$, in other words, the fundamental frequency is gained +12 dB/octave. The sum-frequency $2f$ can be regarded as a second harmonic distortion, and the difference-frequency no longer exist.

Assuming that the sound source is a typical harmonic sound with a fundamental frequency f_0, $f_n = n f_0$ and $|f_{n_1} \pm f_{n_2}| = |n_1 \pm n_2| f_0$ hold. When $n_1 + n_2 > N$, the sum-frequency can be regarded as harmonic distortion because it exceeds the maximum frequency of the sound source's natural components; when $n_1 + n_2 \leq N$, this sum-frequency is added to certain natural components. Because $|n_1 - n_2| < N$ holds for any $n_1, n_2 = 1, 2, \cdots N$, this difference-frequency is always added to certain natural components. Given the above discussion, it can be predicted

that the spectrum of demodulated sound will be distorted for two reasons:
(1) the demodulated sound has an f^2-weighted response;
(2) the harmonic distortions and inter-modulation distortions are generated during the demodulation.

A -12 dB/octave filter has been proposed for equalizing a PAL [40, 43]. However, due to fabrication limitations for ultrasonic transducers, it is difficult to achieve a flat response. Therefore, the demodulated sound does not always show an f^2-weighted response in practice. As an alternative, we can measure the frequency response first and then design a filter using the inverse response for equalization. However, the performance is uncertain when reproducing harmonic sounds because the harmonic distortions and inter-modulation distortions will be much more complicated in this case.

4.2.2 Relationship between Demodulated Frequency Response and Driving Voltage

In this section, preliminary experiments are carried out to investigate the driving voltage's influence on the demodulated sound. At first, Fig. 4.1 gives an example of the relationship between the ultrasonic sound pressure and that of the demodulated sound. Because PAL has a frequency-dependent response, the demodulated sound pressure will be led by the frequency with a higher gain if a wide band signal is used as the sound source. Therefore, a pure tone at 2 kHz was used to modulate an ultrasonic carrier wave at 40 kHz in this experiment. Table 4.1 shows the experimental conditions, and Table 4.2 shows the equipment. The sound pressures of the demodulated sound and the carrier wave were measured 2.0 m away from the PAL using the sound pressure meters. The block diagram of sound reproduction with PAL can be shown as Fig. 4.2. As shown in Fig. 4.2, the instantaneous voltage of the electrical signal can be calculated with the product of the signal level and driving voltage. The driving voltage was set manually by adjusting the output gain of the power amplifier. In Fig. 4.1, the driving voltage on the horizontal axis is given with peak-to-peak voltage. It can be noticed in Fig. 4.1 that when a driving voltage of 24 V is applied, the demodulated sound of the pure tone reaches 70 dB, which is enough for ordinary sound reproduction. In subsequent experiments, we applied 24 V as a default voltage [114].

Then, the characteristics of harmonic distortions in demodulated sounds are investigated

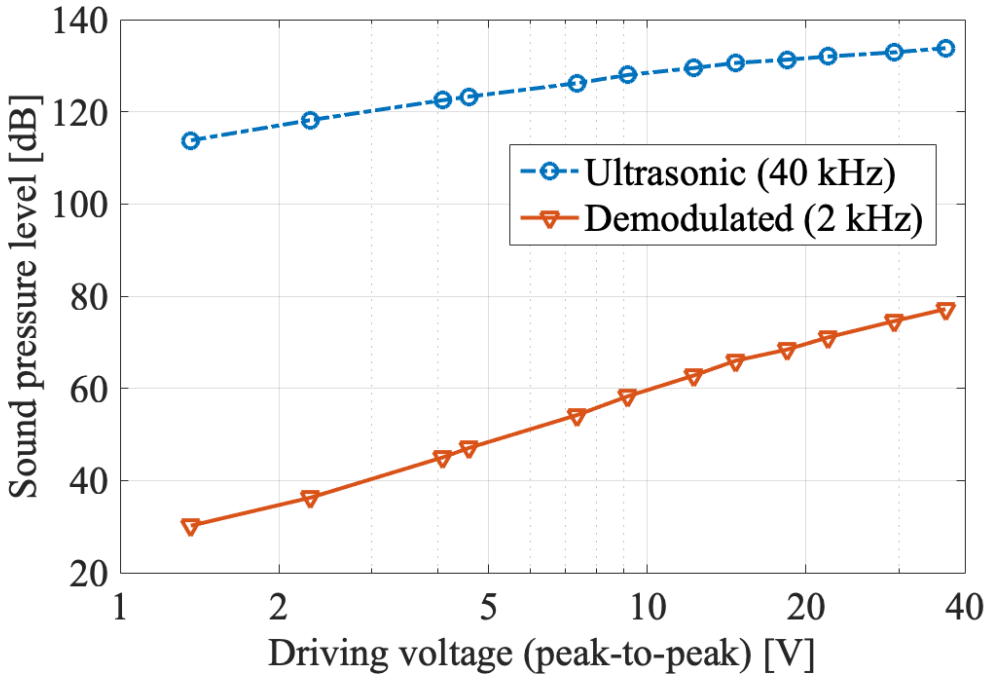


Figure 4.1: Measured sound pressure level.

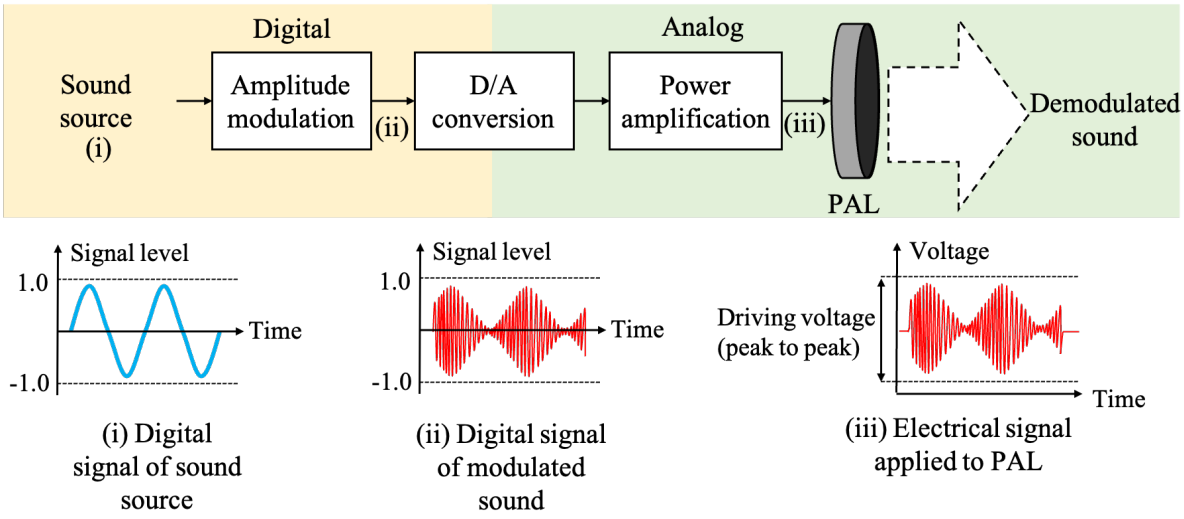


Figure 4.2: Block diagram of sound reproduction with PAL.

with impulse response [121]. The impulse response can be measured with TSP signal [116] by operating the convolution of a measured TSP signal and a reversed TSP signal. In this experiment, log-TSP signal [117, 118] is used for measurement, because the main response and each order of the harmonic distortion can be separated from the calculated impulse response in the time domain, as shown in the waveform in Fig. 4.3.

Table 4.1: Experimental conditions for preliminary experiments.

Reverberation time	650 ms
Ambient sound level	27.3 dB
Sampling frequency	192 kHz
Quantization	32 bits
Driving voltage	v = 18, 24, 29, 36 V
Modulation method	DSB modulation
Carrier frequency	f_{C} = 40 kHz
Modulation factor	m = 1.0

Table 4.2: Experimental equipment for preliminary experiments.

PAL	MITSUBISHI, MSP-50E
Power amplifier	VICTOR, A2002
Sound pressure meter	RION, UN-14 (ultrasonic) RION, NL-27 (audible)
Microphone	SENHEISER, MKH-8020
D/A, A/D converter	RME, FIREFACE UFX

The microphone was set 2.0 m away from the PAL to record demodulated sounds. Since the ultrasounds are emitted intensely from the PAL, there is a risk that the pseudo sound, a nonexistent low-frequency vibration, may be observed in the recorded waves of demodulated sound [122, 123]. In the experiments, the omni-directional microphone is set orthogonal to the radiation axis of the ultrasounds to prevent the pseudo sound from being involved in the recorded signals [124].

The measured main response, second harmonic distortion and third harmonic distortion are shown in Figs. 4.4, 4.5 and 4.6, respectively. It is noticed from Fig. 4.4 that the main response rises with a $p \propto f^{\mu}$ tendency at 200 - 2,000 Hz and fluctuate at 2,000 - 10,000 Hz. Here, μ is an exponential coefficient, and μ is approximately 2 around 500 - 1,000 Hz. The second and third harmonic distortion increase as the sound source frequency increases up to 2,000 Hz and then fall as the sound source frequency increases. These frequency responses under different driving

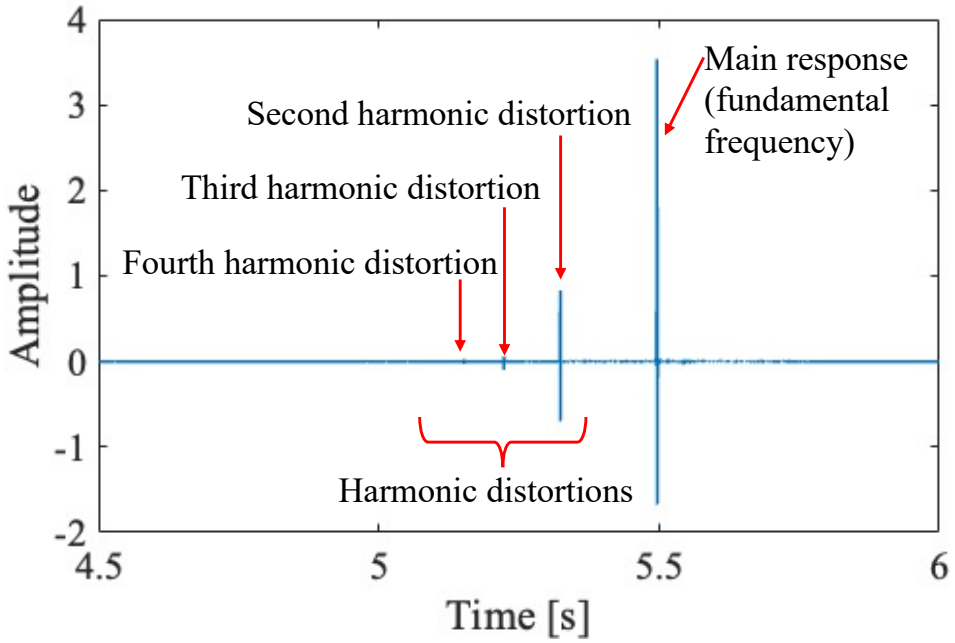


Figure 4.3: Waveform of calculated impulse response with measured log-TSP signal.

voltages are approximately parallel, suggesting that we can predict the frequency response in a certain range of voltage using analytical approaches such as interpolation and regression.

4.2.3 Relationship between Characteristics of Demodulated Harmonic Sounds and Driving Voltage

Another experiment is carried out to investigate the characteristics between the sound sources and demodulated sounds. As discussed, an assumption has been made that the sound quality degrades as the driving voltage on PAL is increased. The assumption will be demonstrated in these preliminary experiments.

Metrics including spectral slope [120], spectral centroid [125], and harmonic distortion are focused on. Ordinarily, the spectral slope is calculated as

$$\text{slope} = \frac{\sum_{k=k_\text{L}}^{k_\text{U}} (f_k - \bar{f})(S_k - \bar{S})}{\sum_{k=k_\text{L}}^{k_\text{U}} (f_k - \bar{f})^2}, \tag{4.11}$$

where k denotes the bin index of the Fourier transform with an upper limit k_U and a lower limit k_L, f_k denotes the frequency of the k-th bin, $\bar{f}$ denotes the average of f_k in the range of $k_\text{L} \leq k \leq k_\text{U}$, S_k denotes the power spectrum of the k-th bin, and $\bar{S}$ denotes the average of S_k

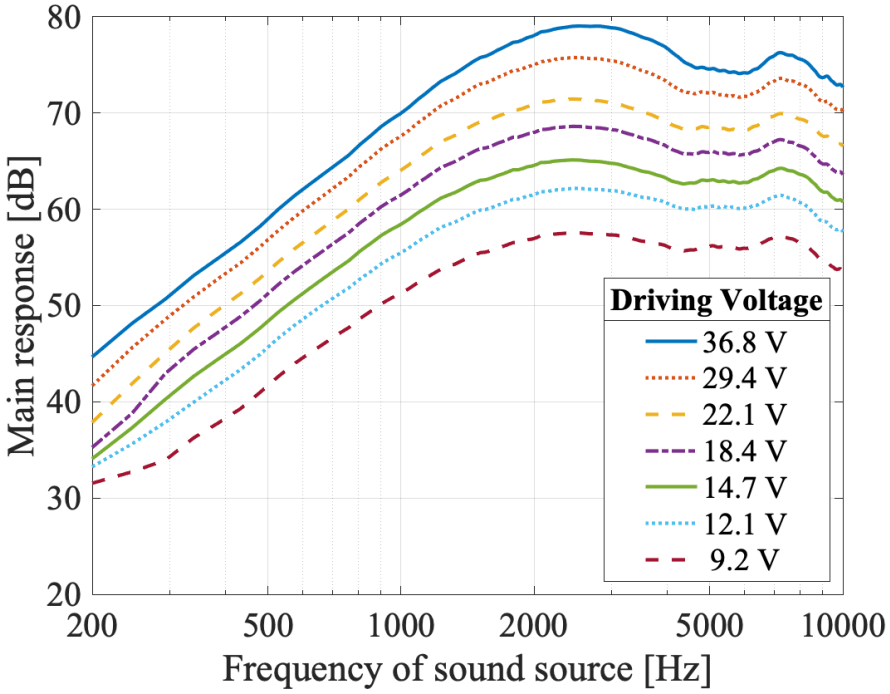


Figure 4.4: Measured Frequency response.

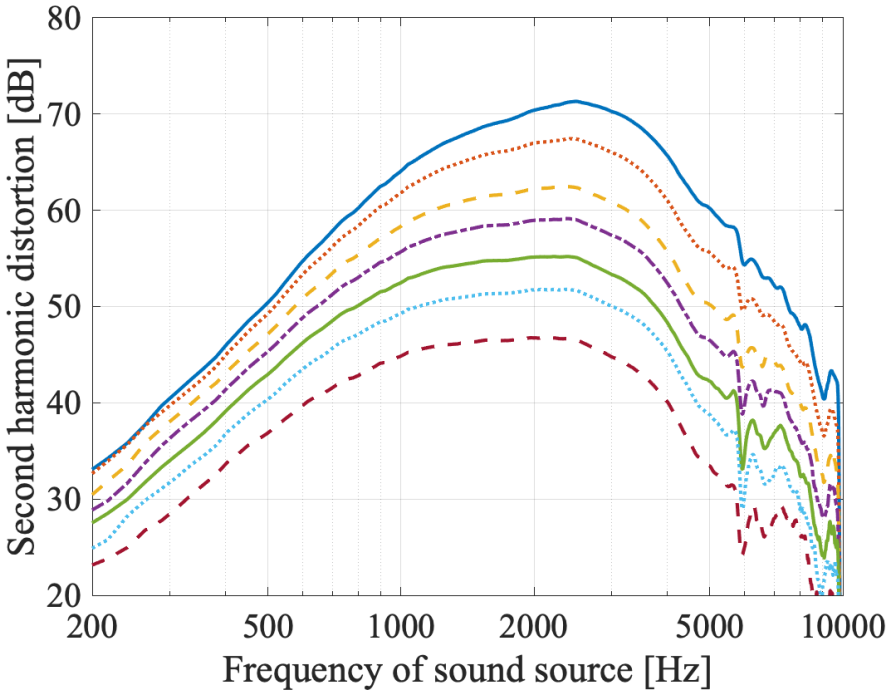


Figure 4.5: Measured second order harmonic distortions.

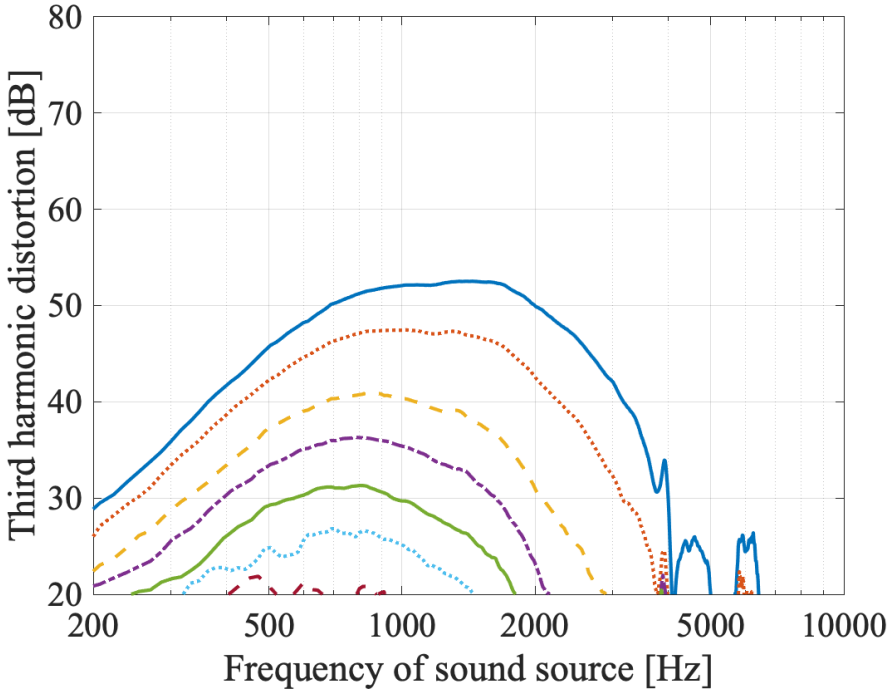


Figure 4.6: Measured third order harmonic distortions.

in the range of $k_{\mathrm{L}} \leq k \leq k_{\mathrm{U}}$. In this research, the $p \propto f^{\mu}$ tendency (μ denotes an exponential coefficient) carries significance in demodulated sounds, so log-slope is calculated by substituting the frequency and power spectrum under logarithmic scale. Sound sources with a $1/f$ power spectrum, such as pink noise, have a log-slope of -1. A larger RMSE of log-slope suggests that the spectral slope drifts more in the reproduction. The spectral centroid represents the “center of gravity" of the spectrum, and it can be used as an indication of brightness of timbre [125]. The calculation of the spectral centroid can be expressed as

$$\mathrm{centroid} = \frac{\sum_{k=k_{\mathrm{L}}}^{k_{\mathrm{U}}} f_k S_k}{\sum_{k=k_{\mathrm{L}}}^{k_{\mathrm{U}}} S_k}. \tag{4.12}$$

A lower RMSE value for the spectral centroid suggests a better reproduction of the sound source. Total harmonic distortion (THD) represents the ratio of harmonic distortions to the original components and can be expressed as

$$\mathrm{THD} = \sqrt{\frac{\sum_{n=N+1}^{n_{\max}} a_n^2}{\sum_{n=1}^{N} a_n^2}}, \tag{4.13}$$

where n denotes the index of the harmonic components, N denotes the number of natural components in the sound source, $n_{\max}$ denotes the total number of harmonic components detected

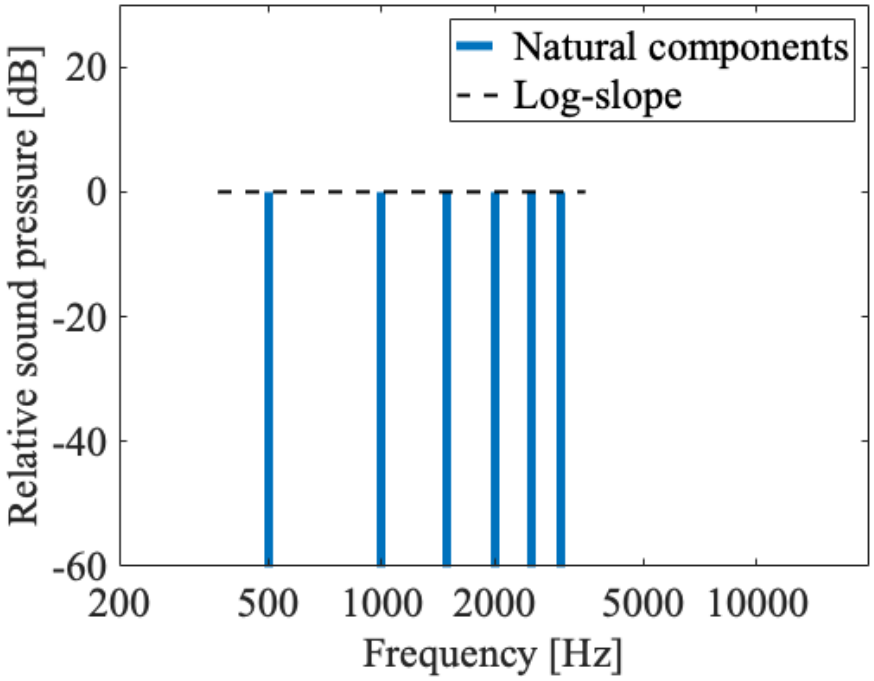
from the demodulated sounds, and a_n denotes the amplitude of n-th harmonic components. A lower value for THD suggests less generation of harmonic distortion.

In this experiment, artificially synthesized harmonic sounds are utilized as sound sources. The fundamental frequency f_0 of the sound sources was set to 250 - 2,000 Hz in 1/2 octave intervals. The number of natural components in N sound sources was set to 6. Therefore, the sound source with lowest f_0 consists of components at 250, 500, 750, ..., 1,500 Hz, and the sound source with highest f_0 consists of components at 2,000, 4,000, 6,000, ..., 12,000 Hz. The log-slopes of the sound sources were set to -3, -2, -1, 0, 1, and 2. The driving voltages were set to 18, 24, 29, and 36 V (peak-to-peak voltage).

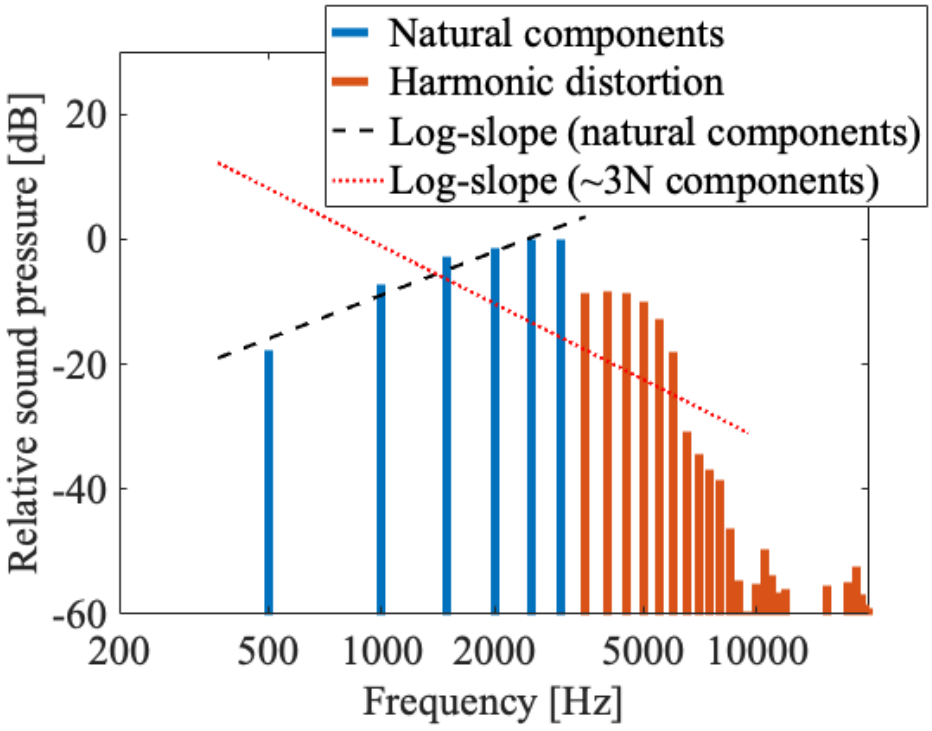
In Fig. 4.7, an example of a pair of the sound source and demodulated sound is given. The sound source has a log-slope of 0, so the dashed line showing the log-slope in Fig. 4.7 (a) is flat. In Fig. 4.7 (b), it can be noticed that natural components and acquired harmonic distortions exist in the demodulated sound. Here, the log-slope is calculated among natural components only, as the dashed line in Fig. 4.7 (b). If the log-slope is calculated among all observed components, for example, $3N$ components as the dotted line in Fig. 4.7, it cannot express the characteristics of natural components in the demodulated sound.

In Fig. 4.8, experimental result for RMSE of log-slope is given in. It can be noticed that the RMSE of log-slope increases as the driving voltage is increased, especially when f_0 is low. The results for root mean square error (RMSE) of spectral centroid are given in Fig. 4.9. Here, both the natural components and harmonic distortions are involved in the calculation of spectral centroid. Because several f_0 patterns have been prepared in the sound sources, the unit is converted from Hz to octave. It can be noticed from Fig. 4.9 that the RMSE of spectral centroid increases as the driving voltage is increased. The results for THD are given in Fig. 4.10. Here, THD is defined to calculate the ratio of the squared sum of levels of natural components ($n = 1, 2, 3, ..., N$) to the squared sum of acquired harmonic distortion levels ($n = N + 1, ...$). It can be noticed from Fig. 4.10 (c) that the amount of harmonic distortions increases as the driving voltage is increased.

The measured results of log-slope and regression lines are shown in Fig. 4.11. RL shows a reference line in each panel of Fig. 4.11. This reference line refers to the case that the log slope of the demodulated sound corresponds to the log slope of the source. Results above reference line mean the log-slope of demodulated sound is higher than that of the sound source, and vice versa.



(a) Sound source



(b) Demodulated sound

Figure 4.7: Example of measured spectra and calculated log-slopes.

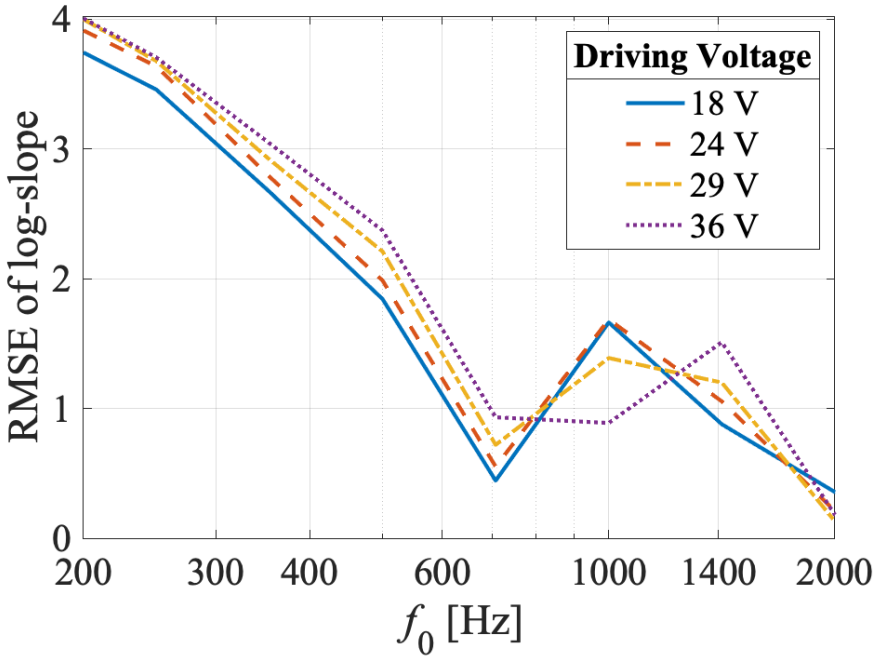


Figure 4.8: Spectral slope of demodulated harmonic sounds.

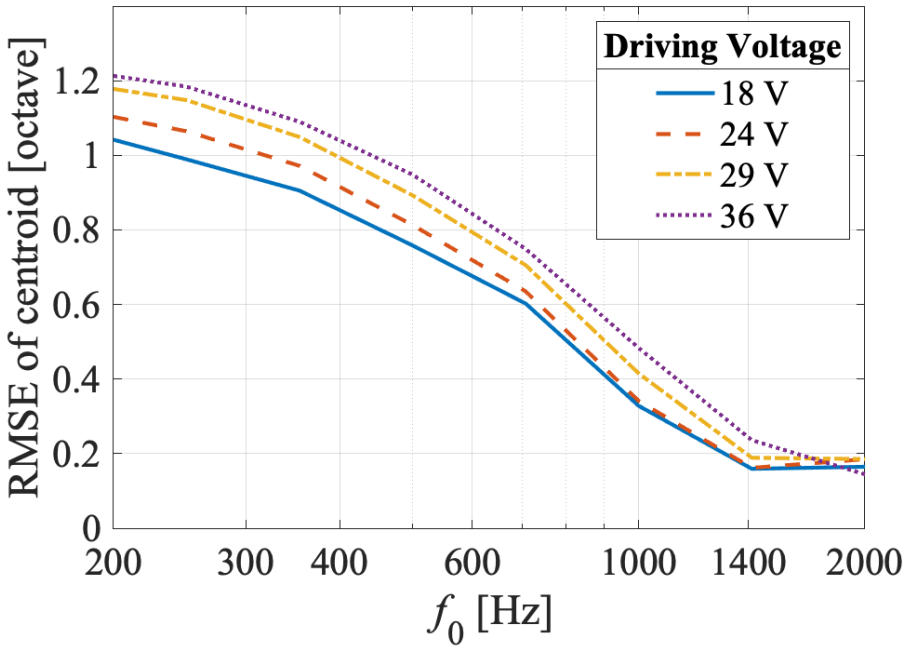


Figure 4.9: Spectral centroid of demodulated harmonic sounds.

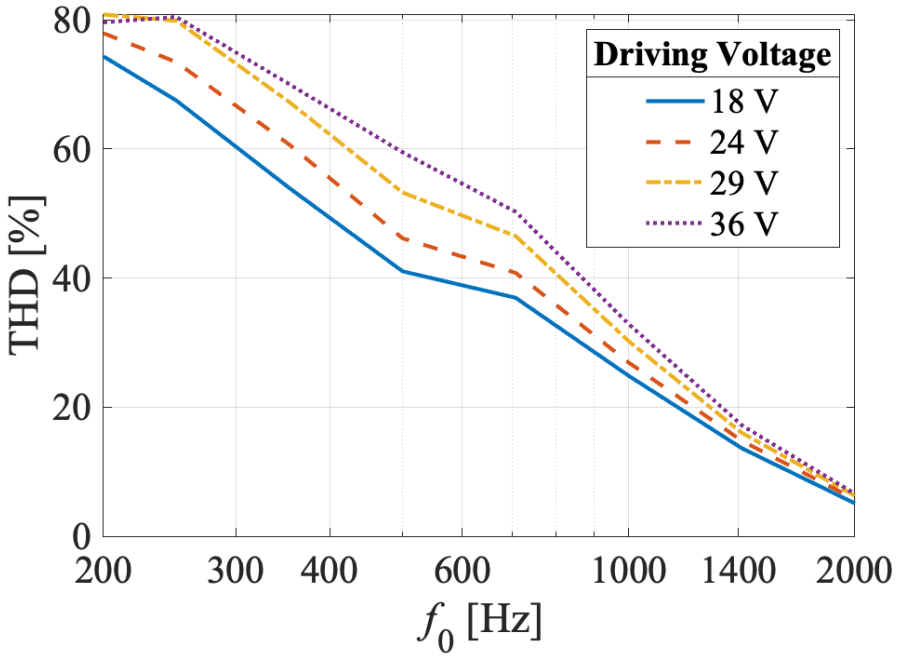
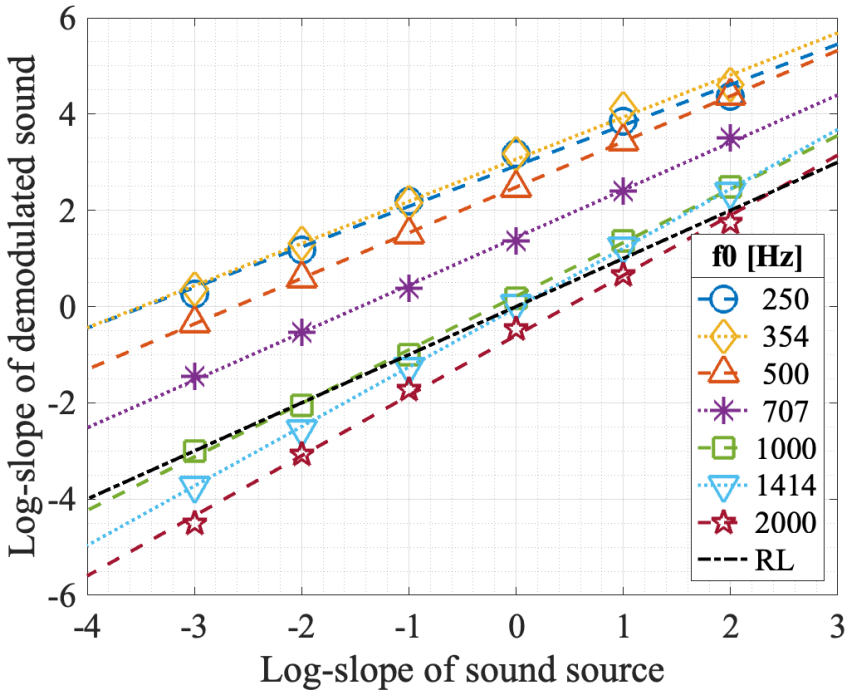


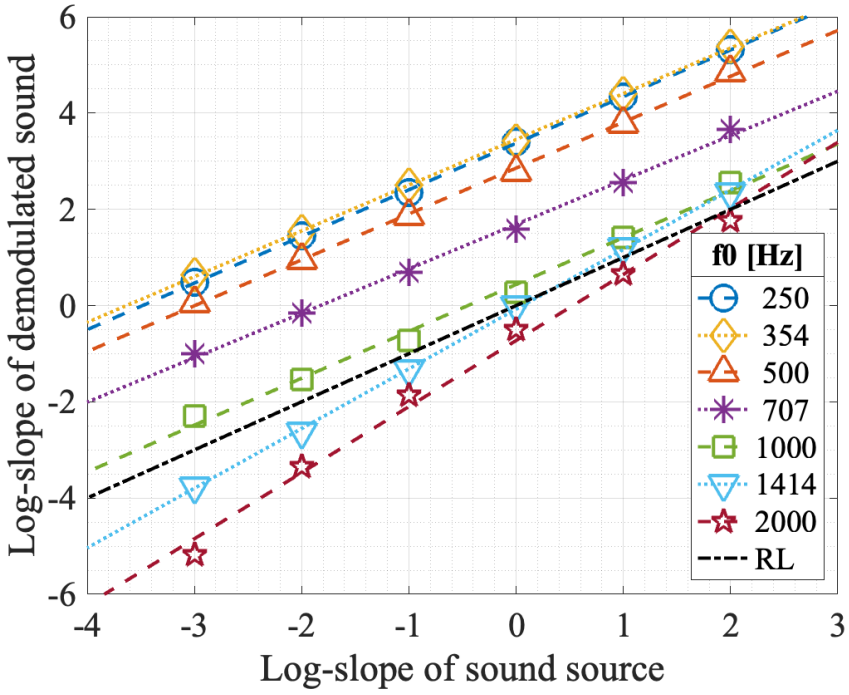
Figure 4.10: Total harmonic distortion of demodulated harmonic sounds.

It can be confirmed that when f_0 is lower, the demodulated sound has a larger log-slope than the sound source. Moreover, the difference between the log-slopes of demodulated sound and sound sources varies with f_0, making an f_0-dependent adjustment necessary. It can also be confirmed that the log-slopes of sound sources and demodulated sounds show a linear relationship. The coefficients of determination for linear regression R^2 are calculated for lines in Fig. 4.11. The values of R^2 show a distribution between 0.98 and 1.0, meaning that it is reasonable to use linear regression to predict the log-slope of demodulated sounds approximately.

From the results above, it can be confirmed that when a larger driving voltage is applied on PAL, the spectral slope drifts, the harmonic distortion increases, so the sound quality degrades. This conclusion suggests that when over-boosting PAL, it is necessary to consider the driving voltage and select the optimized parameters according to it.



(a) Driving voltage: 24 V



(b) Driving voltage: 36 V

Figure 4.11: Relationship of log-slopes between sound source and demodulated sound (natural components only).

4.3 Sound Quality Improvement Based on Spectral Slope Control

4.3.1 Overview

A sound quality improvement based on spectral slope control is proposed for harmonic sound reproduction by a over-boosted PAL. As shown in Fig. 4.12, the process is mainly divided into 3 stages: analysis, modification, and synthesis and post-processing. In the analysis stage, the sound source is decomposed. In the modification stage, the sound source is processed with 2 steps: spectral slope control and harmonic control. The first step, spectral slope control, is to equalize the sound source roughly with a slope-based model. The second step, harmonic control, is to modify the sound source in detail to reduce the harmonic distortion in the demodulated sound. In the synthesis and post-processing stage, the processed sound is synthesized and modulated for emitting from PAL.

Based on the results shown in Section 4.2.3, it is able to construct a model to predict the demodulated spectral slope with the driving voltage and the information of the sound source. Based on the results in Section 4.2.2, the harmonic distortions generated from low-frequency components can be partially predicted with the driving voltage and the information of the sound source. Since some of these harmonic distortions will add to high-frequency natural components in demodulation due to the same frequency, it is able to exploit these harmonic distortions to reduce the harmonic distortion.

Utilizing the proposed method, the sound source is processed adaptively to the driving voltage which has been determined in advance to apply to PAL. Therefore, the degradation of sound quality when over-boosting PAL can be counteracted.

4.3.2 Analysis

In the analysis stage, the sound source $s(t)$ is decomposed into sinusoids using sinusoidal analysis [126]. Several variations of the sinusoidal model have been developed for different purposes [127–129], and in this research, the harmonic plus noise model [127] is sufficient for analysis

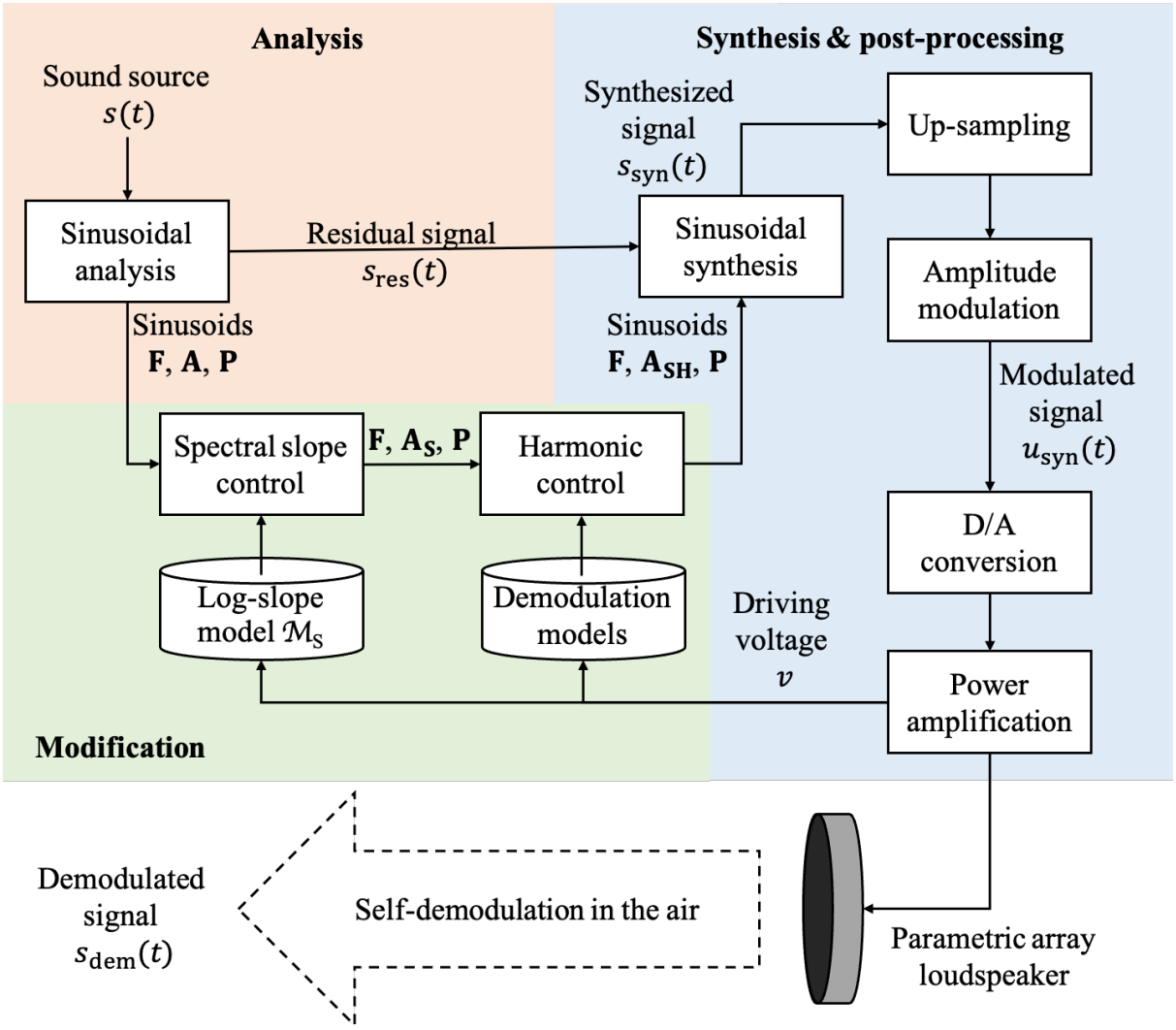


Figure 4.12: Overview of proposed method.

and synthesis. Assuming that the sound source $s(t)$ is a harmonic signal, the following holds:

$$\mathbf{F}, \mathbf{A}, \mathbf{P} = \mathcal{A}[s(t)], \tag{4.14}$$

where $\mathbf{F} = \{f_1, f_2, \cdots, f_N\}$, $\mathbf{A} = \{a_1, a_2, \cdots, a_N\}$, and $\mathbf{P} = \{\varphi_1, \varphi_2, \cdots, \varphi_N\}$ denote the extracted frequencies, amplitudes, and phases, respectively, and $\mathcal{A}$ denotes the operation of the sinusoidal analysis. After decomposing the sound source into sinusoids, sinusoidal synthesis is carried out to obtain the residual signal $s_{\mathrm{res}}(t)$, which can be indicated as follows:

$$s_{\mathrm{res}}(t) = s(t) - \mathcal{S}[\mathbf{F}, \mathbf{A}, \mathbf{P}], \tag{4.15}$$

where $\mathcal{S}$ denotes the operation of sinusoidal synthesis. The amplitudes $\mathbf{A}$ are processed in the modification stage. In contrast, the frequencies $\mathbf{F}$, phases $\mathbf{P}$, and residual signal $s_{\mathrm{res}}(t)$ remain

unprocessed until the synthesis stage.

4.3.3 Modification

In the modification stage, the extracted sinusoid's amplitudes $\mathbf{A}$ are weighted in 2 steps: spectral slope control and harmonic control.

Spectral slope control

It can be noticed from the results shown in Section 4.2.3 that the spectral slope varies in the demodulated sound. Therefore, to restore the sound source's log-slope in demodulated sound, it is necessary to modify the slope of the sound source in advance.

A log-slope control model $\mathcal{M}_{\mathrm{S}}$ is constructed for spectral slope control based on the experimental results shown in Section 4.2.3. The concept of this model is shown in Fig. 4.13. Model $\mathcal{M}_{\mathrm{S}}$ is a 3-D look-up table, containing the demodulated sound's log-slope λ_{dem} for each pattern of sound source's fundamental frequency f_0, log-slope λ_{src} and driving voltage v. Some measured data has been derived in Section 4.2.3, and linear interpolation is carried out for data augmentation.

Utilizing the log-slope model $\mathcal{M}_{\mathrm{S}}$, it is able to predict the spectral slope of demodulated sound approximately with the information of sound source and driving voltage, which can be indicated as:

$$\lambda_{\mathrm{dem}}(\lambda_{\mathrm{src}}, f_0, v) = \mathcal{M}_{\mathrm{S}}[\lambda_{\mathrm{src}}, f_0, v]. \tag{4.16}$$

Then, the optimal log-slope $\lambda_{\mathrm{opt}}(\lambda_{\mathrm{src}}, f_0, v)$ is predicted to let the demodulated sound's log-slope be same as that of the sound source:

$$\lambda_{\mathrm{opt}}(\lambda_{\mathrm{src}}, f_0, v) = \underset{\lambda}{\operatorname{argmin}} \left(\mathcal{M}_{\mathrm{S}}[\lambda, f_0, v] - \lambda_{\mathrm{src}}\right)^2. \tag{4.17}$$

It means that for a sound source with slope λ_{src} and fundamental frequency f_0, the slope should be adjusted to $\lambda_{\mathrm{opt}}(\lambda_{\mathrm{src}}, f_0, v)$ so that the demodulated sound can show the original slope λ_{src} in the case that the modulated sound is emitted under driving voltage v.

The amplitudes $\mathbf{A}$ are weighted to control the spectral slope using the optimal log-slope λ_{opt} as follows:

$$a_n^{\mathrm{S}} = a_n f_n^{(\lambda_{\mathrm{opt}}(\lambda_{\mathrm{src}}, f_0, v) - \lambda_{\mathrm{src}})/2}, \tag{4.18}$$

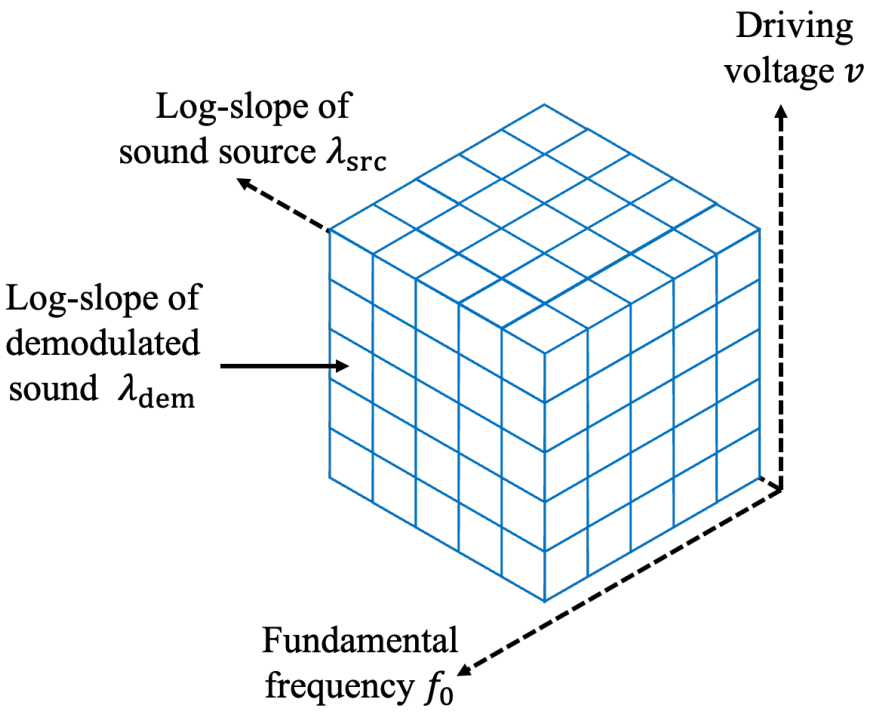


Figure 4.13: Concept of log-slope model $\mathcal{M}_{\mathrm{S}}$.

$$\mathbf{A_S} = \{a_1^{\mathrm{S}}, a_2^{\mathrm{S}}, \cdots, a_N^{\mathrm{S}}\}, \tag{4.19}$$

where $\mathbf{A_S}$ denotes the processed amplitudes after spectral slope control.

Harmonic control

Whereas it is able to restore the sound source's log-slope in the demodulated sound with spectral slope control, the problem appears that harmonic distortion is generated, which limits the sound quality. In this method, a harmonic control is added to exploit the harmonic distortion and improve the sound quality. The concept of harmonic control is shown in Fig. 4.14.

Based on the experimental results of Section 4.2.2, demodulation models $\mathcal{M}_{\mathrm{D1}}$, $\mathcal{M}_{\mathrm{DH2}}$, and $\mathcal{M}_{\mathrm{DH3}}$ were constructed. These models are 2-D look-up tables, containing the main response, second and third harmonic distortions of the demodulated sound for each frequency of the sound source and the driving voltage. As same as the log-slope model $\mathcal{M}_{\mathrm{S}}$, linear interpolation is carried out for unmeasured conditions. The models can be indicated as:

$$h_{\mathrm{dem}}(f, v) = \mathcal{M}_{\mathrm{D}}[f, v], \tag{4.20}$$

$$h_{\mathrm{hd2}}(f, v) = \mathcal{M}_{\mathrm{DH2}}[f, v], \tag{4.21}$$

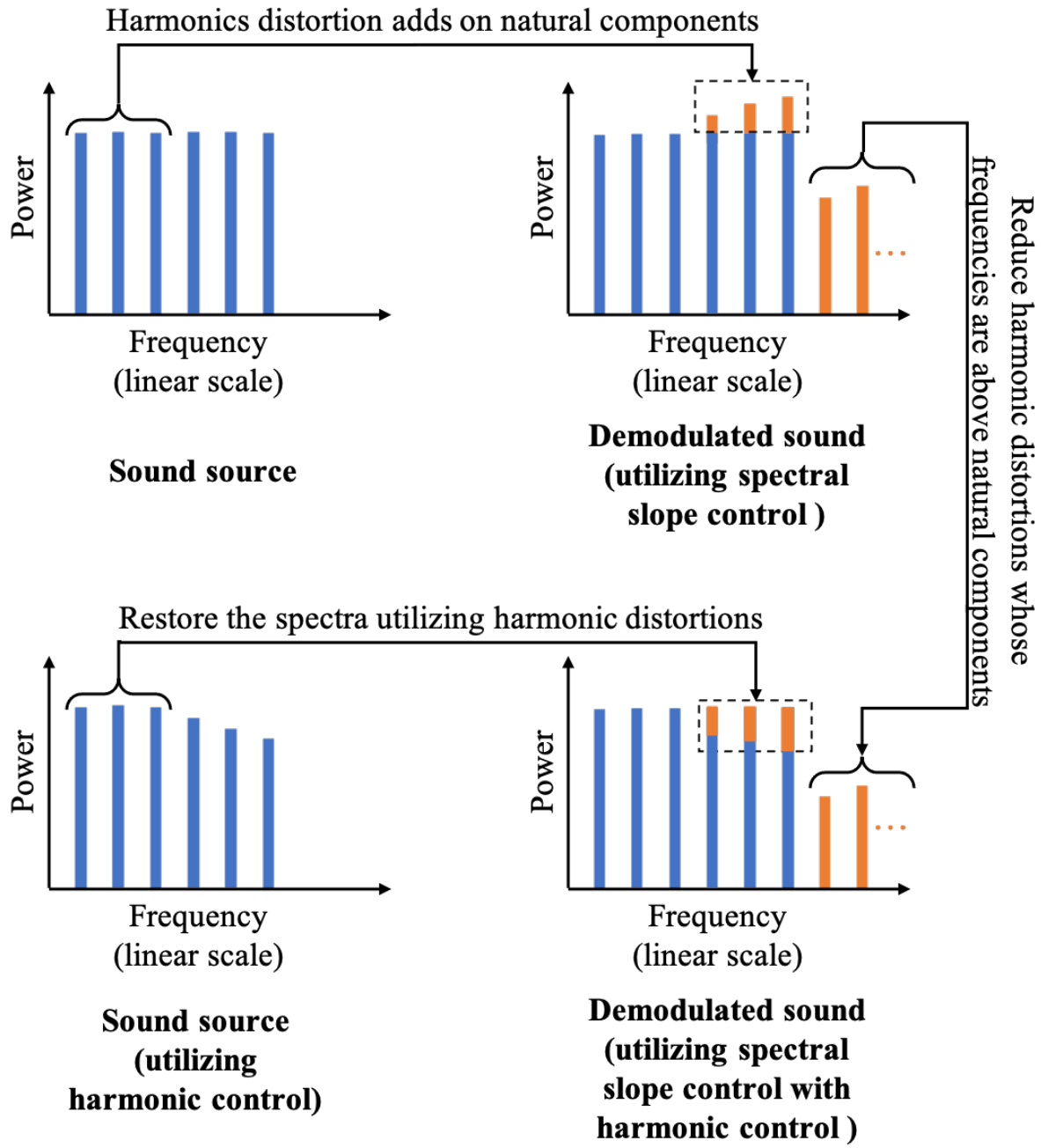


Figure 4.14: Concept of harmonic control.

$$h_{\mathrm{hd3}}(f,v) = \mathcal{M}_{\mathrm{DH3}}[f,v], \tag{4.22}$$

where $h_{\mathrm{dem}}(f,v)$, $h_{\mathrm{hd2}}(f,v)$, and $h_{\mathrm{hd3}}(f,v)$ denote the frequency response of the main response, second harmonic distortion, and third harmonic distortion, respectively, when the frequency of sound source is f and the driving voltage is v. With these models, it is possible to partially predict the harmonic distortions generated in demodulation so that it is able to use them to restore the spectrum of the sound source and reduce harmonic distortions with frequencies above those of the natural components.

In harmonic control, the natural components are divided into two parts by frequency. First, the harmonic distortions that are generated from lower-frequency components and will be added onto higher-frequency components are predicted, as:

$$\eta_n(f_0,v) = \sqrt{\left(a^{\mathrm{S}}_{\frac{n}{2}} r_{\mathrm{hd2}}\left(\frac{n}{2}f_0, v\right)\right)^2 + \left(a^{\mathrm{S}}_{\frac{n}{3}} r_{\mathrm{hd3}}\left(\frac{n}{3}f_0, v\right)\right)^2}. \tag{4.23}$$

This can be reduced from higher-frequency components with the number of harmonic distortions because it will add onto natural components in the demodulation. Taking η as compensation for harmonic distortion, the processed amplitudes of sinusoids $\mathbf{A_{SH}}$ can be indicated as:

$$a_n^{\mathrm{SH}} = \begin{cases} \sqrt{(a_n^{\mathrm{S}})^2 - (\eta_n(f_0,v))^2}, & \text{if } n > \frac{N}{2}, \\ a_n^{\mathrm{S}}, & \text{otherwise}, \end{cases} \tag{4.24}$$

$$\mathbf{A_{SH}} = \mathcal{N}[\{a_1^{\mathrm{SH}}, a_2^{\mathrm{SH}}, \cdots, a_N^{\mathrm{SH}}\}], \tag{4.25}$$

where $\mathcal{N}$ denotes the operation linear normalization, which makes the total power of the processed amplitudes $\mathbf{A_{SH}}$ equal to that of the original amplitudes $\mathbf{A}$.

4.3.4 Synthesis and Post-processing

After modification, the waveform is synthesized with processed sinusoids, and post-processing is carried out to emit the signal from the PAL. The waveform of the processed signal is synthesized with the original frequencies $\mathbf{F}$ and phases $\mathbf{P}$, the processed amplitudes $\mathbf{A_{SH}}$, and the residual signal $s_{\mathrm{res}}(t)$, given by

$$s_{\mathrm{syn}}(t) = \mathcal{S}[\mathbf{F}, \mathbf{A_{SH}}, \mathbf{P}] + s_{\mathrm{res}}(t), \tag{4.26}$$

where $\mathcal{S}$ denotes the operation of sinusoidal synthesis.

The synthesized signal is modulated for reproduction by the PAL with Eq. (4.2). Base on the sampling theorem, a high sampling frequency is needed to carry the ultrasound information. If required, up-sampling is performed to raise the sampling frequency of the sound source. The processed digital signal is converted to an analog signal and amplified with driving voltage v. From the results of preliminary experiments, it can be predicted that the sound quality will degrade if the PAL is over-boosted. However, because both spectral slope control and harmonic control involve driving voltage v in the parameter generation, the degradation of sound quality when over-boosting PAL can be partially neutralized.

4.4 Evaluation Experiment

4.4.1 Experimental Conditions

To demonstrate the effectiveness of the proposed method, evaluation experiment has been carried out. The experimental conditions are listed in Table 4.3 and equipment for the evaluation experiment is listed in Table 4.4. The driving voltage v_{def} was set to 24 V (peak-to-peak) as default, and the driving voltage when over-boosting the PAL v_{ovb} was set to 36 V (peak-to-peak).

Table 4.3: Experimental conditions for evaluation experiments.

Reverberation time	650 ms
Ambient sound level	27.3 dB
Sampling frequency	192 kHz
Quantization	32 bits
Modulation method	DSB modulation
Carrier frequency	f_{C} = 40 kHz
Modulation factor	m = 1.0
Driving voltage	v_{def} = 24 V (default voltage) v_{ovb} = 36 V (over-boost voltage)

Table 4.4: Experimental equipment for evaluation experiments.

PAL	MITSUBISHI, MSP-50E
Power amplifier	VICTOR, A2002
Microphone	SENHEISER, MKH-8020
D/A, A/D converter	RME, FIREFACE UFX

Artificially synthesized harmonic signals were used as sound sources because they are steady in the time and frequency domains, causing very little analysis and synthesis error compared with real-world harmonic sounds, such as speech and music. To validate the effectiveness, sound sources with a $p \propto f^{\mu}$ tendency is used, similar to the sound sources in the preliminary experiments. 6 patterns of the fundamental frequency f_0 are prepared, including 250, 300, 500,

600, 1000, and 1200 Hz. The number of natural components N is set to 6, as in the preliminary experiments. So the sound source with the lowest frequency contains components at 250, 500, 750, ..., 1,500 Hz, the sound source with highest frequency contains components at 1,200, 2,400, 3,600, ..., 7,200 Hz. Frequencies below 250 Hz have very low demodulated sound pressures, and the upper limit of f_0 is 1,200 Hz, therefore, it is enough to cover the frequency range of speech signal. 5 patterns of log-slope λ_{src} are prepared, including -1, -0.5, 0, 0.5, and 1. In total, 30 sound sources were prepared for each case. It can be noticed that some of the conditions did not appear in the preliminary experiments. The expanded range of parameters was intended to validate the effectiveness of the interpolations and regressions used in the proposed method.

5 methods are prepared for comparison: 3 baseline methods, and an incomplete proposed method excluding the voltage adaption, and the complete proposed method. “NON” stands for no process but amplitude modulation. “SS” stands for applying a solid spectral slope, which is equivalent to -12 dB/ octave equalization. “IFR” stands for weighting the sound source’s spectrum with the inverse frequency response of PAL. “PV” stands for the proposed method without voltage adaption. “PM” stands for the proposed method described in Section 4.3.

In the proposed method, the driving voltage v is involved when generating parameters. In the case of PM, when we carried out the experiment under default voltage v_{def}, v_{def} is substituted for v in the proposed method. When the experiment is carried out under over-boost voltage v_{ovb}, v_{ovb} is substituted for v in the proposed method. Therefore, voltage-adapted parameters are generated to process the sound sources. On the other hand, in the case of PV, v_{def} is always substituted for v, regardless of the driving voltage. So PV cannot give suitable parameters if the driving voltage is different from the default voltage.

4.4.2 Evaluation Metrics

The observed demodulated sounds were analyzed in the audible band and evaluated with the metrics described below.

Log-slope:

In this study, because the aim is to restore the sound source’s spectral slope in the demodulated sound, the log-slopes among natural components were evaluated. Since several log-slope patterns

are prepared in the sound sources, RMSE of the log-slope is calculated to evaluate the degree of restoration of the log-slope. A lower RMSE value for the log-slope suggests a better reproduction of the sound source.

Spectral centroid:

The spectral centroid represents the center of the spectrum, and it can be used as an indication of timbre [125]. It is used here as a metric to evaluate the difference between sound sources and demodulated sounds. The calculation has been defined in Eq. 4.12 A lower RMSE value for the spectral centroid suggests a better reproduction of the sound source.

Total harmonic distortion (THD):

THD represents the ratio of harmonic distortions to the original components. The calculation of THD has been defined in Eq. 4.13. A lower value for THD suggests less generation of harmonic distortion.

Cosine similarity:

The spectrum of a specific harmonic sound is ordinarily sparse, consisting of ambient noise and several narrow peaks of harmonic components. Here, the amplitudes of harmonic components are converted into vectors and compared the sound source's amplitude vectors and the demodulated sound. The cosine similarity was used to evaluate the extent to which the spectrum of the sound source was reproduced, given by

$$\text{cosine similarity} = \frac{\mathbf{A}_{\text{src}} \cdot \mathbf{A}_{\text{dem}}}{|\mathbf{A}_{\text{src}}||\mathbf{A}_{\text{dem}}|}, \tag{4.27}$$

where $\mathbf{A}_{\text{src}}$ denotes the amplitudes of harmonic components in the sound source and $\mathbf{A}_{\text{dem}}$ denotes the amplitudes of harmonic components in the demodulated sound, including natural components and harmonic distortion. A value closer to 1.0 suggests a better reproduction of the sound source.

Sound pressure:

The sound pressure of the demodulated sounds was also measured. Because the PAL has a frequency-dependent sound pressure, the sound pressures for the NON under the default voltage

are set as 0 dB references and the relative sound pressures are calculated.

4.4.3 Experimental Results

The overall experimental results under the default voltage and the over-boost voltage are shown in Table 4.5. As discussed, PM and PV process the sound sources with the same parameters under the default voltage, but they process the sound sources with different parameters under the over-boost voltage. It can be seen that the PM shows the best performance for all metrics of sound quality under both default voltage and over-boost voltage. However, the demodulated sound pressure decreases compared with other methods.

Comparing Table 4.5 (a) and (b), it is noticed that there is a larger error of spectral slope and there are more harmonic distortions, as the driving voltage increases from the default voltage to the over-boost voltage. It demonstrates that the demodulated sound quality degrades as the driving voltage increases. Comparing PM and PV under the over-boost voltage, it can be confirmed that the sound quality degradation due to the over-boosting can be reduced by considering the driving voltage when processing the sound source. The PAL shows a low sound pressure level at low frequencies but a relatively acceptable level at high frequencies. Therefore, over-boosting is more meaningful for low-frequency sounds than high-frequency sounds. Figs. 4.15 – 4.19 show the experimental results for each metric against f_0, to confirm whether the proposed method shows a steady performance as f_0 varies. Each result is an average of five results for different log-slopes in the sound sources.

The results for the log-slope are shown in Fig. 4.15. Due to the frequency-dependent response of the PAL, NON and SS show frequency-dependent performance, while the IFR shows relatively steady performance. However, because the effect of harmonic distortion is not considered, this case shows the worst performance when more harmonic distortion exists. With the spectral slope control, the PM case achieves a much better reproduction of the spectral slope, especially when f_0 is low. It is also seen that when emitting sound with an over-boost voltage, PM shows better performance than PV, which proves the effectiveness of the voltage adaptation.

The results for the spectral centroid and THD are shown in Figs. 4.16 and 4.17, respectively. It can be seen that PM achieves better and steadier performance over the baselines and PV. It seems like the SS case achieves less spectral centroid error and causes less harmonic distortion

Table 4.5: Experimental results for each metrics.

(a) Default voltage

	NON	SS	IFR	PM/PV
RMSE of log-slope	2.25	1.60	0.98	**0.35**
RMSE of centroid[oct]	0.73	0.54	0.58	**0.44**
THD[%]	43.3	33.0	36.6	**29.0**
Cosine similarity	0.84	0.86	0.90	**0.93**
Sound pressure[dB]	**0**	-1.0	-1.5	-2.8

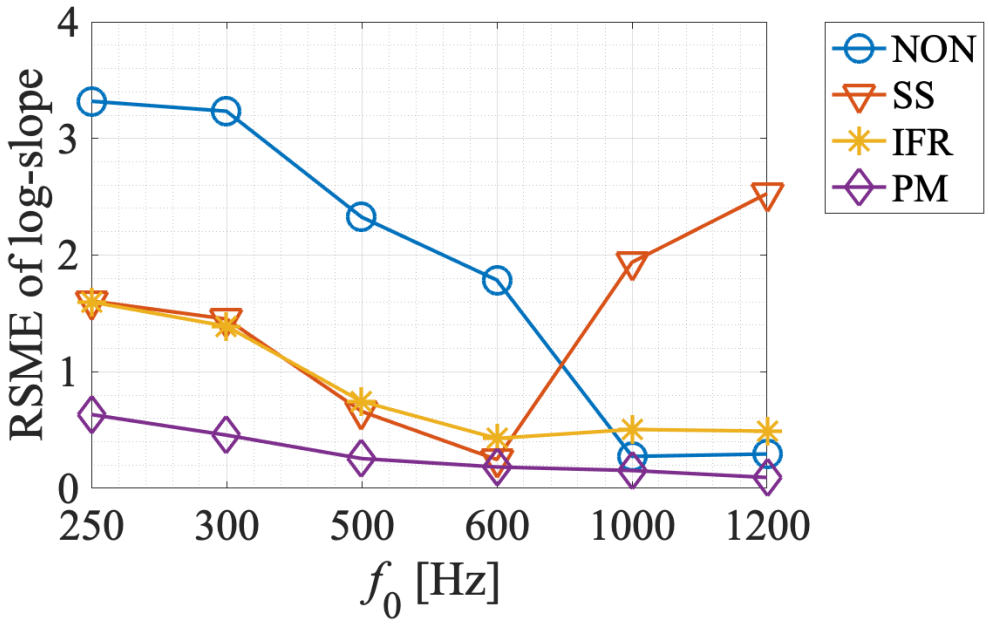
(b) Over-boost voltage

	NON	SS	IFR	PV	PM
RMSE of log-slope	2.41	1.42	1.07	0.46	**0.33**
RMSE of centroid[oct]	0.84	0.66	0.70	0.57	**0.53**
THD[%]	49.5	40.1	43.5	36.2	**33.8**
Cosine similarity	0.83	0.87	0.89	0.91	**0.92**
Sound pressure[dB]	**7.0**	5.9	5.4	3.9	3.6

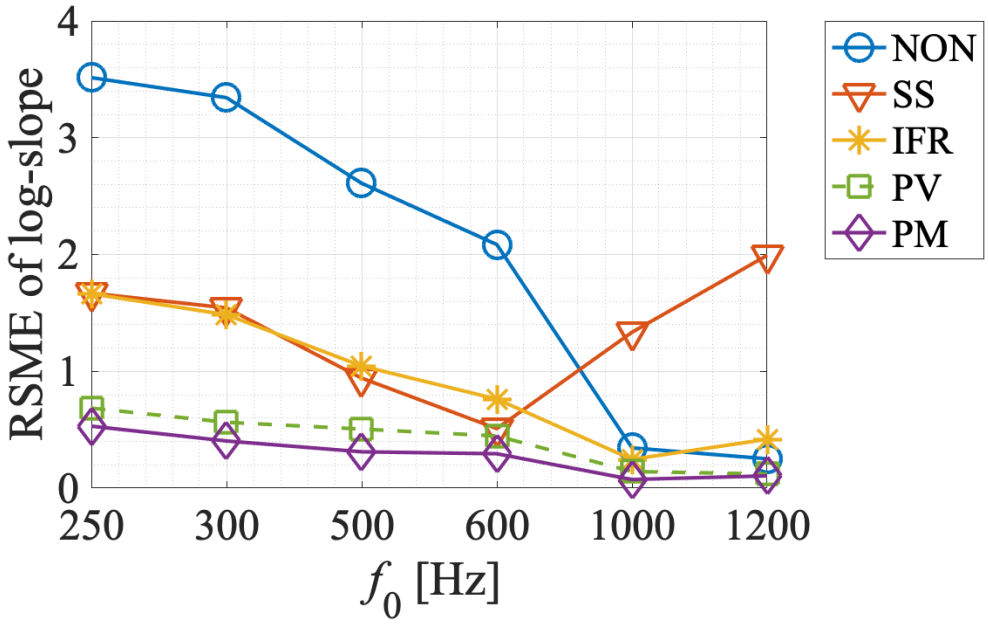
when f_0 is high. In this case, because the PAL exhibits a relatively flat response, it is apparent that excessively concentrating the power of the sound source at low frequencies results in lower RSME levels for the centroid and THD but distorts the spectral slope greatly, as shown in Fig. 4.15. The results for cosine similarity are shown in Fig. 4.18. It can be confirmed the proposed method achieves better and steadier performance when f_0 is low. The results for sound pressure are shown in Fig. 4.19. As predicted, concentrating the power of the sound source at low frequencies decreases the demodulated sound pressure, especially when f_0 is low. It can be seen that PM with an over-boost voltage reaches a level near that of the NON with the default voltage, which reveals that over-boosting the PAL can compensate for the decreased sound pressure.

4.4.4 Summary and discussions

The results presented above demonstrated that the proposed method can improve the demodulated sound quality. Though over-boosting the PAL increases distortion, the proposed method can

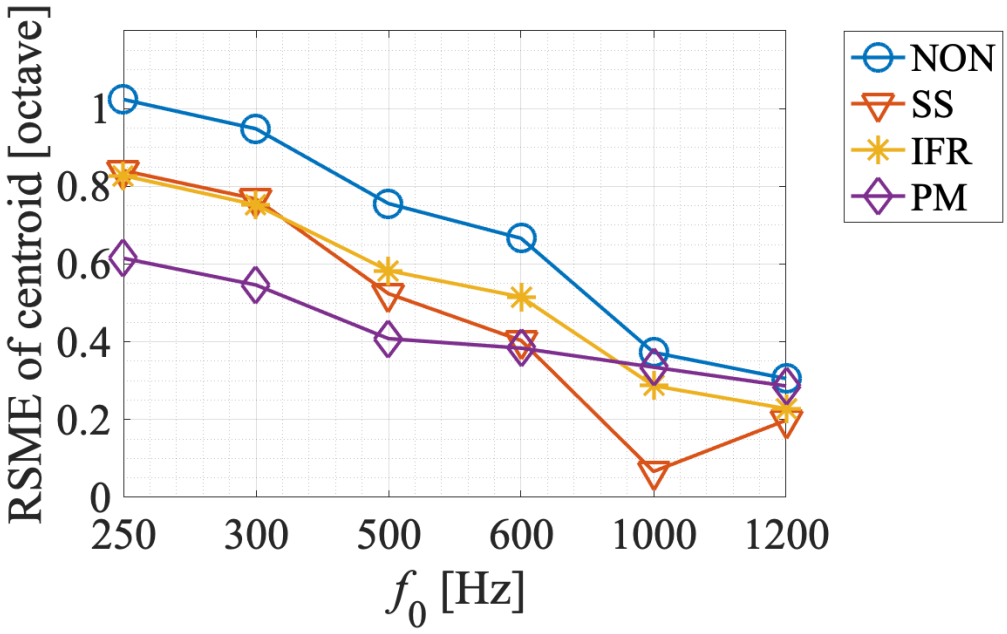


(a) Default voltage

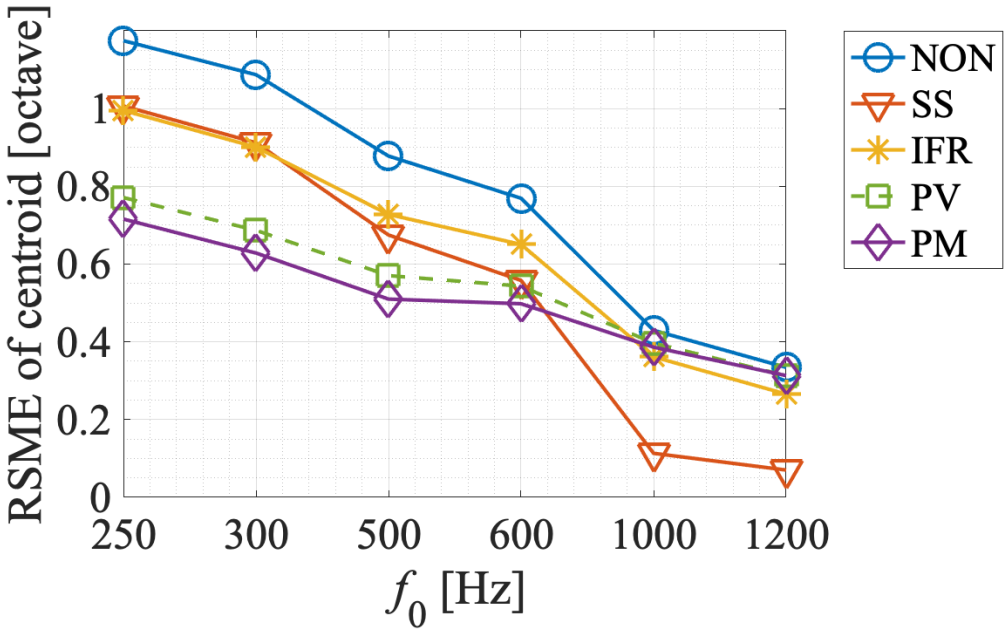


(b) Over-boost voltage

Figure 4.15: Results for RMSE of log-slope (lower is better).

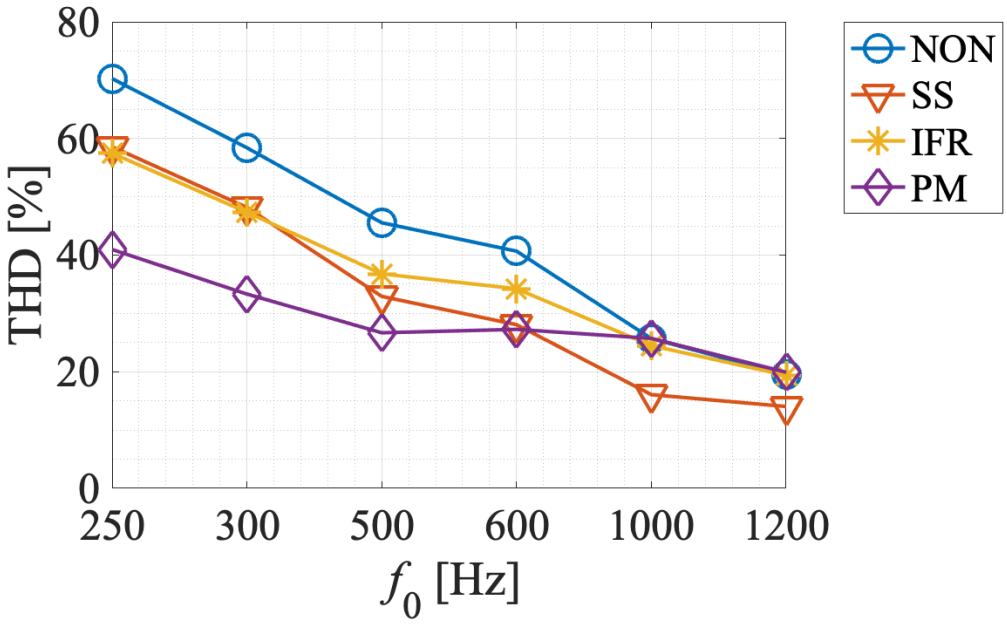


(a) Default voltage

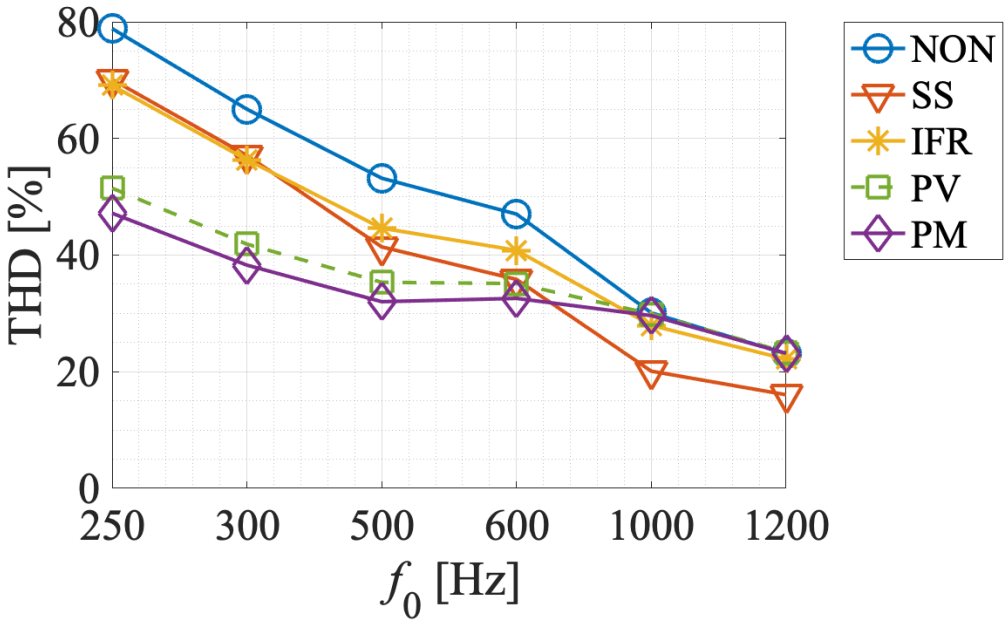


(b) Over-boost voltage

Figure 4.16: Results for RMSE of spectral centroid in octave (lower is better).

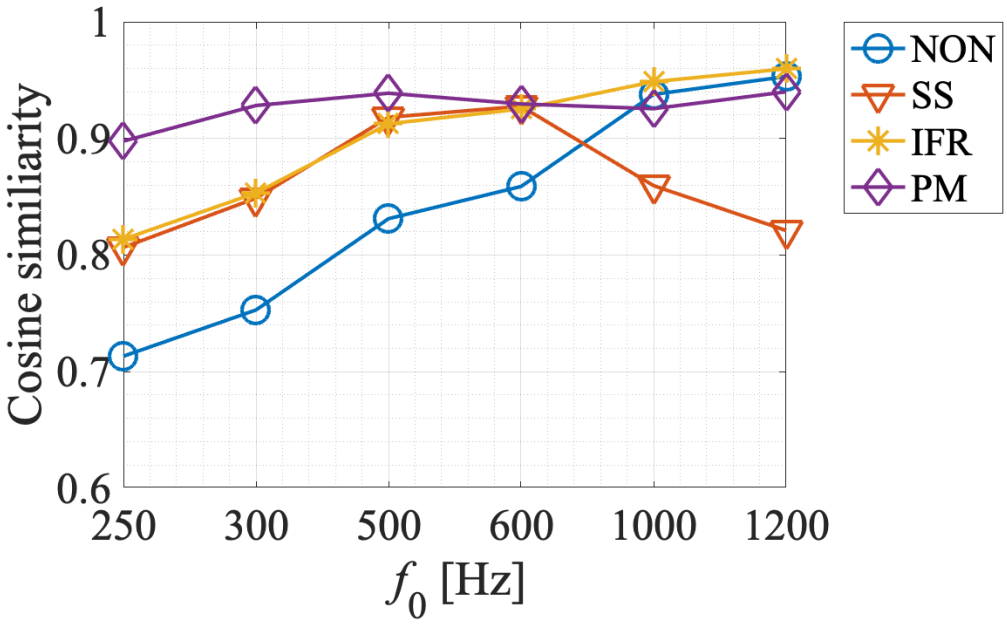


(a) Default voltage

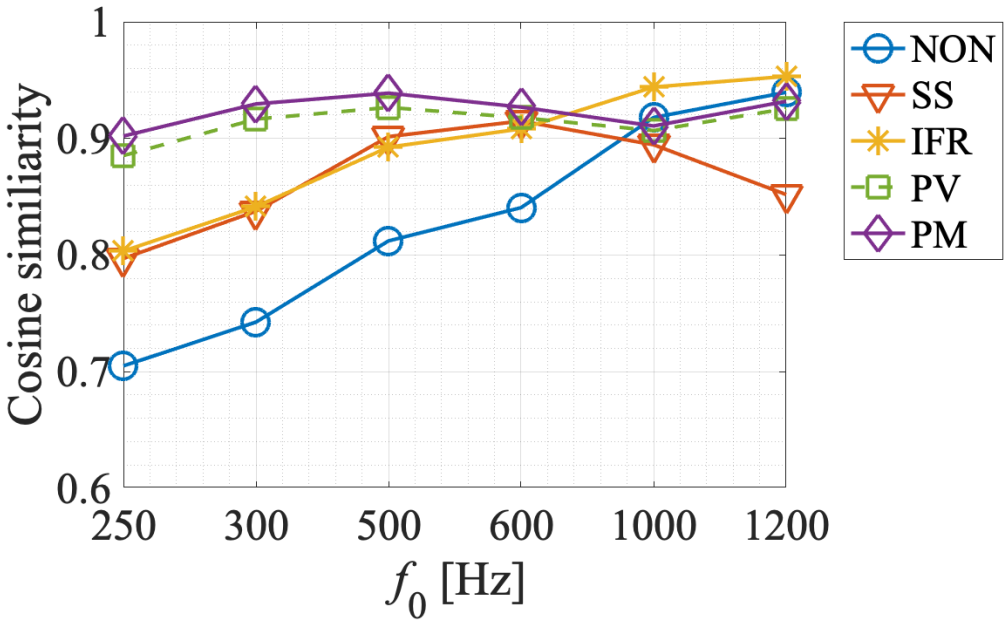


(b) Over-boost voltage

Figure 4.17: Results for THD (lower is better).

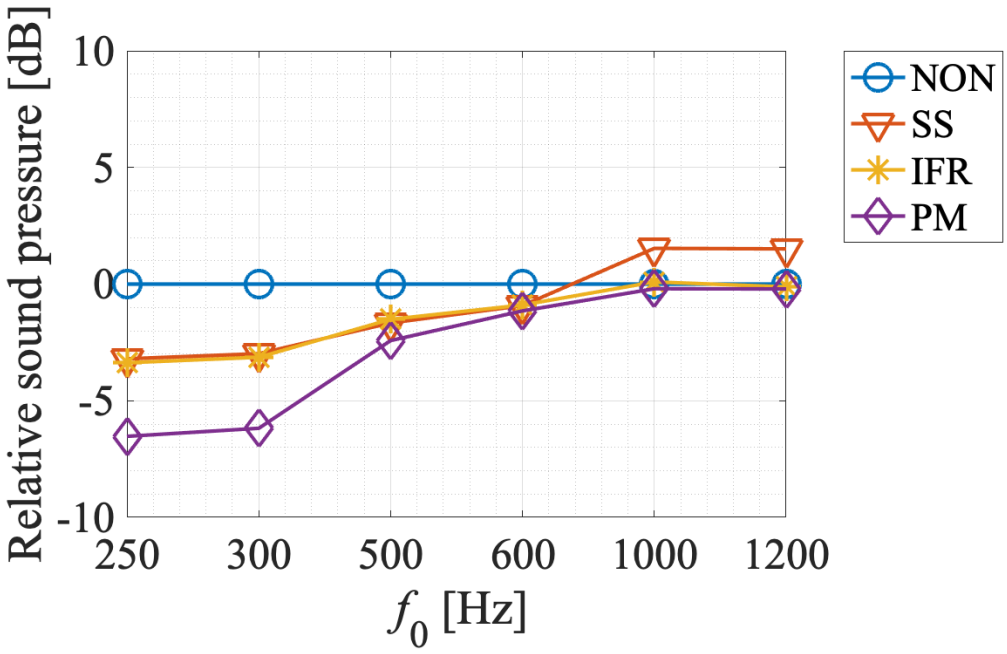


(a) Default voltage

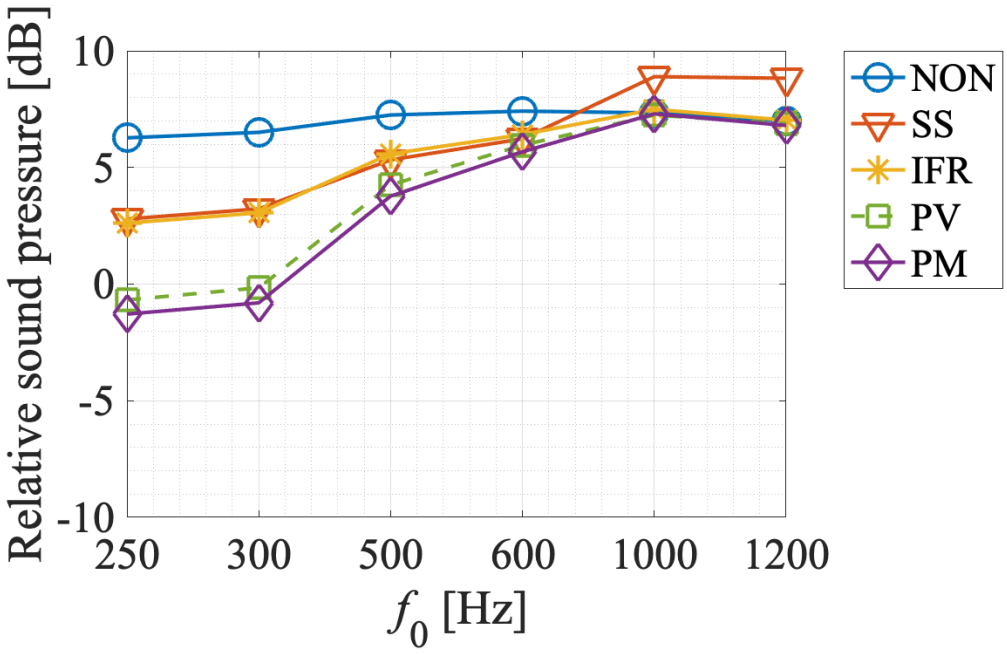


(b) Over-boost voltage

Figure 4.18: Results for cosine similarity (higher is better).



(a) Default voltage



(b) Over-boost voltage

Figure 4.19: Results for demodulated sound pressure (higher is better).

counteract this influence with voltage adaptation. The proposed method shows relatively steadier performance over baseline methods when f_0 varies because f_0-adapted models are constructed in the proposed method. The proposed method shows significant improvement over baseline methods for low-frequency sounds, which are difficult for a PAL to reproduce.

Hence, the effectiveness of the proposed method it is confirmed in improving sound quality and adapting the voltage for harmonic sounds that have the same spectral features as the models. The sound sources used in this experiment were static sounds with fixed slope, fixed number of harmonics. Therefore, for realistic sounds with varying slopes and numbers of harmonics, a further evaluation using sound sources with dynamically changing features is necessary in the future. To apply this method to various sounds, its models should be expanded to involve more patterns for expressing each sound source.

In application, the trade-off between the sound pressure and the sound quality needs to be considered. For sounds with low f_0, the proposed method can be used to improve the sound quality. Over-boosting the PAL can compensate for the lack of sound pressure, and the negative influence of over-boosting can also be counteracted. For sounds with high f_0, the harmonic distortion is reduced, so the inverse frequency response is sufficient to guarantee higher sound quality.

4.5 Conclusion

The reproduction of harmonic sounds with an over-boosted PAL is focused on. The proposed method is voltage-adapted to suppress the influence of boosting the voltage. Two models are employed to adjust the spectral slope and exploit the harmonic distortion. And the effectiveness of the proposed method is demonstrated through evaluation experiments. Especially, the proposed method achieved an average at 0.33 in the error of the log-spectral slope, and an average at 0.53 octave in the error of the spectral centroid. These results suggest that the proposed method can improve the sound quality over conventional methods when reproducing harmonic sounds.

In the future, application to real-world sounds, such as music and speech is necessary. The real signals usually have a complicated structure. So it is considerable to replace the scalar characteristic, the spectral slope with a vector characteristic, the spectral envelope. Another aim should be put on the phase, investigating whether it is predictable, and improving harmonic control by considering the phase. Moreover, under the over-boost voltage condition, it is noticed that low-frequency sounds achieve an equal sound pressure to non-processed sounds, while high-frequency sounds are over-amplified. Automatic gain control to set the optimal voltage should also be considered.

The work presented in this chapter is published in Applied Acoustics [130].

Chapter 5

Directivity Control Based on Phase-inversion for Parametric Array Loudspeaker

5.1 Introduction

Conventional electro-dynamic loudspeaker has a wide directivity commonly, so the spreading sound may become noise for non-listeners. On the other hand, PAL [1, 3, 4, 10] can realize a sharp directivity utilizing ultrasounds. When emitting two ultrasounds intensely in the air, the difference frequency is generated due to the nonlinearity of the air [5]. Utilizing this self-demodulation, it is able to reproduce a target sound by emitting two ultrasounds that the difference frequency is same as the target sound from PAL. The demodulated sound of PAL inherits the straightness of the ultrasounds and achieves the sharp directivity.

A PAL consists of dozens of ultrasonic transducers, and its directivity becomes sharper as the number of ultrasonic transducers increases. This is because the wavefront of the ultrasound is aligned and the parametric effect is strengthened as the array scale increases. Therefore, small-scale PALs tend to have the problem that the beam of demodulated sound spreads. In this study, we aimed to control the beam of demodulated sound for a small-scale PAL.

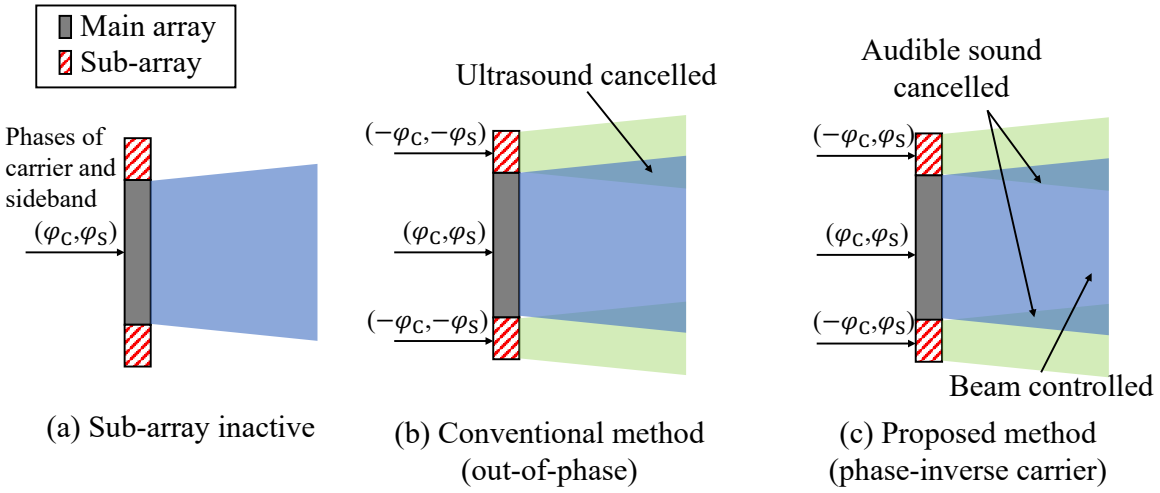
Various conventional approaches can be used to control the directivity of PAL. For example, by choosing a suitable array configuration for the ultrasonic transducers, the demodulated sound

beam can be expanded [31] or focused [131]. Array signal processing approaches have also been proposed, such as beamforming using a delay-and-sum array [132] and beam steering [25,91]. Recently, multiple beamforming by a single PAL has been proposed [96,98].

Alternatively, the parametric effect in PALs can be focused on. If the parametric effect is masked spatially, a narrow-edged beam can be formed. A phase-cancellation excitation has been proposed [123,133]. In this method, two PALs are employed to emit anti-phase ultrasound beams. The sound pressure of ultrasound decreases significantly due to the phase cancellation. Therefore, this method is applied to measurement because the influence of intense ultrasound can be avoided [134]. In this conventional method, the demodulated sound pressure decreases due to the weakening of the parametric effect.

The AM wave emitted from a PAL consists of a carrier wave and a sideband wave. In the conventional method [133], both the carrier and sideband waves are phase-inverse, so the ultrasound is canceled but the demodulated sound retains its original phase. If the phase of either the carrier wave or the sideband wave is inverted, the demodulated sound will have the opposite phase in comparison with the conventional method.

This thesis reviews a narrow-edged beamforming method for PALs. The concept is shown in Fig. 5.1. In this method, PAL is divided into two parts: a main array of ultrasonic transducers in the interior emits a normal AM wave, and a split sub-array of ultrasonic transducers on both sides of the main array emits a phase-inverse AM wave. The characteristics of the modulation method are considered, and suitable signal processing is carried out. The phase-inverse component is canceled spatially to weaken the parametric effect. Moreover, the demodulated sound is canceled spatially to make the beam of demodulated sound narrower. Evaluation experiments were carried out to demonstrate the effectiveness of the proposed method.



(a) Sub-array inactive

(b) Conventional method (out-of-phase)

(c) Proposed method (phase-inverse carrier)

Figure 5.1: Comparison between the proposed method and the conventional method.

5.2 Phase-inversion in Parametric Array Loudspeaker

PAL is an application of a parametric array [4] to reproduce audible sound in the air. The far-field response of a PAL can be predicted by Berktay's model [6], which can be expressed as

$$p(t) \propto \frac{\partial^2 E^2(t)}{\partial t^2}, \tag{5.1}$$

where t denotes the time index, $p(t)$ denotes the demodulated sound, and $E(t)$ denotes the envelope of the emitted ultrasound. To convert an audible target signal to ultrasound, amplitude modulation is carried out. Almost all AM methods wave holds a carrier component, except for methods without carrier such as DSB-SC [48]. Therefore, an AM wave to be emitted from PAL can be indicated as

$$u(t) = c(t) + b(t), \tag{5.2}$$

where $u(t)$ denotes the ultrasonic AM wave, $c(t)$ denotes the carrier wave, and $b(t)$ denotes the sideband wave. The carrier wave $c(t)$ is usually a sinusoidal wave, which can be expressed as

$$c(t) = A_{\mathrm{C}} \cos(2\pi f_{\mathrm{C}} t), \tag{5.3}$$

where A_{C} denotes the amplitude of the carrier wave and f_{C} denotes the carrier frequency. The sideband wave $b(t)$ is dependent on the target signal $s(t)$. When the target signal $s(t)$ is a pure tone, it can be expressed as

$$s(t) = A_{\mathrm{S}} \cos(2\pi f_{\mathrm{S}} t), \tag{5.4}$$

where A_{S} denotes the amplitude of the target signal and f_{S} denotes the frequency of the target signal.

A DSB-modulated wave can be expressed as

$$\mathrm{DSB}[s(t)] = c(t) + m_{\mathrm{DSB}} s(t), \tag{5.5}$$

where $s(t)$ denotes the audible target sound, $c(t)$ denotes the carrier wave, and $m_{\mathrm{DSB}} = A_{\mathrm{S}}/A_{\mathrm{C}}$ ($0 < m \leq 1$) denotes the modulation factor. In (5.5), the first term on the right-hand side is carrier wave (with frequency f_{C}) and the last term on the right-hand side is a sideband wave (frequency $f_{\mathrm{C}} + f_{\mathrm{S}}$, $f_{\mathrm{C}} - f_{\mathrm{S}}$). The envelope of an DSB-modulated wave can be given as

$$\mathcal{E}[\mathrm{DSB}[s(t)]] = 1 + m_{\mathrm{DSB}} s(t). \tag{5.6}$$

Substituting $E(t)$ in Eq. (5.1) with Eq. (5.6), we have

$$p_{\mathrm{DSB}}(t) \propto m_{\mathrm{DSB}} \frac{\partial^2 s(t)}{\partial t^2} + \frac{m_{\mathrm{DSB}}^2}{2} \frac{\partial^2 s^2(t)}{\partial t^2}. \tag{5.7}$$

When the target signal is a pure tone, the first term on the right-hand side of Eq. (5.7) can be regarded as an f^2-weighted target signal (f denotes the frequency index) and the second term can be regarded as a second harmonic distortion.

SSB modulation can be categorized into LSB modulation and USB modulation according to the utilization of sideband. LSB-modulated wave and USB-modulated wave can be expressed as:

$$\mathrm{LSB}[s(t)] = c(t) + m_{\mathrm{SSB}} s(t) c(t) + m_{\mathrm{SSB}} \hat{s}(t) \hat{c}(t), \tag{5.8}$$

$$\mathrm{USB}[s(t)] = c(t) + m_{\mathrm{SSB}} s(t) c(t) - m_{\mathrm{SSB}} \hat{s}(t) \hat{c}(t), \tag{5.9}$$

where $\hat{s}(t)$ denotes signal $s(t)$ after a phase shift of $\pi/2$ and $\hat{c}(t)$ denotes signal $c(t)$ after a phase shift of $\pi/2$. For both LSB modulation and USB modulation, the envelope of the AM wave can be given as

$$\mathcal{E}[\mathrm{LSB}[s(t)]] = \mathcal{E}[\mathrm{USB}[s(t)]] = \sqrt{1 + 2m_{\mathrm{SSB}} s(t) + m_{\mathrm{SSB}}^2 [s^2(t) + \hat{s}^2(t)]} \tag{5.10}$$

when $s(t)$ is a single sinusoid and $s^2(t) + \hat{s}^2(t) = 1$ always holds. In this case, by substituting $E(t)$ in Eq. (5.1) with Eq. (5.10), we have

$$p_{\mathrm{SSB}}(t) \propto 2m_{\mathrm{SSB}} \frac{\partial^2 s(t)}{\partial t^2}. \tag{5.11}$$

It can be noted from (5.11) that the demodulated sound only contains components of the target signal, therefore, SSB modulation has better sound quality than DSB modulation in theory. In practice, the ultrasonic transducer does not hold a flat frequency response, so the characteristics of demodulated sounds may be different between LSB modulation and USB modulation. Moreover, because the sideband frequency differs between LSB modulation and USB modulation, the directivity of demodulated sound may differ as well.

It can be noted that when the phase is inverted in the target signal by replacing $s(t)$ with $-s(t)$ in Eq. (5.7), the phase becomes inverse in the demodulated sound . Based on Eq. (5.5), it is known that phase-inverting only the sideband wave after the modulation is equivalent to inverting the phase of the target signal before the modulation. Considering the symmetry, phase-inverting only the carrier wave can achieve the same result. In the case of SSB modulation, the same conclusion can be drawn. Therefore, a phase-inverse demodulated sound can be obtained by inverting either the sideband wave or the carrier wave.

5.3 Proposed Narrow-edged Beamforming Based on Phase Inversion

Typically, ultrasound with a higher frequency produces a straighter beam. It can be predicted that a beam of demodulated sound will be narrower if the higher frequency component is utilized to form a beam. With USB modulation, the beam is expected to become narrower if the AM wave with a phase-inverse carrier wave is emitted from the sub-array. With LSB modulation, the beam is expected to become narrower if the AM wave with a phase-inverse sideband wave is emitted from the sub-array. This assumption will be proved correct in the experiments.

Based on these assumptions, beamforming method is proposed for PAL to generate a narrow-edged demodulated sound beam. Because the proposed method considers the different characteristics of LSB modulation and USB modulation, different signal processing is performed depending on the modulation. A block diagram of the proposed method with LSB modulation is shown in Fig. 5.2, and a block diagram of the proposed method with USB modulation is shown in Fig. 5.3.

The PAL is divided into a main array that emits a normal AM wave and a sub-array that emits the phase-inverse AM wave in which either the carrier wave or the sideband wave is phase-inverse.

When LSB modulation is used, the AM wave emitted by the main array can be expressed as

$$u_{\mathrm{main}}^{\mathrm{LSB}}(t) = c(t) + b_{\mathrm{LSB}}(t), \tag{5.12}$$

where $b_{\mathrm{LSB}}(t)$ denotes the sideband wave with LSB modulation. Based on (5.2) and (5.8), the sideband wave can be given as

$$b_{\mathrm{LSB}}(t) = m_{\mathrm{SSB}} s(t) c(t) + m_{\mathrm{SSB}} \hat{s}(t) \hat{c}(t). \tag{5.13}$$

The AM wave emitted by the sub-array can be expressed as

$$u_{\mathrm{sub}}^{\mathrm{LSB}}(t) = c(t) - b_{\mathrm{LSB}}(t). \tag{5.14}$$

Compared with $u_{\mathrm{main}}^{\mathrm{LSB}}(t)$, $u_{\mathrm{sub}}^{\mathrm{LSB}}(t)$ consists of an anti-phase sideband wave and an in-phase carrier wave.

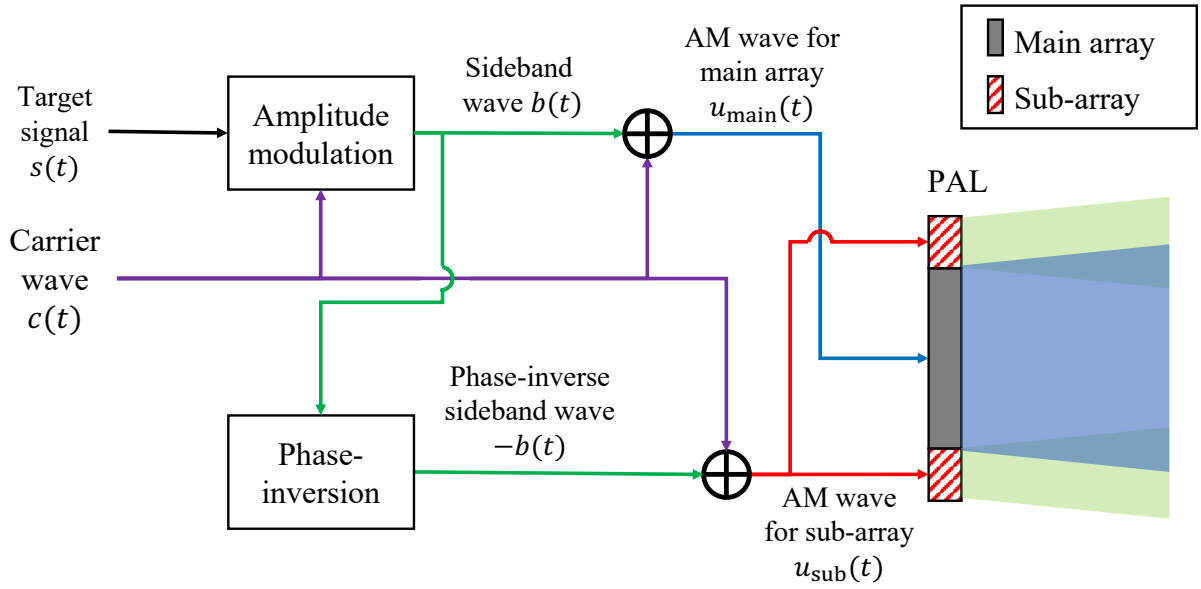


Figure 5.2: Block diagram of the proposed method with LSB modulation.

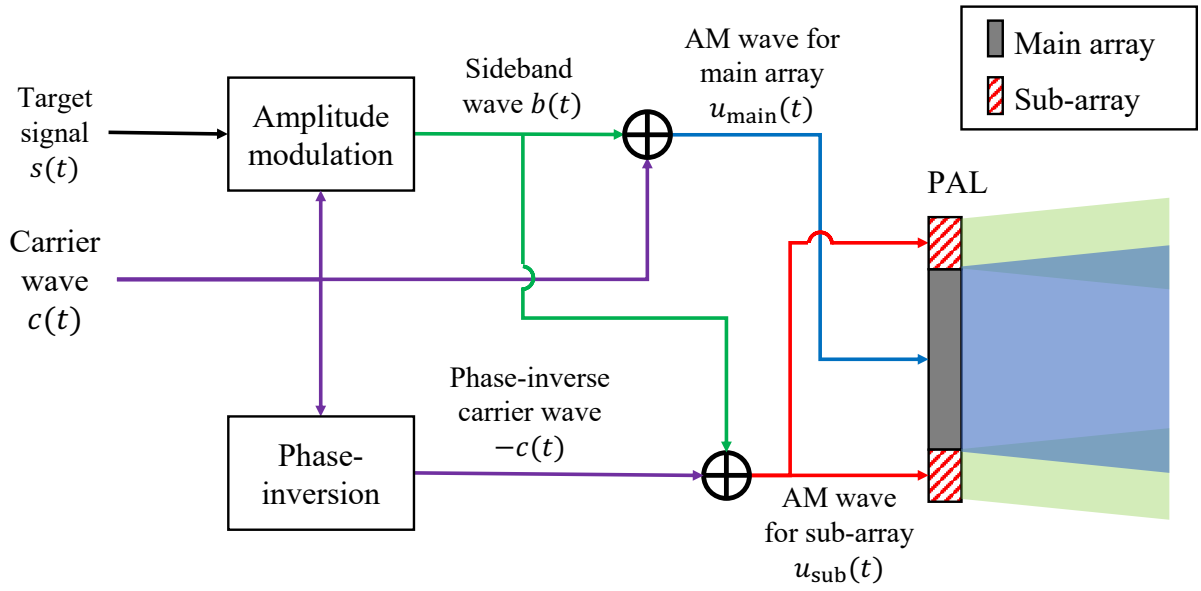


Figure 5.3: Block diagram of proposed method with USB modulation.

When USB modulation is used, the AM wave emitted from the main array can be expressed as

$$u_{\mathrm{main}}^{\mathrm{USB}}(t) = c(t) + b_{\mathrm{USB}}(t), \tag{5.15}$$

where $b_{\mathrm{USB}}(t)$ denotes the sideband wave with USB modulation. Based on (5.2) and (5.9), the sideband waves can be given as

$$b_{\mathrm{USB}}(t) = m_{\mathrm{SSB}} s(t) c(t) - m_{\mathrm{SSB}} \hat{s}(t) \hat{c}(t). \tag{5.16}$$

The AM wave emitted by the sub-array can be expressed as

$$u_{\mathrm{sub}}^{\mathrm{USB}}(t) = -c(t) + b_{\mathrm{USB}}(t). \tag{5.17}$$

Compared with $u_{\mathrm{main}}^{\mathrm{USB}}(t)$, $u_{\mathrm{sub}}^{\mathrm{USB}}(t)$ consists of an in-phase sideband wave and an anti-phase carrier wave.

5.4 Evaluation Experiment

The effectiveness of the proposed method was demonstrated experimentally. The characteristics of demodulated sound beams under different conditions were measured to verify that the beam can be controlled.

5.4.1 Experimental Conditions

The experimental conditions and experimental equipment are listed in Tables 5.1 and 5.2, respectively. The PAL used in the experiment consisted of 150 ultrasonic transducers. As shown in Fig. 5.4, the main array employed 90 transducers in the interior and the sub-array employed 60 transducers, 30 on each side of the main array. As shown in Figs. 5.2 and 5.3, a normal AM wave was emitted from the main array, and a phase-inverse AM wave was emitted from the sub-array. In the experiment, these two AM waves were emitted with the same power.

Table 5.1: Experimental conditions.

Environment	Office room
Ambient noise level	$L_A = 29.7$ dB
Reverberation time	$T_{60} = 650$ ms
Sampling frequency	192 kHz
Quantization	32 bit
Sound source	TSP signal (0–8 kHz)
Modulation factor	$m_{\mathrm{SSB}} = 1.0$
Carrier frequency	$f_{\mathrm{C}} = 40$ kHz

A TSP signal [116] of 0–8 kHz was used as sound source. LSB modulation and USB modulation were carried out with the same carrier frequency and modulation factor shown in Table 5.1.

For comparison, the signal emitted from the main array $u_{\mathrm{main}}(t)$ is held constant while the signal emitted from the sub-array $u_{\mathrm{sub}}(t)$ changes by inverting the phase of the carrier wave or the sideband wave. 4 cases were applied in the experiment:

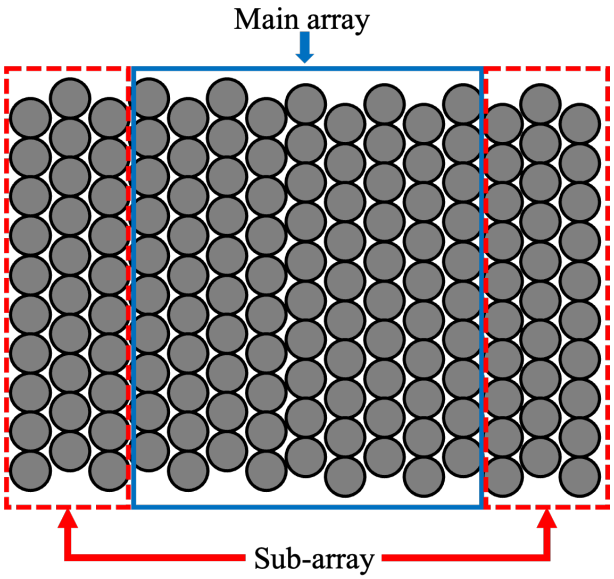


Figure 5.4: Structure of parametric array loudspeaker used in experiment.

- The first case is in-phase (IP), where the signal emitted by the sub-array is identical to that of the main array, $u_{\mathrm{sub}}(t) = u_{\mathrm{main}}(t)$.
- The second case is anti-phase (AP), where the signal emitted by the sub-array is exactly inverse to that of the main array, $u_{\mathrm{sub}}(t) = -u_{\mathrm{main}}(t)$.
- The third case is anti-phase carrier wave with in-phase sideband wave (ACIS), where the signal emitted by the sub-array has a carrier wave with an inverse phase and a sideband wave with the same phase compared with the signal emitted by the main array, $u_{\mathrm{sub}}(t) = -c(t) + b(t)$ and $u_{\mathrm{main}}(t) = c(t) + b(t)$.
- The last case is in-phase carrier wave with anti-phase sideband wave (ICAS), where the

Table 5.2: Experimental equipment.

Microphone	SONY, ECM-88B
Microphone amplifier	THINKNET, MA-2016C
Ultrasonic transducer	SPL Limited, UT1007-Z325R (150 pcs.)
A/D, D/A converter	RME, FIREFACE UFX
Power amplifier	VICTOR, PS-A2002

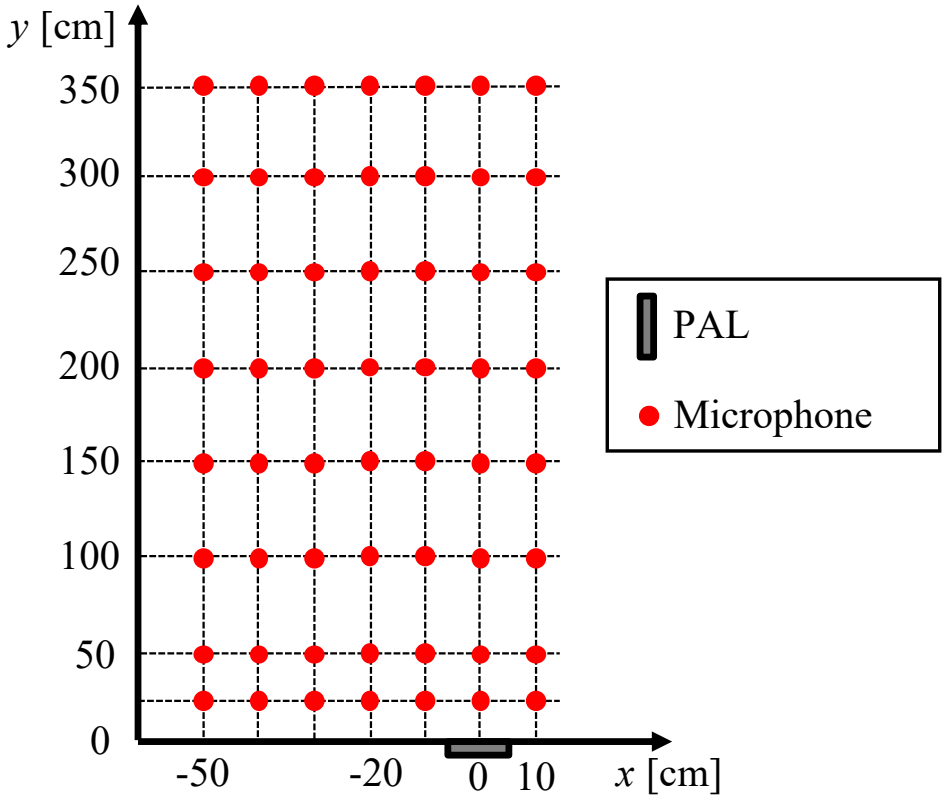


Figure 5.5: Experimental arrangement.

signal emitted by the sub-array has a carrier wave with the same phase and a sideband wave with an inverse phase compared with the signal emitted by the main array, $u_{\mathrm{sub}}(t) = c(t) - b(t)$ and $u_{\mathrm{main}}(t) = c(t) + b(t)$.

As discussed in Section 5.3, the proposed method with LSB modulation and that with USB modulation are different. Here, when LSB modulation is used, ICAS is the same as proposed method. In contrast, when USB modulation is used, ACIS is the same as proposed method.

The experimental arrangement is shown in Fig. 5.5. Microphones were used to record the demodulated sound of the PAL. On the x-axis, the observation positions were set at intervals of 10 cm. On the y-axis, the observation positions were set at intervals of 50 cm. The ultrasonic components were recorded by the microphones, and the observed signals were processed with a low-pass filter to eliminate the ultrasound beyond 20 kHz before analysis.

5.4.2 Experimental Results

The demodulated sound pressure distributions under LSB and USB modulation are shown in Figs. 5.6 and 5.7, respectively. The corresponding beam patterns of each case are shown in Figs. 5.8 and 5.9. In these figures, the maximum sound pressure is normalized to 0 dB, and the contour lines of -3, -6, ..., -15 dB are drawn.

In Figs. 5.6 and 5.8, it can be seen that under LSB modulation, AP and ACIS show a wider but shorter beam, while ICAS (the proposed method) shows a narrower beam compared with IP. Under USB modulation, ACIS (the proposed method) shows a narrower beam, as seen in Figs. 5.7 and 5.9.

The demodulated sound pressure distributions on the radiation axis (x = 0 cm) are shown in Fig. 5.10. It can be seen that IP achieves the highest demodulated sound pressure with both LSB and USB modulation. Inverting the phase of the AM wave decreases the on-axis sound pressure. However, due to the straightness of ultrasound beams, the on-axis sound pressure does not decrease very much. It can be seen in Fig. 5.10 (a) that ICAS causes a smaller decrease in the demodulated sound pressure, while in Fig. 5.10 (b), ACIS causes a smaller decrease in the demodulated sound pressure. This is because that both ICAS under LSB modulation and ACIS under USB modulation invert the AM wave component with a lower frequency.

The -6 dB beamwidths are shown in Fig. 5.11. Here, the beamwidth is calculated as the distance at which the sound pressure decays by 6 dB from the on-axis sound pressure. A larger beamwidth suggests that the beam spreads more. It can be seen that AP achieves a wider beam than IP under both LSB and USB modulation. Because ultrasound is phase-canceled, the parametric effect is weakened and the beam of demodulated sound becomes wider. Under LSB modulation, it can also be seen in Fig. 5.10 (a) that ICAS (the proposed method) achieves an equal or narrower beam. Under USB modulation, it can be seen in Fig. 5.10 (b) that ACIS (the proposed method) achieves an equal or narrower beam compared with IP. In the proposed method, either the carrier wave or the sideband wave is phase-inverted, so the phase-inversed component is canceled and the parametric effect is weakened. Also, the phase of demodulated sound becomes inverted, which spatially cancels the demodulated sound.

As discussed in Section 5.3, the carrier wave has a higher frequency with LSB modulation, and the sideband wave has a higher frequency with USB modulation. It is supposed that inverting

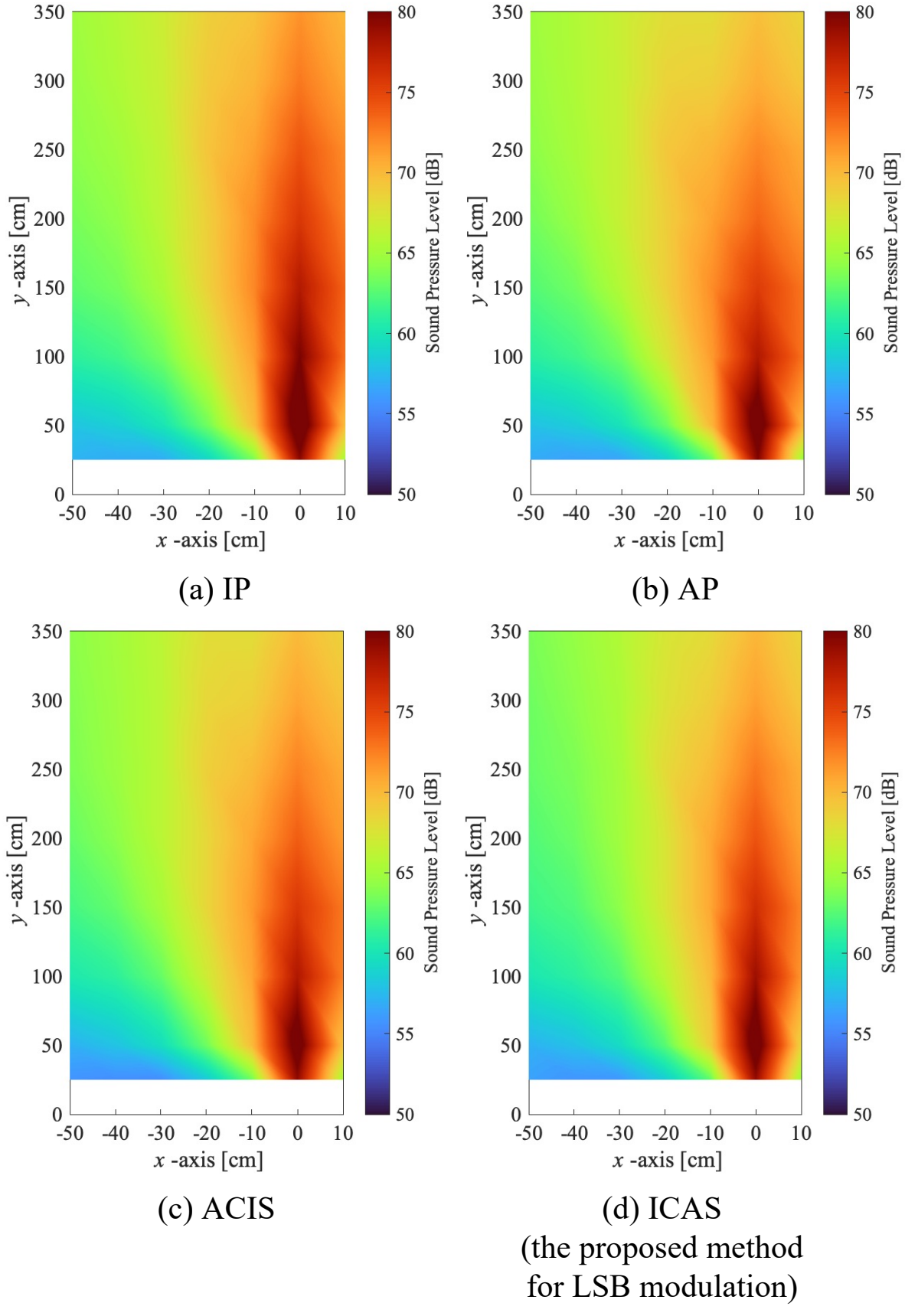
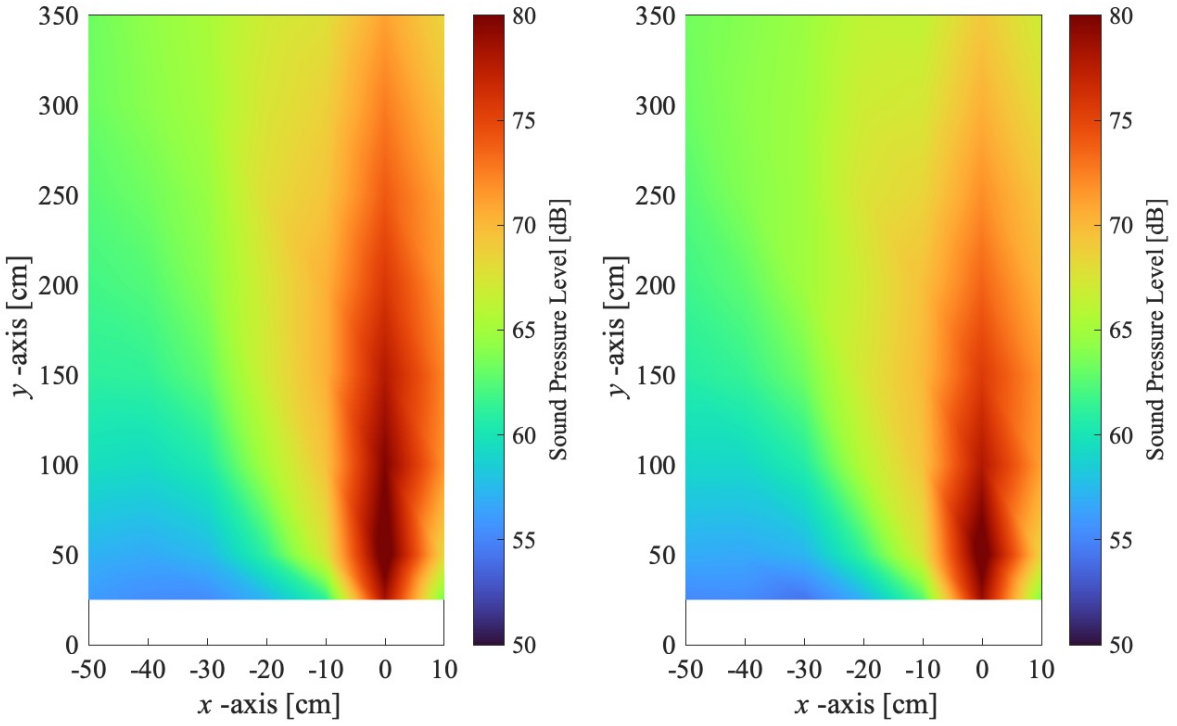


Figure 5.6: Demodulated sound pressure distribution under LSB modulation.



(a) IP

(b) AP

(c) ACIS
(the proposed method
for USB modulation)

(d) ICAS

Figure 5.7: Demodulated sound pressure distribution under USB modulation.

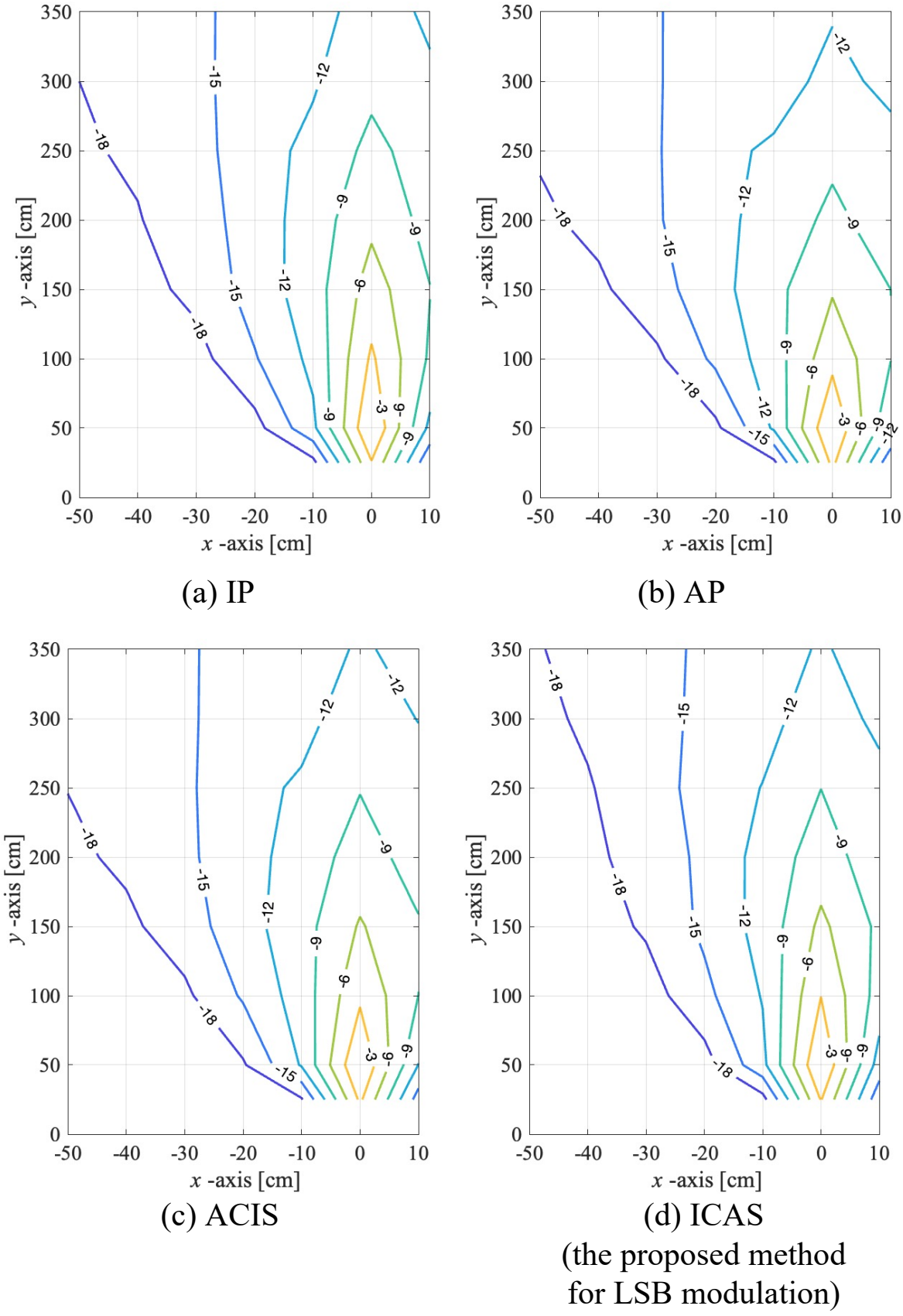
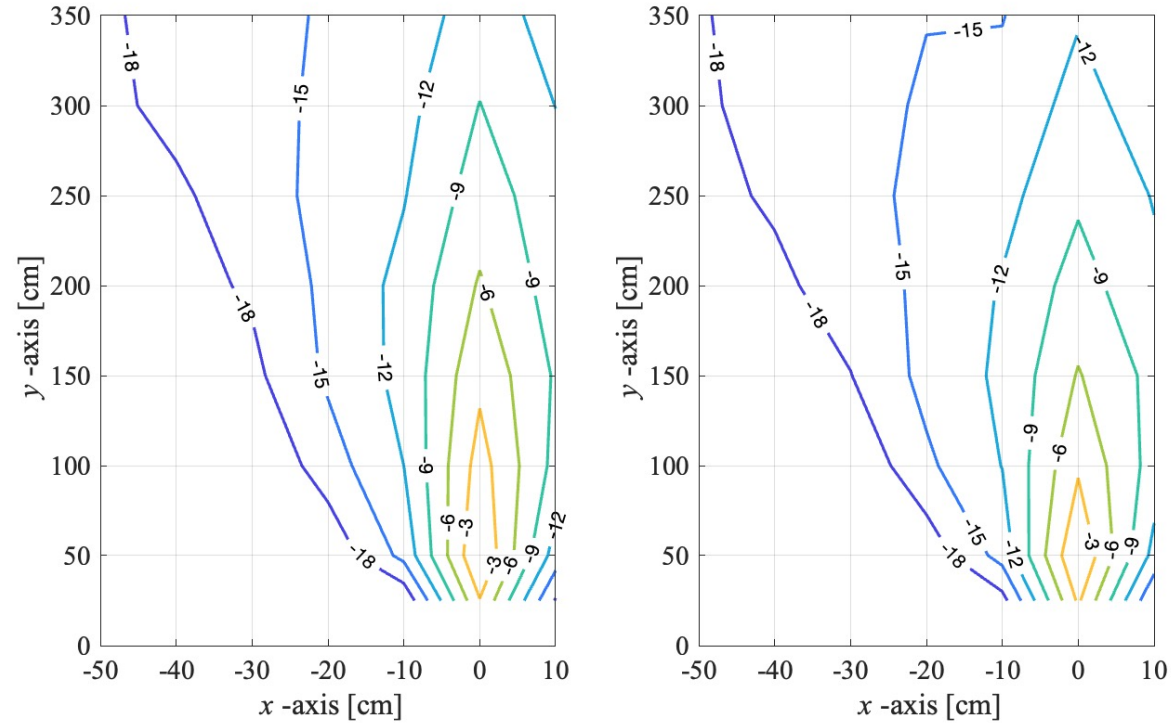


Figure 5.8: Beam pattern of the demodulated sound under LSB modulation.



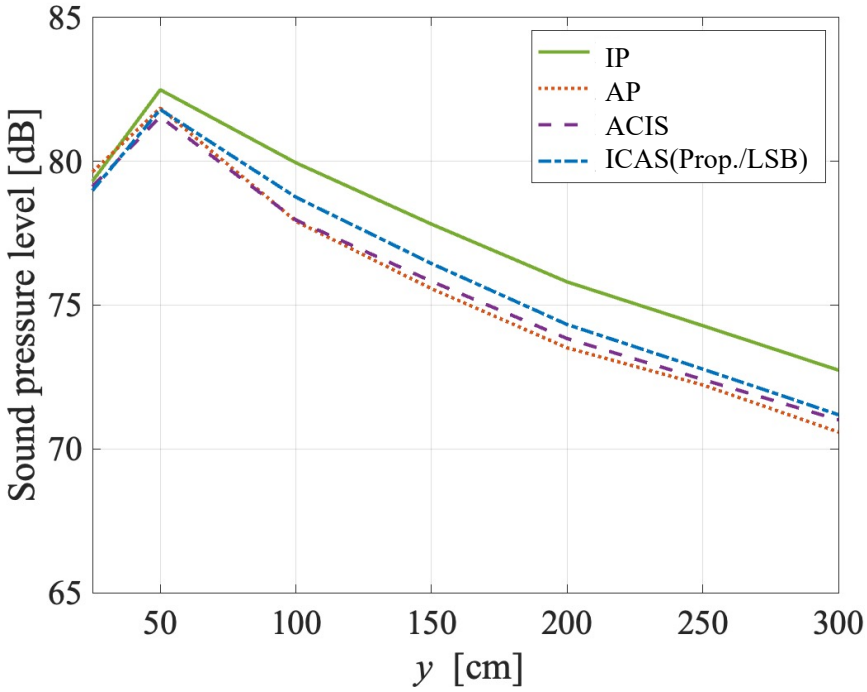
(a) IP

(b) AP

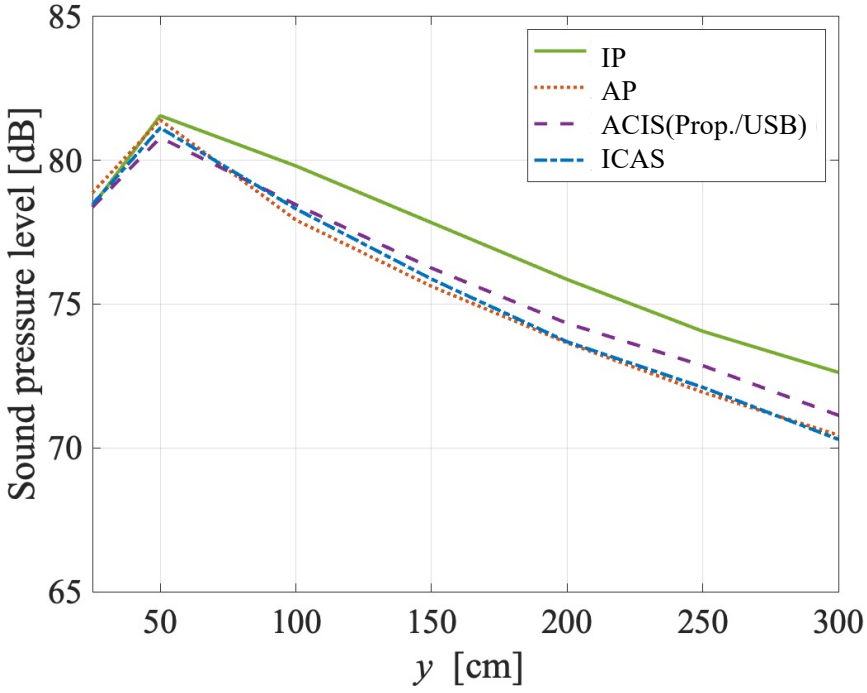
(c) ACIS
(the proposed method
for USB modulation)

(d) ICAS

Figure 5.9: Beam pattern of the demodulated sound under USB modulation.

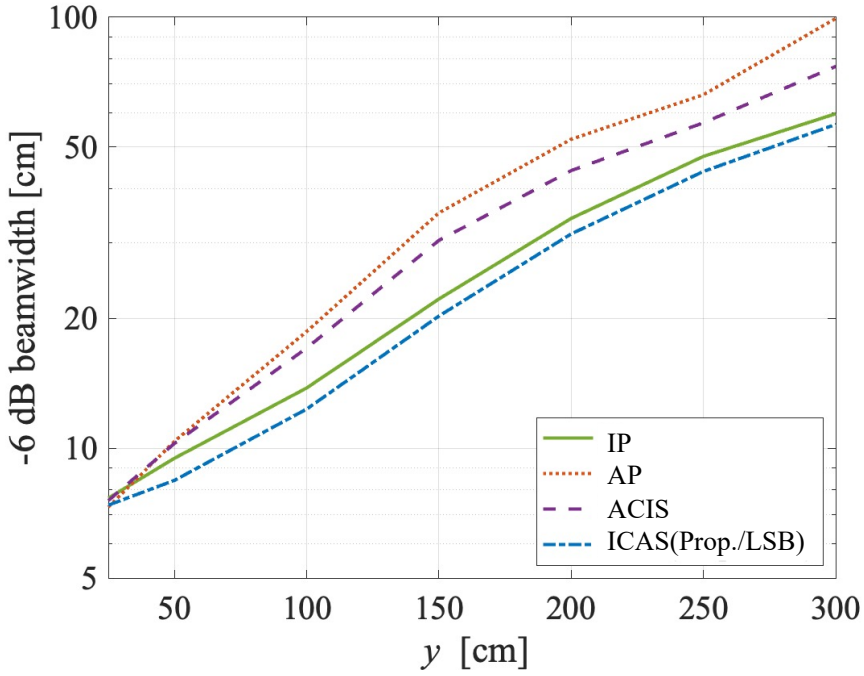


(a) LSB modulation

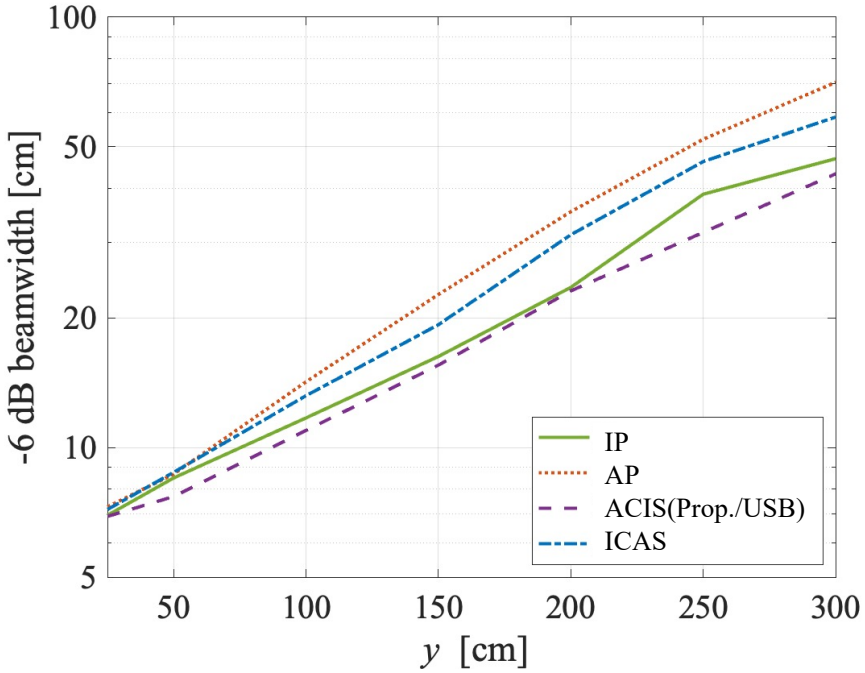


(b) USB modulation

Figure 5.10: Sound pressure distribution on the radiation axis ($x = 0$ cm).



(a) LSB modulation



(b) USB modulation

Figure 5.11: Beamwidth of the demodulated sound.

the phase of the lower-frequency component and holding the phase of the higher- frequency component will result in a narrower beam. The assumption was proved experimentally by the

results showing that inverting the carrier wave under LSB modulation or inverting the sideband wave under USB modulation can narrow the beam of the demodulated sound. In the experiment, the same driving voltages are applied in the main array and the sub-array. If the sub-array is enlarged, or the driving voltage on the sub-array is increased, the performance of the proposed method is expected to improve.

5.5 Conclusion

A narrow-edged beamforming method is proposed for PAL based on parametric effect. By inverting the phase of either the carrier wave or sideband wave, both the ultrasound and the demodulated sound are canceled, which narrows the beam of demodulated sound. The experimental results demonstrated the effectiveness of the proposed method. In future work, it is considerable to investigate the influence of the size of the sub-array and the power level driving the signal emitted from the sub-array to control the beam even more.

The work presented in this chapter is published in the proceeding of APSIPA-ASC 2021, © 2021 IEEE. [135].

Chapter 6

Conclusions and Future Works

6.1 Conclusions

PAL is renowned for projecting directional acoustic beams in the air due to its sharp directivity. Therefore, the practical application of PAL is in strong demand in the construction of the next-generation community. Compared with conventional electro-dynamic loudspeakers, PAL has a poorer sound quality because of the nonlinearity. Moreover, the directivity control needs to be investigated aiming for practical use. This thesis focused on the sound quality and directivity control of PAL and introduced two methods for sound quality improvement along with one method for directivity control.

The first proposed method which was introduced in Chapter 3 aims at realizing a flat frequency response by enhancing the demodulated sound pressure in low-frequency. It is known theoretically that it is difficult for PAL to reproduce low-frequency, therefore, a multi-way structure is introduced to PAL. The proposed multi-way PAL consists of Woofer PAL for low-frequency reproduction and Tweeter PAL for high-frequency reproduction. It is known that the demodulated sound pressure shows a positive relevance to the scale of PAL, i.e. the number of ultrasonic transducers. Facing the difficulty of low-reproduction, more ultrasonic transducers are employed in Woofer PAL. Suitable modulation methods are selected for each part of the multi-way PAL. Especially for Woofer PAL, the harmonic distortions are utilized efficiently to enhance the low-frequency. Measurements on the prototype multi-way PAL show that the proposed method realized a flat frequency response in a wide band when reproducing wideband

signals.

The second proposed method which was introduced in Chapter 4 aims at realizing a high fidelity of spectral characteristics when reproducing harmonic sounds like a speech signal or music signal. Typically, a harmonic sound is composed of several components whose frequency is in multiples of the fundamental frequency. The amount of components leads to a complicated condition in the generation of nonlinear distortion. Therefore, the difference of spectrum between the demodulated sound and the original signal was investigated in an overall way and summarized into models. For small-scale PAL, the most efficient way to realize a flat frequency response at a sufficient sound pressure is to increase the driving voltage, or over-boosting the PAL in other words. Here appears another issue to tackle the nonlinearity that varies as the driving voltage. Therefore, this proposed method takes the driving voltage into consideration to provide optimized parameters to counteract the influence of over-boosting. The effectiveness has been demonstrated by the experimental results, which show that the proposed method can help demodulated sound better restore the spectral slope of the original target signal and bring less harmonic distortion.

In addition, a novel directivity control was introduced in Chapter 5. PAL shows a tendency that the directivity gets sharper as the scale of PAL increases. It suggests a latent problem that a small-scale PAL may suffer a diffuse acoustic beam. This method puts the focus on the parametric effect, which is the fundamental principle of PAL. If the parametric effect can be weakened spatially, the directivity will be controlled. In the proposed method, PAL is decomposed into a main array and a sub-array. Phase-inverse ultrasounds are emitted from the sub-array to weaken the parametric effect. In order to achieve a narrower beam, the strategies of phase inversion are designed for different modulation methods. The measured results validated that the proposed method can achieve a narrower acoustic beam.

6.2 Future Works

Through the study on the sound quality improvement and directivity control for PAL introduced in this thesis, several tasks remain and can be explored in the future.

For example, though the multi-way PAL introduced in Chapter 3 can achieve a sufficient enhancement for wideband signals, the performance on harmonic sound needs to be improved. Moreover, the prototype built for the evaluation experiments is on a relatively large scale, therefore, the downscaling of the device aiming for practical use is in demand. In addition, the hardware arrangement can be studied further. In this proposed method, the focal point is set to meet the condition that the distances to Woofer and Tweeter PAL are the same. In the future, the phase characteristic and wavefront synthesis should also be considered in the design of hardware arrangement.

It has been validated that the proposed spectral slope control introduced in Chapter 4 can restore the spectrum with higher fidelity than conventional methods. However, the performance is limited to sounds with a linear envelope, which can be expressed by a one-dimension scalar spectral slope. In the future, application to real-world sounds with complicated structures is necessary. So the spectral slope should be expanded to the spectral envelope. Moreover, automatic gain control should also be considered to achieve sound pressure without being affected by the frequency of the target signal.

In the proposed directivity control based on phase inversion which was introduced in Chapter 5, the phenomenon has been confirmed, however, part of the mechanism needs further investigation. Aiming to the practical use, how the scale and emission power of the main and sub-array influence the beam width is also needed to be investigated.

List of publications

Journal Paper

1. **Yuting Geng**, Masato Nakayama, and Takanobu Nishiura, “Demodulated sound quality improvement for harmonic sounds in over-boosted parametric array loudspeaker,” Applied Acoustics, vol. 168, pp. 108460, 2022. (Accepted, Impact factor: 2.639)

2. **Yuting Geng**, Haonan Wang, Masato Nakayama, and Takanobu Nishiura, “Realization of Normalized Frequency Characteristics in Wide Band for Parametric Array Loudspeaker Based on Multi-way Structure,” IEICE Transactions on Information and Systems, vol. J-105D, no. 3, 2022 (in Japanese). (Accepted)

International Conference

1. **Yuting Geng**, Haonan Wang, Masato Nakayama, and Takanobu Nishiura, “Narrow-Edged Beamforming Using Masked Parametric Array Loudspeakers," APSIPA-ASC 2021, Tokyo, Japan, Dec. 2021. (Accepted)

2. **Yuting Geng**, Masato Nakayama and Takanobu Nishiura, “Evaluation on Parametric Array Loudspeaker with Virtual Bass Enhancer Utilized Harmonic Distortions," ICSV27, Paper ID: 1150, Online, Jul. 2021.

3. **Yuting Geng**, Masato Nakayama and Takanobu Nishiura, “Evaluation of a Multi-way Parametric Array Loudspeaker Based on Multiplexed Double Sideband Modulation," APSIPA-ASC 2020, pp. 409-415, Online, Dec. 2020.

4. **Yuting Geng**, Yusei Nakano, Masato Nakayama, and Takanobu Nishiura, "Development of Multi-way Parametric Array Loudspeaker Using Multiplexed Double Sideband Modulation," ICA 2019, pp. 1069-1076, Aachen, Germany, Sep. 2019.

5. **Yuting Geng**, Masato Nakayama, and Takanobu Nishiura, "Virtual Bass Enhancement Based on Harmonics Control Using Missing Fundamental in Parametric Array Loudspeaker," INTER-NOISE 2019, Paper ID: 1517, Madrid, Spain, Jun. 2019.

Technical Report

1. **Yuting Geng**, Haonan Wang, Masato Nakayama, and Takanobu Nishiura, "Beam Control Based on Multi-phased Array for Parametric Array Loudspeaker", Technical report of IEICE, vol.121, no.112, EA2021-5, pp. 24-29, Online, Jul. 2021 (in Japanese).

Domestic Conference

1. **Yuting Geng**, Haonan Wang, Masato Nakayama, and Takanobu Nishiura, "A Study on Demodulated Sound Beamwidth Control Using Phase-inverse Carrier in Parametric Array Loudspeaker," pp. 247-250, Online, Sep. 2021 (in Japanese).

2. **Yuting Geng**, Masato Nakayama, and Takanobu Nishiura, "Sound Quality Improvement Based on Spectral Slope Control for Harmonic Structure in Parametric Array Loudspeaker," pp. 271-272, Online, Mar. 2021 (in Japanese).

3. **Yuting Geng**, Masato Nakayama, and Takanobu Nishiura, "Speech-enhanced Parametric Array Loudspeaker Using Formant Enhancement Based on Emitted Energy Optimization," pp. 229-230, Online, Sep. 2020 (in Japanese).

4. **Yuting Geng**, Masato Nakayama, and Takanobu Nishiura, "A Study on Sound Enhancement for Music Using Missing Fundamental in Parametric Array Loudspeaker," pp. 395-396, Saitama, Mar. 2020 (in Japanese).

5. **Yuting Geng**, Masato Nakayama, and Takanobu Nishiura, "A Study on Sound Enhancement for Parametric Array Loudspeaker Based on Missing Fundamental Using Harmonic

Distortions," pp. 715-716, Shiga, Sep. 2019 (in Japanese).

6. **Yuting Geng**, Masato Nakayama, and Takanobu Nishiura, "A Study on Sound Enhancement for Parametric Array Loudspeaker Based on Quality Compensation with Missing Fundamental and Low-frequency Suppression for Dry Source," pp. 351-354, Tokyo, Mar. 2019 (in Japanese).

Award

1. **Yuting Geng**, 20th Student Presentation Award, Acoustical Society of Japan, 2019.

Bibliography

[1] K. Aoki, T. Kamakura, and Y. Kumamoto, "Parametric loudspeaker – characteristics of acoustic field and suitable modulation of carrier ultrasound," Electronics and Communications in Japan (Part III: Fundamental Electronic Science), vol.74, no.9, pp.76–82, 1991.

[2] F.J. Pompei, Sound from ultrasound: The parametric array as an audible sound source, Ph.D. thesis, Massachusetts Institute of Technology, 2002.

[3] W.S. Gan, J. Yang, and T. Kamakura, "A review of parametric acoustic array in air," Applied Acoustics, vol.73, no.12, pp.1211–1219, 2012.

[4] P.J. Westervelt, "Parametric acoustic array," The Journal of the Acoustical Society of America, vol.35, no.4, pp.535–537, 1963.

[5] A. Thuras, R. Jenkins, and H. O'Neil, "Extraneous frequencies generated in air carrying intense sound waves," The Journal of the Acoustical Society of America, vol.6, no.3, pp.173–180, 1935.

[6] H. Berktay, "Possible exploitation of non-linear acoustics in underwater transmitting applications," Journal of Sound and Vibration, vol.2, no.4, pp.435–461, 1965.

[7] U. Ingard and D.C. Pridmore-Brown, "Scattering of sound by sound," The Journal of the Acoustical Society of America, vol.28, no.3, pp.367–369, 1956.

[8] J. Bellin, P.J. Westervelt, and R. Beyer, "Experimental investigation of a parametric end array," The Journal of the Acoustical Society of America, vol.32, no.7, pp.935–935, 1960.

[9] J. Bellin and R. Beyer, "Experimental investigation of an end-fire array," The Journal of the Acoustical Society of America, vol.34, no.8, pp.1051–1054, 1962.

[10] M. Yoneyama, J. Fujimoto, Y. Kawamo, and S. Sasabe, "The audio spotlight: An application of nonlinear interaction of sound waves to a new type of loudspeaker design," The Journal of the Acoustical Society of America, vol.73, no.5, pp.1532–1536, 1983.

[11] J.J. Croft, "Theory, history, and the advancement of parametric loudspeaker," White paper of American Technology Corporation, 2001.

[12] S. Sakai and T. Kamakura, "Dynamic single sideband modulation for realizing parametric loudspeaker," AIP Conference Proceedings, pp.613–616, American Institute of Physics, 2008.

[13] Y. Wang, X. Li, L. Xu, and L. Xu, "Ssb modulation of the ultrasonic carrier for a parametric loudspeaker," 2009 International Conference on Electronic Computer Technology, pp.669–673, IEEE, 2009.

[14] M.F. Hamilton, D.T. Blackstock, *et al.*, Nonlinear acoustics, Academic press San Diego, 1998.

[15] K. Aoki, T. Kamakura, and Y. Kumamoto, "A parametric loudspeaker—applied examples," Electronics and Communications in Japan (Part III: Fundamental Electronic Science), vol.77, no.1, pp.64–74, 1994.

[16] T. Miyachi, J.J. Balvig, W. Kisada, K. Hayakawa, and T. Suzuki, "A quiet navigation for safe crosswalk by ultrasonic beams," International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, pp.1049–1057, Springer, 2007.

[17] K. Ishii, Y. Yamamoto, M. Imai, and K. Nakadai, "A navigation system using ultrasonic directional speaker with rotating base," Symposium on Human Interface and the Management of Information, pp.526–535, Springer, 2007.

[18] W.S. Gan, E.L. Tan, and S.M. Kuo, "Audio projection," IEEE Signal Processing Magazine, vol.28, no.1, pp.43–57, 2010.

[19] S. Aoki, M. Toba, and N. Tsujita, "Sound localization of stereo reproduction with parametric loudspeakers," Applied Acoustics, vol.73, no.12, pp.1289–1295, 2012.

[20] Y. Sugibayashi, S. Kurimoto, D. Ikefuji, M. Morise, and T. Nishiura, "Three-dimensional acoustic sound field reproduction based on hybrid combination of multiple parametric loudspeakers and electrodynamic subwoofer," Applied Acoustics, vol.73, no.12, pp.1282–1288, 2012.

[21] E.L. Tan, W.S. Gan, and C.H. Chen, "Spatial sound reproduction using conventional and parametric loudspeakers," Proceedings of The 2012 Asia Pacific Signal and Information Processing Association Annual Summit and Conference, pp.1–9, IEEE, 2012.

[22] S. Aoki, K. Shimizu, and K. Itou, "Study of vertical sound image control with parametric loudspeakers," Applied Acoustics, vol.116, pp.164–169, 2017.

[23] N. Shimada, K. Iwai, M. Nakayama, and T. Nishiura, "High-presence sharp sound image based on sound blending using parametric and dynamic loudspeakers," Journal of Signal Processing, vol.24, no.4, pp.171–174, 2020.

[24] L.A. Brooks, A.C. Zander, and C.H. Hansen, "Investigation into the feasibility of using a parametric array control source in an active noise control system," Proceedings of ACOUSTICS, pp.39–45, 2005.

[25] N. Tanaka and M. Tanaka, "Active noise control using a steerable parametric array loudspeaker," The journal of the acoustical society of America, vol.127, no.6, pp.3526–3537, 2010.

[26] T. Komatsuzaki and Y. Iwata, "Active noise control using high-directional parametric loudspeaker," Journal of Environment and Engineering, vol.6, no.1, pp.140–149, 2011.

[27] D. Ikefuji, T. Matsui, M. Nakayama, T. Nishiura, and Y. Yamashita, "Fundamental study on active noise control with audio-spot for minimum area using parametric loudspeaker," INTER-NOISE and NOISE-CON Congress and Conference Proceedings, pp.2524–2529, Institute of Noise Control Engineering, 2015.

[28] K. Tanaka, C. Shi, and Y. Kajikawa, "Binaural active noise control using parametric array loudspeakers," Applied Acoustics, vol.116, pp.170–176, 2017.

[29] S. Nakagawa, Y. Okamoto, and Y. Fujisaka, "Development of a bone-conducted ultrasonic hearing aid for the profoundly sensorineural deaf," Transactions of Japanese Society for Medical and Biological Engineering, vol.44, no.1, pp.184–189, 2006.

[30] H. Wang, T. Fujii, M. Nakayama, and T. Nishiura, "A wearable parametric array loudspeaker demodulating inside head," 27th International Congress on Sound and Vibration, Paper ID:1194, 2021.

[31] D. Ikefuji, H. Tsujii, S. Masunaga, M. Nakayama, T. Nishiura, and Y. Yamashita, "Reverberation steering and listening area expansion on 3-d sound field reproduction with parametric array loudspeaker," 2014 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA-ASC 2014), pp.1–5, IEEE, 2014.

[32] J. Donley, C. Ritz, and W.B. Kleijn, "Reproducing personal sound zones using a hybrid synthesis of dynamic and parametric loudspeakers," 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA-ASC 2016), pp.1–5, IEEE, 2016.

[33] S. Sayama, Y. Ogami, M. Nakayama, and T. Nishiura, "Acoustic space-sharing with multiple parametric array loudspeakers for daily sports and exercise," Forum Acusticum 2020, pp.2599–2605, 2020.

[34] Y. Ochiai, T. Hoshi, and I. Suzuki, "Holographic whisper: Rendering audible sound spots in three-dimensional space by focusing ultrasonic waves," Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, pp.4314–4325, 2017.

[35] S. Komori, Y. Ogami, M. Nakayama, and T. Nishiura, "Acoustic hologram with multiple parametric array loudspeakers," IEICE Transactions on Information and Systems (Japanese Edition), vol.J101-D, no.3, pp.578–587, 2018.

[36] Y. Ogami, M. Nakayama, and T. Nishiura, "Virtual sound source construction based on radiation direction control using multiple parametric array loudspeakers," The Journal of the Acoustical Society of America, vol.146, no.2, pp.1314–1325, 2019.

[37] S. Lee, T. Katsuura, T. Towatari-Ueno, and Y. Shimomura, "The effects of parametric speaker sound on salivary hormones and a subjective evaluation," Neuroendocrinology Letters, vol.31, no.4, p.524, 2010.

[38] B.W. Lawton, Damage to human hearing by airborne sound of very high frequency or ultrasonic frequency, Health & Safety Executive, 2001.

[39] I.O. Wygant, M. Kupnik, J.C. Windsor, W.M. Wright, M.S. Wochner, G.G. Yaralioglu, M.F. Hamilton, and B.T. Khuri-Yakub, "50 khz capacitive micromachined ultrasonic transducers for generation of highly directional sound with parametric arrays," IEEE transactions on ultrasonics, ferroelectrics, and frequency control, vol.56, no.1, pp.193–203, 2009.

[40] P. Ji, E.L. Tan, W.S. Gan, and J. Yang, "A comparative analysis of preprocessing methods for the parametric loudspeaker based on the khokhlov–zabolotskaya–kuznetsov equation for speech reproduction," IEEE Transactions on audio, speech, and language processing, vol.19, no.4, pp.937–946, 2010.

[41] T. Kamakura, M. Yoneyama, and K. Ikegaya, "Developments of parametric loudspeaker for practical use," Proc. 10th International Symposium on Non-linear Acoustics, 1984, pp.147–150, 1984.

[42] T. Kamakura, M. Yoneyama, and K. Ikegaya, "Studies for the realization of parametric loud-speaker," Journal of the Acoustical Society of Japan, vol.41, no.6, pp.378–385, 1985.

[43] T.D. Kite, J.T. Post, and M.F. Hamilton, "Parametric array in air: Distortion reduction by preprocessing," Proc. 16th int. cong. acoust, vol.2, pp.1091–1092, 1998.

[44] F.J. Pompei, "The use of airborne ultrasonics for generating audible sound beams," Journal of the Audio Engineering Society, vol.47, no.9, pp.726–731, 1999.

[45] P. Ji, W.S. Gan, E.L. Tan, and J. Yang, "Performance analysis on recursive single-sideband amplitude modulation for parametric loudspeakers," 2010 IEEE International Conference on Multimedia and Expo, pp.748–753, IEEE, 2010.

[46] E.L. Tan, P. Ji, and W.S. Gan, "On preprocessing techniques for bandlimited parametric loudspeakers," Applied Acoustics, vol.71, no.5, pp.486–492, 2010.

[47] C. Shi and W.S. Gan, "A preprocessing method to increase high frequency response of a parametric loudspeaker," 2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, pp.1–5, IEEE, 2013.

[48] C. Shi and Y. Kajikawa, "A comparative study of preprocessing methods in the parametric loudspeaker," 2014 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA-ASC 2014), pp.1–5, IEEE, 2014.

[49] D. Ikefuji, M. Nakayama, T. Nishiura, and Y. Yamashita, "Weighted double sideband modulation toward high quality audible sound on parametric loudspeaker," 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp.843–847, IEEE, 2013.

[50] Y. Nakano, T. Fukumori, M. Nakayama, and T. Nishiura, "A study on audible low-frequency sound emphasis based on multiplexed double sideband modulation in parametric loudspeaker," INTER-NOISE and NOISE-CON Congress and Conference Proceedings, pp.2054–2065, Institute of Noise Control Engineering, 2018.

[51] Y. Hatano, C. Shi, and Y. Kajikawa, "Compensation for nonlinear distortion of the frequency modulation-based parametric array loudspeaker," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol.25, no.8, pp.1709–1717, 2017.

[52] T. Iwasaki, D. Ikefuji, M. Nakayama, and T. Nishiura, "Parametric loudspeaker for speech signal based on the combination of amplitude and frequency modulations," Proceedings of 21st International Congress on Acoustics (ICA2013), p.055047, Acoustical Society of America, 2013.

[53] M. Nakayama and T. Nishiura, "Synchronized amplitude-and-frequency modulation for a parametric loudspeaker," 2017 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA-ASC 2017), pp.130–135, IEEE, 2017.

[54] K. Mori, T. Fukumori, M. Nakayama, and T. Nishiura, "Spectral peak noise reduction with frequency modulated carrier wave for parametric loudspeaker," INTER-NOISE and

NOISE-CON Congress and Conference Proceedings, pp.2066–2076, Institute of Noise Control Engineering, 2018.

[55] K. Mori, M. Nakayama, and T. Nishiura, "Evaluation of thermal runaway control based on frequency modulated carrier wave in parametric array loudspeaker," Proceedings of the 23rd International Congress on Acoustics (ICA 2019), pp.1045–1052, 2019.

[56] K. Asahi, K. Mori, M. Nakayama, and T. Nishiura, "Spectral peak noise reduction based on frequency-modulated carrier wave with temporal adaptation for parametric array loudspeaker," Forum Acusticum, pp.2745–2752, 2020.

[57] K. Asahi, K. Mori, M. Nakayama, and T. Nishiura, "Audible sound enhancement and spectral peak noise reduction based on advanced carrier for parametric loudspeaker," IEICE Transactions on Information and Systems (Japanese Edition), vol.J104-D, no.4, pp.186–197, 2021.

[58] K. Asahi, M. Nakayama, and T. Nishiura, "Audible sound enhancement based on carrier with periodic square-wave for parametric loudspeaker," 2020 IEEE 9th Global Conference on Consumer Electronics (GCCE), pp.12–15, IEEE, 2020.

[59] C. Shi and Y. Kajikawa, "Effect of the ultrasonic emitter on the distortion performance of the parametric array loudspeaker," Applied Acoustics, vol.112, pp.108–115, 2016.

[60] R.D. Nowak and B.D. Van Veen, "Random and pseudorandom inputs for volterra filter identification," IEEE Transactions on Signal Processing, vol.42, no.8, pp.2124–2135, 1994.

[61] R.D. Nowak and B.D. Van Veen, "Volterra filter equalization: A fixed point approach," IEEE Transactions on signal processing, vol.45, no.2, pp.377–388, 1997.

[62] W. Ji and W.S. Gan, "Identification of a parametric loudspeaker system using an adaptive volterra filter," Applied Acoustics, vol.73, no.12, pp.1251–1262, 2012.

[63] Y. Mu, P. Ji, W. Ji, M. Wu, and J. Yang, "On the linearization of a parametric loudspeaker system by using third-order inverse volterra filters," 2013 IEEE China Summit

and International Conference on Signal and Information Processing, pp.570–574, IEEE, 2013.

[64] Y. Mu, P. Ji, W. Ji, M. Wu, and J. Yang, "Modeling and compensation for the distortion of parametric loudspeakers using a one-dimension volterra filter," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol.22, no.12, pp.2169–2181, 2014.

[65] C. Shi and Y. Kajikawa, "Identification of the parametric array loudspeaker with a volterra filter using the sparse nlms algorithm," 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.3372–3376, IEEE, 2015.

[66] C. Shi and Y. Kajikawa, "Volterra model of the parametric array loudspeaker operating at ultrasonic frequencies," The Journal of the Acoustical Society of America, vol.140, no.5, pp.3643–3650, 2016.

[67] C. Shi and Y. Kajikawa, "Synthesis of volterra filters for the parametric array loudspeaker," 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.4229–4233, IEEE, 2016.

[68] C. Shi and Y. Kajikawa, "Automatic gain control for parametric array loudspeakers," 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp.589–593, IEEE, 2016.

[69] J.C.R. Licklider, "A duplex theory of pitch perception," The Journal of the Acoustical Society of America, vol.23, no.1, pp.147–147, 1951.

[70] D.R. Chialvo, "How we hear what is not there: A neural mechanism for the missing fundamental illusion," Chaos: An Interdisciplinary Journal of Nonlinear Science, vol.13, no.4, pp.1226–1230, 2003.

[71] T. Matsuoka and K. Ito, "Perception of missing fundamental and consideration on its characteristics," Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp.2059–2062, IEEE, 2003.

[72] E. Larsen and R.M. Aarts, "Reproducing low-pitched signals through small loudspeakers," Journal of the Audio Engineering Society, vol.50, no.3, pp.147–164, 2002.

[73] C. Shi, H. Mu, and W.S. Gan, "A psychoacoustical preprocessing technique for virtual bass enhancement of the parametric loudspeaker," 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp.31–35, IEEE, 2013.

[74] Y. Geng, M. Nakayama, and T. Nishiura, "Virtual bass enhancement based on harmonics control using missing fundamental in parametric array loudspeaker," INTER-NOISE and NOISE-CON Congress and Conference Proceedings, pp.2660–2671, Institute of Noise Control Engineering, 2019.

[75] Y. Geng, M. Nakayama, and T. Nishiura, "Evaluation on parametric array loudspeaker with virtual bass enhancer utilized harmonic distortions,," 27th International Congress on Sound and Vibration (ICSV27), (Paper ID: 1150), 2021.

[76] S. Tao, J. Ren, and C. Shi, "Virtual bass preprocessing and carrier frequency optimization for the parametric array loudspeaker," INTER-NOISE and NOISE-CON Congress and Conference Proceedings, pp.1676–1682, Institute of Noise Control Engineering, 2021.

[77] C. Shi, Y. Kajikawa, and W.S. Gan, "An overview of directivity control methods of the parametric array loudspeaker," APSIPA Transactions on Signal and Information Processing, vol.3, e20, pp.1–12.

[78] D. Olszewski, F. Prasetyo, and K. Linhard, "Steerable highly directional audio beam loudspeaker," Interspeech, pp.137–146, 2005.

[79] M. Morise, D. Ikefuji, H. Tujii, K. Hirokawa, and T. Nishiura, "A design of reflective audio spot with reflective objects," Proceedings of the 20th International Congress on Acoustics (ICA 2010), pp.3003–3006, 2010.

[80] S. Masunaga, D. Ikefuji, M. Nakayama, and T. Nishiura, "Steering for listening area of reflective audio spot with parametric loudspeaker array," Proceedings of 21st International Congress on Acoustics (ICA2013), p.055026, Acoustical Society of America, 2013.

[81] S. Komori, D. Ikefuji, T. Fukumori, M. Nakayama, and T. Nishiura, "Design of moving sound image on wall with curved-type parametric loudspeaker," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences (Japanese Edition), vol.J99-A, no.11, pp.1–4, 2016.

[82] J. Zhong, S. Wang, R. Kirby, and X. Qiu, "Reflection of audio sounds generated by a parametric array loudspeaker," The Journal of the Acoustical Society of America, vol.148, no.4, pp.2327–2336, 2020.

[83] S. Komori, S. Masunaga, D. Ikefuji, M. Nakayama, and T. Nishiura, "Control of audio spot with flexible parametric loudspeaker," Transactions of the Virtual Reality Society of Japan, vol.20, no.3, pp.189–198, 2015.

[84] D. Ikefuji, M. Nakayama, T. Nishiura, and Y. Yamashita, "Robust sound image localization for moving listener with curved-type parametric loudspeaker," 2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA-ASC 2015), pp.1045–1049, IEEE, 2015.

[85] U. Sayin, P. Artís, and O. Guasch, "Realization of an omnidirectional source of sound using parametric loudspeakers," The Journal of the Acoustical Society of America, vol.134, no.3, pp.1899–1907, 2013.

[86] O. Guasch and P. Sánchez-Martín, "Far-field directivity of parametric loudspeaker arrays set on curved surfaces," Applied Mathematical Modelling, vol.60, pp.721–738, 2018.

[87] M. Arnela, O. Guasch, P. Sánchez-Martín, J. Camps, R.M. Alsina-Pagès, and C. Martínez-Suquía, "Construction of an omnidirectional parametric loudspeaker consisting in a spherical distribution of ultrasound transducers," sensors, vol.18, no.12, p.4317, 2018.

[88] Y. Zhang, L. Wang, L. Qin, C. Zhong, and S. Hao, "Spherical–omnidirectional piezoelectric composite transducer for high-frequency underwater acoustics," IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol.68, no.5, pp.1791–1796, 2020.

[89] T. Matsui, D. Ikefuji, M. Nakayama, and T. Nishiura, "Audio spot design based on separating emission of carrier and sideband waves," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences (Japanese Edition), vol.J97-A, no.4, pp.304–312, 2014.

[90] T. Matsui, D. Ikefuji, M. Nakayama, and T. Nishiura, "Multiple audio spots design based on separating emission of carrier and sideband waves," INTER-NOISE and NOISE-CON

Congress and Conference Proceedings, pp.2924–2928, Institute of Noise Control Engineering, 2014.

[91] W.S. Gan, J. Yang, K.S. Tan, and M.H. Er, "A digital beamsteerer for difference frequency in a parametric array," IEEE Transactions on audio, speech, and language processing, vol.14, no.3, pp.1018–1025, 2006.

[92] C. Shi and W.S. Gan, "Analysis and calibration of system errors in steerable parametric loudspeakers," Applied acoustics, vol.73, no.12, pp.1263–1270, 2012.

[93] S. Wu, M. Wu, C. Huang, and J. Yang, "Fpga-based implementation of steerable parametric loudspeaker using fractional delay filter," Applied Acoustics, vol.73, no.12, pp.1271–1281, 2012.

[94] C. Shi, Investigation of the steerable parametric loudspeaker based on phased array techniques, Ph.D. thesis, Nanyang Technological University, Singapore, 2013.

[95] C. Shi, H. Nomura, T. Kamakura, and W.S. Gan, "Development of a steerable stereophonic parametric loudspeaker," 2014 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA-ASC 2014), pp.1–5, IEEE, 2014.

[96] C. Shi, Y. Kajikawa, and W.S. Gan, "Generating dual beams from a single steerable parametric loudspeaker," Applied Acoustics, vol.99, pp.43–50, 2015.

[97] K. Nakagawa, C. Shi, and Y. Kajikawa, "Beam steering of portable parametric array loudspeaker," 2019 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), pp.1824–1827, IEEE, 2019.

[98] C. Shi, R. Bai, J. Gou, and J. Liang, "Multi-beam design method for a steerable parametric array loudspeaker," 2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA-ASC 2020), pp.416–420, IEEE, 2020.

[99] M. Nakayama, T. Nishiura, R. Okuno, and N. Nakasako, "A study on acoustic beam steering with parametric loudspeaker based on individual delay-filtering for carrier and sideband waves," 2014 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), pp.447–450, IEEE, 2014.

[100] M. Nakayama, R. Konabe, T. Fukumori, and T. Nishiura, "Near-sound-field propagation based on individual beam-steering for carrier and sideband waves with parametric array loudspeaker," 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), pp.1–8, IEEE, 2016.

[101] C. Shi, H. Nomura, T. Kamakura, and W.S. Gan, "Spatial aliasing effects in a steerable parametric loudspeaker for stereophonic sound reproduction," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, vol.97, no.9, pp.1859–1866, 2014.

[102] C. Shi and W.S. Gan, "Grating lobe elimination in steerable parametric loudspeaker," IEEE transactions on ultrasonics, ferroelectrics, and frequency control, vol.58, no.2, pp.437–450, 2011.

[103] C. Shi and W.S. Gan, "Product directivity models for parametric loudspeakers," The Journal of the Acoustical Society of America, vol.131, no.3, pp.1938–1945, 2012.

[104] C. Shi and Y. Kajikawa, "A convolution model for computing the far-field directivity of a parametric loudspeaker array," The Journal of the Acoustical Society of America, vol.137, no.2, pp.777–784, 2015.

[105] O. Guasch, P. Sánchez-Martín, and M. Arnela, "Convolution model for the far-field directivity of curved parametric loudspeaker arrays," INTER-NOISE and NOISE-CON Congress and Conference Proceedings, pp.1735–1745, Institute of Noise Control Engineering, 2017.

[106] K.S. Tan, W.S. Gan, J. Yang, and H.E. Meng, "Constant beamwidth beamformer for difference frequency in parametric array," 2003 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp.V–361, IEEE, 2003.

[107] C. Shi and R. Bai, "Design of a constant beamwidth beamformer for the parametric array loudspeaker," Proceedings of the 23rd International Congress on Acoustics (ICA 2019), pp.3077–3081, 2019.

[108] J. Liang, J. Gou, R. Bai, and C. Shi, "Design of a constant beamwidth beamformer for the length-limited parametric array loudspeaker," INTER-NOISE and NOISE-CON

Congress and Conference Proceedings, pp.3658–3664, Institute of Noise Control Engineering, 2020.

[109] C. Hedberg, K. Haller, and T. Kamakura, "A self-silenced sound beam," Acoustical Physics, vol.56, no.5, pp.637–639, 2010.

[110] H. Nomura, C.M. Hedberg, and T. Kamakura, "Numerical simulation of parametric sound generation and its application to length-limited sound beam," Applied Acoustics, vol.73, no.12, pp.1231–1238, 2012.

[111] E. Skinner, M. Groves, and M.K. Hinders, "Demonstration of a length limited parametric array," Applied Acoustics, vol.148, pp.423–433, 2019.

[112] Y. Geng, Y. Nakano, M. Nakayama, and T. Nishiura, "Development of multi-way parametric array loudspeaker using multiplexed double sideband modulation," Proceedings of the 23rd International Congress on Acoustics (ICA 2019), pp.1069–1076, 2019.

[113] Y. Geng, M. Nakayama, and T. Nishiura, "Evaluation of a multi-way parametric array loudspeaker based on multiplexed double sideband modulation," 2020 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA-ASC 2020), pp.409–415, IEEE, 2020.

[114] R. San Martín Murugarren, P. Tello, A. Valencia Leoz, and A. Marzo Pérez, "Experimental evaluation of distortion in amplitude modulation techniques for parametric loudspeakers," Applied Sciences, vol.10, no.6, pp.2070–2081, 2020.

[115] T. Fujii, K. Ariyoshi, M. Nakayama, and T. Nishiura, "Control of maximum demodulation distance in parametric loudspeaker using variable-type acoustic lens with gas layer," IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences (Japanese Edition), vol.J102-A, no.12, pp.299–309, 2019.

[116] N. Aoshima, "Computer-generated pulse signal applied for sound measurement," The Journal of the Acoustical Society of America, vol.69, no.5, pp.1484–1488, 1981.

[117] N. Moriya and Y. Kaneda, "Study of harmonic distortion on impulse response measurement with logarithmic time stretched pulse," Acoustical science and technology, vol.26, no.5, pp.462–464, 2005.

[118] S. Masunaga, D. Ikefuji, M. Morise, M. Nakayama, and T. Nishimura, "Harmonic distortion measurement for a parametric loudspeaker with logarithmic time stretched pulse," Proc. Acoust. ' 12, 2012.

[119] Y. Geng, H. Wang, M. Nakayama, and T. Nishiura, "Realization of normalized frequency characteristics in wide band for parametric array loudspeaker based on multi-way structure," IEICE Transactions on Information and Systems (Japanese Edition), vol.J105-D, no.3, 2022.

[120] G. Peeters, B.L. Giordano, P. Susini, N. Misdariis, and S. McAdams, "The timbre toolbox: Extracting audio descriptors from musical signals," The Journal of the Acoustical Society of America, vol.130, no.5, pp.2902–2916, 2011.

[121] G.B. Stan, J.J. Embrechts, and D. Archambeau, "Comparison of different impulse response measurement techniques," Journal of the Audio engineering society, vol.50, no.4, pp.249–262, 2002.

[122] P. Ji, W. Hu, and J. Yang, "Development of an acoustic filter for parametric loudspeaker using phononic crystals," Ultrasonics, vol.67, pp.160–167, 2016.

[123] P. Ji, W. Liu, S. Wu, J. Yang, and W.S. Gan, "An alternative method to measure the on-axis difference-frequency sound in a parametric loudspeaker without using an acoustic filter," Applied acoustics, vol.73, no.12, pp.1244–1250, 2012.

[124] H.S. Ju and Y.H. Kim, "Near-field characteristics of the parametric loudspeaker using ultrasonic transducers," Applied Acoustics, vol.71, no.9, pp.793–800, 2010.

[125] J.M. Grey and J.W. Gordon, "Perceptual effects of spectral modifications on musical timbres," The Journal of the Acoustical Society of America, vol.63, no.5, pp.1493–1500, 1978.

[126] R. McAulay and T. Quatieri, "Speech analysis/synthesis based on a sinusoidal representation," IEEE Transactions on Acoustics, Speech, and Signal Processing, vol.34, no.4, pp.744–754, 1986.

[127] Y. Stylianou, "Harmonic plus noise models for speech, combined with statistical methods, for speech and speaker modification," Ph. D thesis, Ecole Nationale Superieure des Telecommunications, 1996.

[128] Y. Pantazis, O. Rosec, and Y. Stylianou, "Adaptive am–fm signal decomposition with application to speech analysis," IEEE Transactions on Audio, Speech, and Language Processing, vol.19, no.2, pp.290–300, 2010.

[129] G. Degottex and Y. Stylianou, "Analysis and synthesis of speech using an adaptive full-band harmonic model," IEEE Transactions on Audio, Speech, and Language Processing, vol.21, no.10, pp.2085–2095, 2013.

[130] Y. Geng, M. Nakayama, and T. Nishiura, "Demodulated sound quality improvement for harmonic sounds in over-boosted parametric array loudspeaker," Applied Acoustics, vol.186, p.108460, 2022.

[131] N. Tanaka and M. Tanaka, "Mathematically trivial control of sound using a parametric beam focusing source," The Journal of the Acoustical Society of America, vol.129, no.1, pp.165–172, 2011.

[132] J. Yang, W.S. Gan, K.S. Tan, and M.H. Er, "Acoustic beamforming of a parametric speaker comprising ultrasonic transducers," Sensors and Actuators A: Physical, vol.125, no.1, pp.91–99, 2005.

[133] T. Kamakura, S. Sakai, H. Nomura, and M. Akiyama, "Parametric audible sounds by phase-cancellation excitation of primary waves," Acoustics' 08 in Paris, pp.4365–4370, 2008.

[134] A. Sugahara, H. Lee, S. Sakamoto, and S. Takeoka, "Measurements of acoustic impedance of porous materials using a parametric loudspeaker with phononic crystals and phase-cancellation method," Applied Acoustics, vol.152, pp.54–62, 2019.

[135] Y. Geng, H. Wang, M. Nakayama, and T. Nishiura, "Narrow-edged beamforming using masked parametric array loudspeakers," 2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA-ASC 2021), IEEE, 2021.