

円調和展開と多重極展開の関係と 音場再現への応用に関する研究

電気通信大学大学院 情報理工学研究科
博士（工学）の学位申請論文

堤 公孝

2020年3月

円調和展開と多重極展開の関係と 音場再現への応用に関する研究

博士論文審査委員会

主査 羽田 陽一 教授

委員 梶本 裕之 教授

委員 広田 光一 教授

委員 高橋 裕樹 准教授

委員 野村 英之 准教授

著作権保有者

2020年 堤 公孝

copyright 2020 Kimitaka Tsutsumi

Abstract

This research aims to reproduce arbitrary sound fields using loudspeaker arrays comprising multiple speakers. Both sound quality and directivity patterns (patterns of the powers radiated from a sound source in each direction) are important for a quality experience. If recorded sound fields could be reproduced in remote theaters, the quality of viewing of the audience members would be highly improved. This research focuses on sound field synthesis that can reproduce not only the sounds themselves but also the positions and directivity patterns of the sound sources.

Sound field synthesis technology, which uses a loudspeaker array comprising speakers arranged in a line or a circle, reproduces arbitrary sound fields by playing acoustic signals filtered by driving functions. One of the salient features of the proposed sound field synthesis technology is the in-depth control of sound position by moving the sources in horizontal directions, which cannot be achieved by existing sound reproduction systems. Human utterances and musical instruments have their own directivity patterns. Therefore, it is important for the improvement of the quality of experience to reproduce not only the positions but also the directivity patterns of sound sources. Directivity control technologies based on sound field synthesis have long been investigated; however, it is difficult for existing technologies to reduce algorithmic complexity.

Thus, this research focuses on sound field reproduction of a target directional sound source with low complexity. Focused sources are created using an actual linear loudspeaker array, which operates as virtual loudspeakers for the sound field synthesis of the directional source. The proposed method reflects the change of the sound field of a rotated or shifted directional sound source by moving the focused sources without changing the output from the virtual sound sources. This process corresponds to the use of focused sources as the virtual loudspeaker array. Thus, the sound field of moving directional sources can be accurately reproduced without increasing the

algorithmic complexity, which is not the case with existing methods.

In Chapter 1, existing spatial sound reproduction technologies, their relation to our study, and the aim of our research are introduced. In Chapter 2, technologies related to existing research on sound field synthesis and directivity control are described: the derivation of filters (driving functions) for each actual loudspeaker, circular harmonic expansion (spatial Fourier transformation in 2-dimensional space), and multipole expansion. In Chapter 3, a method is proposed to reproduce the sound field of a directional source using a linear loudspeaker array. The sound field modeled with circular harmonic expansion coefficients is reproduced through a virtual circular loudspeaker array created using a linear loudspeaker array. It is shown that this method, which updates the sound field by moving the virtual loudspeakers, requires less algorithmic complexity than existing methods. In Chapter 4, a method for reproducing the sound field of a directional source is described. This method also uses a virtual multipole loudspeaker array. The virtual array is created using a linear loudspeaker array and is driven with filters obtained based on the multipole expansion of the sound field. It is shown that the method operates at a lower algorithmic complexity than existing methods when the sound field is expressed by a multipole expansion of an order lower than three. In Chapter 5, the relation between two methods of modeling sound field synthesis, namely, circular harmonic expansion and multipole expansion, is discussed. This relation has long been investigated; however, it has not yet been defined analytically. An analytical conversion from circular harmonic expansion coefficients of a directional sound source to weighting coefficients of multipole sources is derived. It is shown that the sound field can be reproduced based on the superposition of multipole sources controlled by weighting coefficients converted from circular harmonic expansion coefficients. Finally, Chapter 6 concludes this dissertation.

論文要旨

本研究は、複数のスピーカからなるスピーカアレイを用いて所望の音場を再現することを目的とする。例えば、歌舞伎などの演舞場において発声や鈴の音、客席からの掛け声をマイクロフォンなどで収録し、遠く離れたサテライト会場において距離や方向などを忠実に再現しながら収録音響信号を再生できれば、サテライト会場にいる観客はまるで演舞場に潜り込んだような臨場感や一体感が得られると考える。本研究では、単に音響信号を再生するだけでなく前後左右に動きまわる音源の位置や音の伝搬方向といった空間情報をも再現するための手法として波面合成技術に着目する。波面合成技術では多数のスピーカを直線上に並べて構成した直線状スピーカアレイや、スピーカを円上に並べて構成した円状スピーカアレイに対し、スピーカ毎に求めたフィルタ（駆動関数）を入力音響信号に畳み込んで得られる出力信号を与えることで、所望の音場を再現する。この波面合成技術では、音源の左右の動きだけでなく物理的なスピーカ位置よりも前方の受聴者側に飛び出すような奥行き方向の動きも再現できることが知られており、従来の音響再生技術と比較した際の特徴のひとつとなっている。さらに、人の発話や楽器などには音の放射指向特性（音が特定の方向にのみ強く伝搬するといった、音の強弱に関する方向依存性）があるため、奥行き方向の制御に加えて音源の指向特性も再現することが音環境の再現性の向上には重要となる。波面合成技術を基盤として音源の指向特性を再現する研究についても従来から検討が進められてきているが、音源が回転したり平行移動したりする際の駆動信号の更新に多くの演算量を必要とし、また再現性も低下するといった課題があった。

そこで、本研究においては、演算量を抑えながら所望の指向特性を再現することに主眼を置き、スピーカアレイで生成した焦点音源を仮想的なスピーカ（仮想音源）とみなし、この焦点音源で生成した仮想音源を複数準備し、これに独立に駆動信号を与える手法を検討する。本手法では指向性音源の回転や平行移動による音場の変化を、仮想音源から出力される音響信号は変更せず、仮想音源である焦点音源の位置の更新により実現する。これらの処理は、仮想音源群を新たな仮想スピーカアレイとして捉えていることに相当し、実スピーカアレイの駆動信号の更新に伴う演算量の増加を抑

えつつ指向特性が作る音場の再現性を改善することが期待できる。

1章では、音源の位置を含めた空間的な情報をスピーカで再現する音響再生の研究を概説し、本研究の位置づけを述べる。2章では、波面合成技術と指向性制御技術について、所望の音場を再現するため音響信号に畳み込むフィルタ（駆動関数）の導出と従来研究について述べる。また、波面合成と指向制御のフィルタを導出する際に基礎となる平面上の空間フーリエ変換の一つである円調和級数展開と多重極展開について説明する。3章では、指向性音源が作る音場を直接的に再現するために、音場を円状に捉えることで円調和級数展開によりモデル化し、直線状スピーカアレイで生成する手法を提案する。円状アレイを用いて指向性音源が作る音場をモデル化することにより、実スピーカアレイの駆動信号を更新するのに必要な演算量を低減できることを示す。4章では、指向性音源が作る音場を、無指向性の音源を格子状に密な間隔で並べてできる多重極音源によりモデル化し、直線上スピーカアレイで生成する手法を提案する。多重極音源の重ね合わせにより音場をモデル化することにより、特に低次の多重極音源で表現可能な指向特性について、実スピーカアレイの駆動信号の更新に必要な演算量を低減できることを示す。5章では、従来研究において関係性が指摘されながらも陽に示されることがなかった円調和級数展開と多重極展開の間の解析的關係について考察し、所望の指向性音源が作る音場をモデル化した円調和展開係数を多重極展開係数に解析的に変換できることを示す。また、解析的変換で得られた多重極展開係数を用いて4章で示した方法で、指向特性を持つ音源を含んだ音場を再現できることを示す。6章で本論文の結論を述べる。

目次

第1章 序論	1
1.1 研究背景と目的	1
1.2 従来技術	2
1.3 研究の位置づけ	6
1.4 本論文の構成	8
第2章 波面合成と指向性制御	11
2.1 波面合成技術	11
2.1.1 キルヒホッフ-ヘルムホルツの積分方程式	12
2.1.2 2次音源と駆動関数	13
2.1.3 焦点音源法	13
2.1.4 焦点音源法の時間領域実装	16
2.1.5 時間領域実装の音場再現精度改善	17
2.2 指向性制御技術	22
2.2.1 空間フーリエ変換に基づく指向性制御	22
2.2.2 多重極音源に基づく指向性制御	30
2.2.3 多重極展開を用いた指向性制御	34
2.3 直線状アレイを用いた指向性音源の音場再現	36
2.4 音場再現と客観評価尺度	39
2.5 音源が持つ指向特性	39
2.5.1 人物発話や楽器における水平面上の指向特性	39
2.5.2 有限次元の円調和級数展開による打ち切り誤差	41
第3章 複数の焦点音源で生成した仮想円状スピーカアレイによる指向性再生	44
3.1 円状スピーカアレイを用いた指向特性の制御	45
3.2 直線状スピーカアレイを用いた仮想円状スピーカアレイの生成	47
3.3 時間領域での実装	47

3.4	時間領域での実装における音場再現の精度改善	48
3.4.1	Thiran フィルタを用いた時間領域実装による精度改善	48
3.5	性能評価実験	49
3.5.1	実験条件	49
3.5.2	客観評価尺度	50
3.5.3	円状アレイの半径が再現精度に与える影響	50
3.5.4	音場の再現精度	52
3.5.5	演算量の評価	58
3.6	本章のまとめ	60
3.6.1	本章の貢献と関連研究	60
第 4 章	複数の焦点音源から生成した仮想多重極音源による指向性再生	62
4.1	焦点音源からなる多重極音源の生成	62
4.2	多重極音源の効率的表現方法	64
4.3	時間領域での駆動関数	66
4.4	時間領域での実装における音場再現の精度改善	68
4.4.1	Thiran フィルタを用いた多重極音源の時間領域実装	68
4.5	性能評価	69
4.5.1	実験条件	70
4.5.2	客観評価尺度	71
4.5.3	モノポール音源同士の間隔が再現精度に与える影響	71
4.5.4	音場と再現精度	73
4.5.5	演算量の評価	78
4.6	本章のまとめ	79
4.6.1	本章の貢献と関連研究	79
第 5 章	音場再現における円調和級数展開と多重極展開	80
5.1	直線アレイによる音場再現手法の比較	80
5.2	単位円上での係数比較による円調和級数の解析的変換	81
5.2.1	関係式の導出	82
5.2.2	予備実験	83
5.3	多重極展開級数と円調和展開級数の係数比較による解析的変換	86
5.3.1	関係式の導出	86

5.3.2	単位円上での係数比較による解析的変換との関係性	87
5.3.3	予備実験	88
5.4	性能評価実験	92
5.4.1	実験条件	92
5.4.2	客観評価尺度	92
5.4.3	音場と指向特性の再現精度	92
5.5	本章のまとめ	97
5.5.1	本章の貢献と関連研究	97
第 6 章	結論	99
付 録 A	波面合成技術を活用したパブリックビューイング	103
A.1	Kirari! と波面合成技術	103
A.2	システム構成	104
	参考文献	108
	謝辞	119
	学位論文を構成する論文	122
	研究業績リスト	123
	著者略歴	126

第1章 序論

1.1 研究背景と目的

本研究は、歌舞伎などの演舞場においてマイクロフォンなどで収録した音響信号を、音源の位置や音源が持つ方向別の強弱のパターン（以降、指向特性と呼ぶ）を含めて遠隔地で再現する空間音響技術に関する研究である。空間音響技術は、1) 音源の位置を受聴者に知覚させる音像定位と、2) 残響のように空間的な広がりを知覚させる環境音再現、の2要素からなる [1]。本研究は、これらのうち音像定位技術に焦点を当てる。音像定位技術には、少数のスピーカに与える音圧差や再生時間差を制御することで仮想的な音源の位置を受聴者に知覚させるステレオ再生やサラウンド再生技術 [2] [3]、人間の頭部や耳における音の伝搬特性を考慮した信号処理によりヘッドフォン等を装着した受聴者に音響空間を知覚させるバイノーラル再生技術 [4]、波動方程式などでモデル化された収録場所の音響空間を多数のスピーカを制御して再現する波面合成技術 [5] [6] などが検討されてきた。本研究は、複数人で視聴するスタイルを前提とした音響再生技術に焦点を当てる。そのような音響再生技術の適用先として、特にライブイベントなど比較的大人数で同じコンテンツを楽しむ際に個別の音源の空間的な位置関係や方向別の音の強弱のパターンを再現する音像定位技術に着目する。個人で視聴する場合と大勢で視聴する場合の大きな違いは、別の観客との会話や別会場で視聴する観客との間の歓声等の共有であるとの視点に立つ。この場合、ヘッドフォンなどを用いた受聴スタイルは観客同士の会話の妨げとなる。また、別会場の観客の歓声を再現するためには、受聴者の近くに音像を生成しなければならず、受聴者の近くまで飛び出る音像（以降、焦点音源と呼ぶ）の生成が必要となる。このような観点から、本研究ではキルヒホッフ-ヘルムホルツの積分方程式に基づきモデル化された音場を多数のスピーカを制御して再現する波面合成技術に着目する。波面合成技術を用いれば、スピーカより前面の客席近くまで音像を飛び出させることができるため、他の会場の歓声などを受聴者の隣の席から発せられたかのように再生したり、解説音声を対象となる受聴者のみに呈示するといったことが可能となる。

一方、楽器や人間の発声器官は、個々に異なる形状と発音メカニズムを有するため様々な指向特性を持つ。音源が持つ指向特性の再現と受聴者が感じる原音らしさの間

には関連があることも指摘されている [7] [8]. そこで, 本研究では, 収録地点の音環境の再構築を目的とし, 前後左右といった音源の位置情報のみならず音源が持つ指向特性を再現することを目指す. 従来, 複数のスピーカからなるスピーカアレイを用いた指向特性の再現には, 円調和級数展開などの空間フーリエ変換に基づく手法や, 直交格子状に密に並んだ音源が生成する多重極音源の重み付き和である多重極展開に基づく手法などが検討されてきた. 本研究では, 実スピーカアレイで生成した複数の焦点音源を新たな仮想スピーカとみなし, この仮想スピーカ群に対して, 円調和展開や多重極展開に基づく指向性制御技術を適用することで, スピーカよりも前面に指向特性を有する音源が作る音場を生成する研究を行う. さらに, 従来, 関連が指摘されつつも陽に定式化されることがなかった円調和展開級数と多重極展開の展開係数の間の解析的な関係について定式化を行う.

1.2 従来技術

マイクロフォンで収録した音響信号を, 音源の空間的な位置関係も含めて再現性高く再生する空間音響技術については, 1881 年に電話を 2 回線用いてコンサートを中継したもの [9] [10] を始めとして様々なものが検討されてきた [11] [12] [13] [14]. 空間音響技術は, 聴覚心理の成果を活用して所望の音を知覚させる心理音響と, 収録場所における音の物理的な現象を可能な限り正確に再現する物理音響に分類できる [14].

心理音響による音像定位技術には, 家庭や映画館などでも広く用いられているステレオ再生やサラウンド再生がある. これらの再生手法は, 2 つのスピーカの音圧差や時間差によって, 実際に音源がない場所に仮想的な音源を感じる phantom source (仮想音源) と呼ばれる聴覚心理的現象に基づき実現される [1] [15]. 聴覚心理の研究成果として, 2 つのスピーカの音圧差や時間差と, 受聴者が知覚する仮想音像の位置の関係が体系化され, サイン則やタンゼント則と呼ばれる音圧制御を用いた仮想音源の位置制御手法が確立された [16]. このような聴覚心理の成果を応用し, 音響信号に空間的な位置情報を与えて音響の空間的な広がりを表現するステレオ再生 (stereophony) の技術が広く用いられるようになった [17] [3] [18]. 適切なスピーカの組を選んでサイン則やタンゼント則を適用することで, 5.1 チャンネル [19] や, 22.2 チャンネル [20] といった多チャンネルのスピーカ配置に対しても音像位置の制御が可能となる. このように知覚する仮想音源の位置と, それを実現するためのスピーカの選択方法, 各スピーカへの音圧や遅延の付与方法を数学的に体系化した Vector based amplitude panning (以降, VBAP と呼ぶ) が提案され, 左右方向のみならず上下方向にも仮想的な音源を

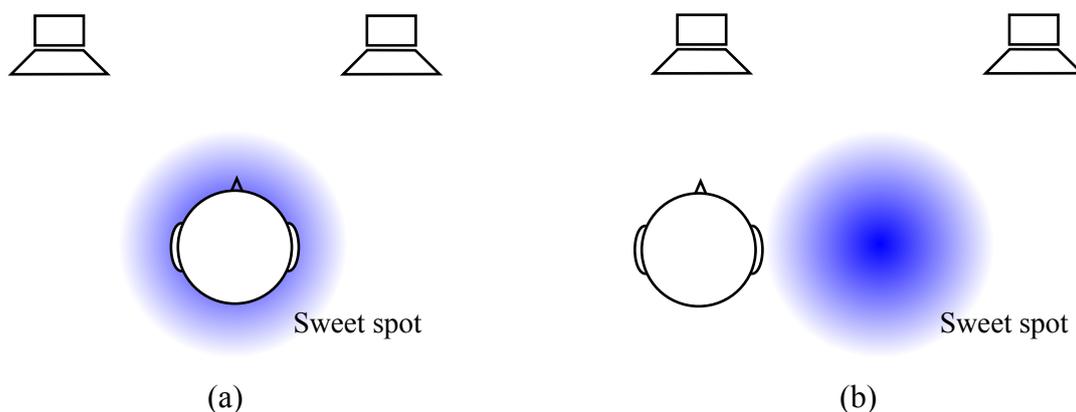


図 1.1: ステレオ再生方式の模式図. (a) ステレオ再生における受聴者とスイートスポットの位置関係. (b) 受聴者がスイートスポットの外にいる場合 (ステレオ感が損なわれる).

生成する手法が確立された [21].

一方, これらの心理音響に基づく再生方式では, 複数の受聴者が同じ位置に仮想音像を知覚できない課題がある. これは, 各スピーカから再生するコンテンツを制作する際に図 1.1 (a) に示すように参照用の観測地点を設定し, そこで受聴することを前提としてスピーカ間の音圧差や時間差を設定するためである (図では例としてステレオ再生方式を示す). 限られた空間は一般にスイートスポットと呼ばれている. 図 1.1 (b) に示すように参照用の観測地点以外でコンテンツを受聴すると, 本来意図した音圧差や時間差を再現できず, 仮想音像の位置が変化する. また, 心理音響に基づく再生方式では, 仮想音像の位置がスピーカとスピーカの間に制限され, スピーカよりも客席近くまで迫るような音像を生成できない課題もある.

さて, 受聴者にスピーカを用いて音響信号を呈示する方式のほかに, ヘッドフォンを用いて音響信号を呈示するバイノーラル方式も知られている [4]. 3次元空間で起こるあらゆる音響信号は, 頭部や耳介での反射や拡散を経て2つの耳で観測される. そのため, 両耳の入り口での音圧が正確に再現されれば, 理論上あらゆる空間音響を受聴者に呈示可能となる. このような考えに基づき, 図 1.2 (a) に示すように, 頭部や耳介を経て耳の入り口で観測されるインパルス応答 (頭部伝達関数) を事前に測定しておき, これを入力音響信号に畳み込んで得られた音響信号を図 1.2 (b) に示すように, ヘッドフォンで呈示する音響再生方式がバイノーラル方式である. 頭部伝達関数は, マネキンの耳にマイクを設置したダミーヘッドなどで収録されることが多く, 様々な機関が頭部伝達関数のライブラリを公開している [22] [23] [24] [25] [26]. しかし, このように公開されている頭部伝達関数を用いてバイノーラル再生を行っても, 期待した

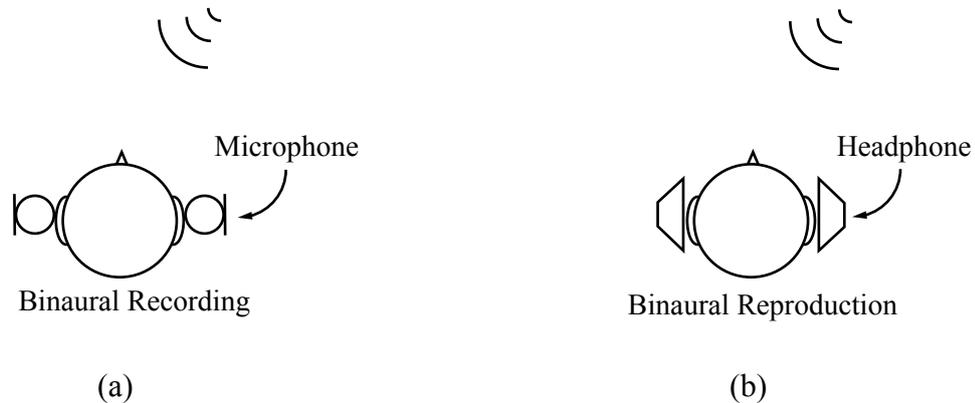


図 1.2: バイノーラル再生方式の模式図. (a) ダミーヘッドなどを用いたバイノーラル録音. (b) ヘッドフォンを通じたバイノーラル再生 (受聴者は音源の空間的な位置を知覚可能).

ような空間音響を受聴者に呈示できない課題がある. 例えば, 前方への音像定位を意図して呈示した仮想音像が頭の中や, 後方から知覚されることが起こる. このような現象は, 1. 受聴者本人の頭部伝達関数とバイノーラル再生に用いた頭部伝達関数が一致しない [27], 2. 頭部伝達関数の収録・再生を行う音響系に対する適切な補正がなされていない [28] [29], 3. 頭部の運動に伴う動的な頭部伝達関数の適用がなされていない [30] [31] [32], といった要因により生じると考えられている [4]. 境界要素法などを用いて頭部の形状データから個人に適応した頭部伝達関数を生成する, あるいは受聴者の頭部の運動に追従して頭部伝達関数を更新する, といった取り組みも盛んに行われている [33] [34] [35] [36]. しかし, 受聴者の頭部の動きを追跡しながらリアルタイムに頭部伝達関数を生成するのは難しく, また, 複数人で同じコンテンツを楽しむ場合に受聴者同士の会話がヘッドフォンに遮られる課題もある.

一方, 音響場における物理的な現象を正確に再現する物理音響に基づく技術の一つとして, アンビソニックス (Ambisonics) [37] や, その拡張方式である高次アンビソニックス [38] [39] [40] [41] [42] が知られている. アンビソニックスは, モノラル, ステレオ, 3次元音響を階層的に表現する方式であり, 受音点における音圧を球面調和展開を用いて表現する [37]. アンビソニックスは, 受音点にあらゆる方向から到来する平面波を近似的に表現する手法であるため, 受音点から離れた位置で受聴した場合に, 意図した音場が再現できない課題がある. 通常アンビソニックスと高次アンビソニックスの違いは, アンビソニックスが球面調和展開の係数を1次までしか用いないのに対し, 高次アンビソニックスは2次以上の展開係数も利用することである. これは, 受音点に入射する平面波の重ね合わせで音場を表現する際の再現精度に関連し

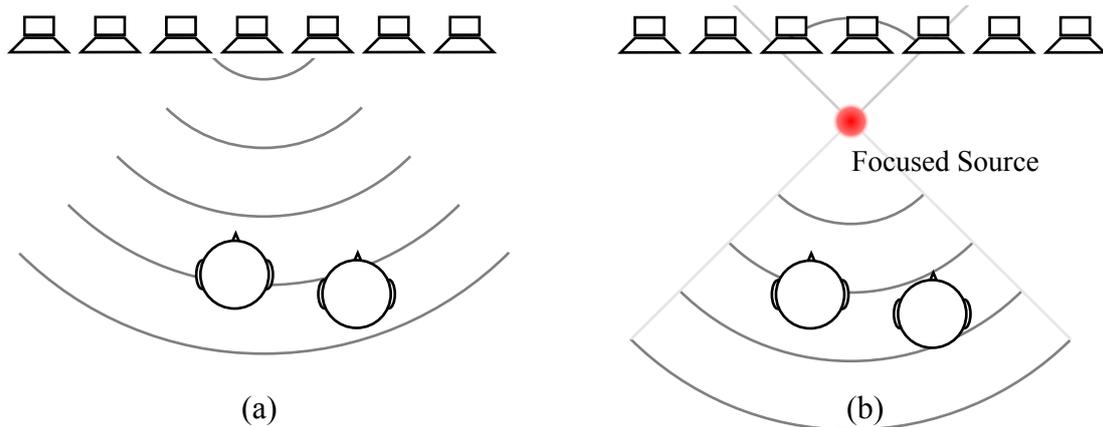


図 1.3: 波面合成技術の模式図. (a) 直線状スピーカアレイと受聴者の位置関係. (b) 焦点音源法における焦点音源と受聴者の位置関係.

ており, 高次アンビソニックスでは, より複雑な音場が表現できるだけでなく, 意図した音場を再現できる受聴エリアが拡大することも期待できる [43] [44]. しかし, 高次の球面調和展開級数を再現するためには多数のスピーカを用いる必要があるため受聴エリアを拡大するのは実装上難しく, ステレオ再生と同様に受聴エリアや仮想音像位置の自由度と言った課題を解決することは難しい.

物理音響に基づく別の空間音響技術として, 音の物理的な波面を再現する波面合成技術がある [5] [6] [45]. 波面合成技術は音の波面を物理的に厳密に再生するため, 他の再生技術よりも, 広い受聴範囲 (スイートスポット) が得られることが特徴である. 波面合成技術を用いれば, スイートスポット内の受聴者が複数人であったり移動していても良好な音響効果を与えることができる. また, ステレオ再生方式とは異なり, 音源の方向はもちろん, 音源の奥行きを含めた立体的な空間音響の再現が可能である [46]. しかし, スピーカ間隔を再現対象の音波の $1/2$ 以下としなければ物理的に正確な波面を形成できないため, システムの規模が膨大になる課題がある [44]. また, 理論上は無限の長さを持つ連続な音源で生成する波面を, 図 1.3 (a) に示すように, 有限長かつ離散的に配置されたスピーカを用いて再現するため, 特に図 1.3 (b) に示すように焦点音源法を用いて仮想音源の位置をスピーカよりも前面に設定した際に, 受聴エリアが制約を受けるといった課題がある [47] [48] [49].

さて, 実スピーカアレイを用いた指向性制御には, 図 1.4 (a) のように円状アレイを用いるもの [50] [51] [52] [53], 図 1.4 (b) のように多重極展開を用いるもの [54] [55] [56] [57] [58] [59] [60], 等がある. 円状アレイは円調和級数展開に基づき駆動関数を導出するため, 指向特性の再現性を保ったままアレイのサイズを小さくすることが難し

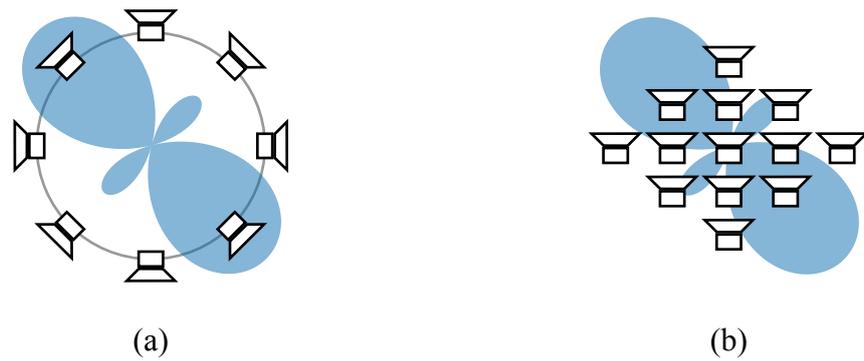


図 1.4: 指向性制御技術の模式図. (a) 円調和級数展開に基づく指向性再生. (b) 多重極展開に基づく指向性再生.

い [52] [53]. 一方, 多重極展開に基づくものはアレイのサイズを小さくすることができる [60] が, 多重極音源の構成に必要な無指向性音源の数が次数の増大に伴って膨大になる課題がある [59].

1.3 研究の位置づけ

本研究では, 同一空間内において多人数が同じ音体験を同時に受聴して感動を共有できる空間音響技術の実現を目指す. 1.2 節でも述べた通り, ヘッドフォンを通じた受聴では受聴者同士の会話などが妨げられるため, スピーカを用いた音像定位技術の確立を目指す. 別会場の観客と同一音響空間を共有するためには, その会場における歓声などを客席近くに音像定位させる必要がある. これを実現するため, 波面合成技術を用いて, スピーカよりも前面に飛び出し客席側に迫る仮想音源の生成を行う. また, 客席側に迫る仮想音像の再現性を高めて, あたかも元の音源が受聴者の近くに現れたかのような音場を作り出すため, 図 1.5 に示すように客席側に迫る仮想音源に指向特性を与える研究を行う.

波面合成技術には, 複数のマイクを密に並べたマイクアレイで収録した音響信号に処理を行い, 複数のスピーカを密に並べて構成したスピーカアレイから収録地点の音場を再現する波面再構成フィルタに基づく波面合成技術もある [61] [62] [63] [64]. 波面再構成フィルタは収録地点の音場をあらかじめ決めた境界上で記録して遠隔地で再現できるため, 収録対象の音源が作る音場を指向特性まで含めて再現できる. しかし, 歌舞伎やサッカーなどの競技場では音響上最適な場所にマイクアレイを設置することは, 演技や競技の妨げとなるため, 一般的に難しい. そこで本研究では, 入力としてモノラルの音声信号と位置情報が与えられた場合に生成される目標音場を再現するタイプのオブジェクトベースと呼ばれる波面合成技術 [6] に焦点を当てて検討を行う.

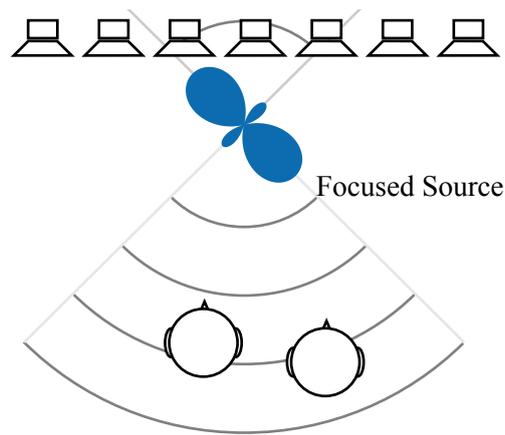


図 1.5: 本研究で目指す仮想音源の指向性制御.

さて、これまでも波面合成技術を用いて、指向性音源が作る音場を再現する技術については検討が行われてきた [65] [66] [67] [68] [69]. これらの技術では、指向性音源が生成する音場の空間勾配から導出される駆動関数を用いることで目標となる音場を再現する. しかし、指向性音源が平行移動したり回転したりした場合、方向別の音の強弱のパターンそのものは大きくは変わっていないにも関わらず、従来の技術では、指向性音源の平行移動や回転に対しても駆動関数の再計算が必要となるため、演算量の削減が困難となる課題があった. そこで本研究では、図 1.6 に示すように、波面合成で生成した複数の焦点音源で仮想のスピーカアレイを生成するための駆動関数と、仮想スピーカアレイの制御により指向特性の形状を生成するための駆動関数を独立に算出し、これらを結合して得られる駆動関数により指向性音源が作る音場を再現する. このような駆動関数を用いた場合、指向性音源の平行移動や回転を仮想スピーカアレイの生成に関わる駆動関数の更新のみで実現できるため、指向性パターンの再計算に必要な演算量を低減できると考える.

仮想スピーカアレイの指向性制御には、実スピーカからなるスピーカアレイで用いられる指向性制御技術を利用することができる. 本研究では、図 1.7 に示すように、円状アレイに基づく方法と多重極展開に基づく方法を仮想スピーカアレイの指向性制御に適用して比較を行うことで、各手法の優位性を議論する. 円調和級数展開と多重極展開の間には密接な関係があることが指摘されており [54], 多重極展開により生成される空間は、円調和級数の線形和により表現できることも数学的に証明されている [70]. しかし、指向性音源が作る音場をモデル化した円調和展開級数と、同じ音場をモデル化した多重極音源の重み係数の間の解析的な関係については陽に定式化されていない. 本研究では、これらの展開級数の間の解析的な変換を導出する検討を行い、指向

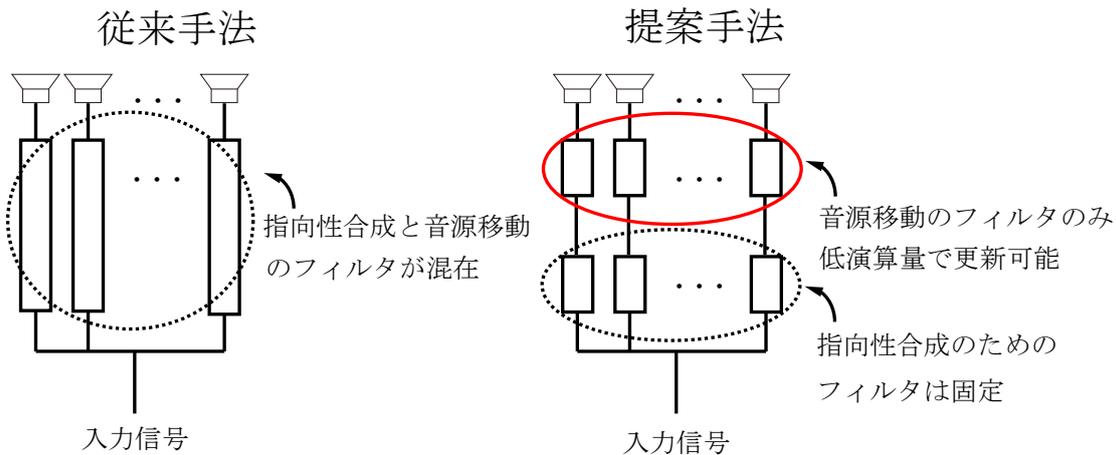


図 1.6: 本研究で目指す指向性音源の平行移動や回転における演算量削減. 従来手法では, 指向特性の変化だけでなく指向性音源の移動に対してもスピーカごとにフィルタの更新が必要. 一方, 提案手法は指向特性の変化と音源の移動を別々のフィルタで表現し, 音源の移動については片方のフィルタのみ低演算量で更新.

性音源が作る音場をモデル化した円調和展開級数で求めた駆動関数で多重極音源を駆動して元の音場を再現できることを示す.

1.4 本論文の構成

図1.8に, 本論文の構成を示す. 2章では, 波面合成技術と指向性制御技術について, 所望の音場を再現するため音響信号に畳み込むフィルタ (駆動関数) の導出と従来研究について述べる. また, 波面合成と指向制御のフィルタを導出する際に基礎となる平面上の空間フーリエ変換の一つである円調和級数展開と多重極展開について説明する. 3章では, 指向性音源が作る音場を直接的に再現するために, 音場を円状に捉えることで円調和級数展開によりモデル化し, 直線状スピーカアレイで生成する手法を提案する. 円状アレイを用いて指向性音源が作る音場をモデル化することにより, 必要な焦点音源の数を低減できることを示す. 4章では, 3章と同様に指向性音源が作る音場を直接的に再現する手法を検討する. 4章では無指向性のモノポール音源を直交格子に密に並べてできる多重極音源の重ね合わせで音場をモデル化し, 指向性音源が生成する音場を直線状スピーカアレイで生成する手法を提案する. 多重極音源を用いて指向性音源が作る音場をモデル化することにより, 従来円状アレイでモデル化されてきた指向性音源の音場を多重極音源の重ね合わせでも高精度に再現できることを示す. 5章では, 従来研究において関係性が指摘されながらも陽に示されることがなかった円調和展開と多重極展開の間の解析的關係について考察し, 所望の指向性音源

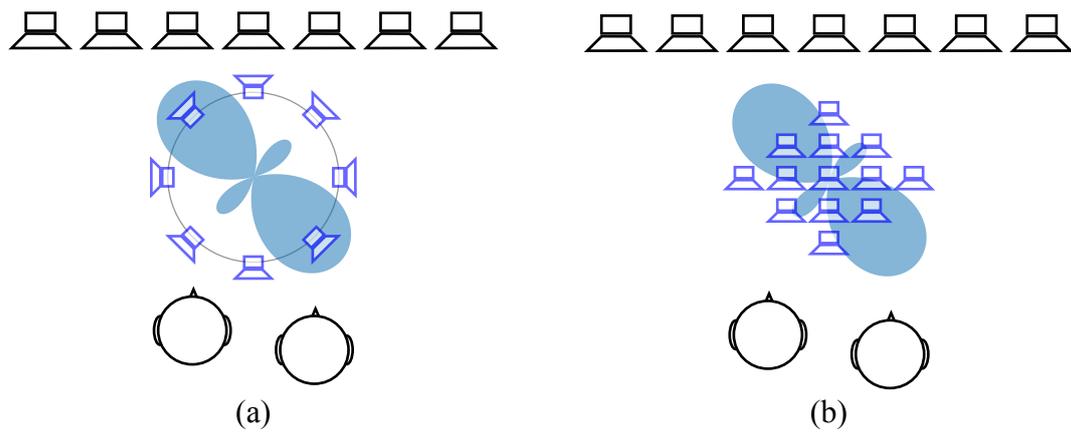


図 1.7: 本研究で目指す仮想音源の指向性制御. (a) 仮想の円状アレイに基づく仮想音源の指向性制御 (3 章で議論). (b) 仮想の多重極音源に基づく仮想音源の指向性制御 (4 章で議論).

が作る音場をモデル化した円調和展開係数を多重極展開係数に解析的に変換できることを示す. また, 解析的変換で得られた多重極展開係数を用いて 4 章で示した方法で, 指向特性を持つ音源を含んだ音場を再現できることを示す. 最後に, 6 章で本論文の結論を述べる.

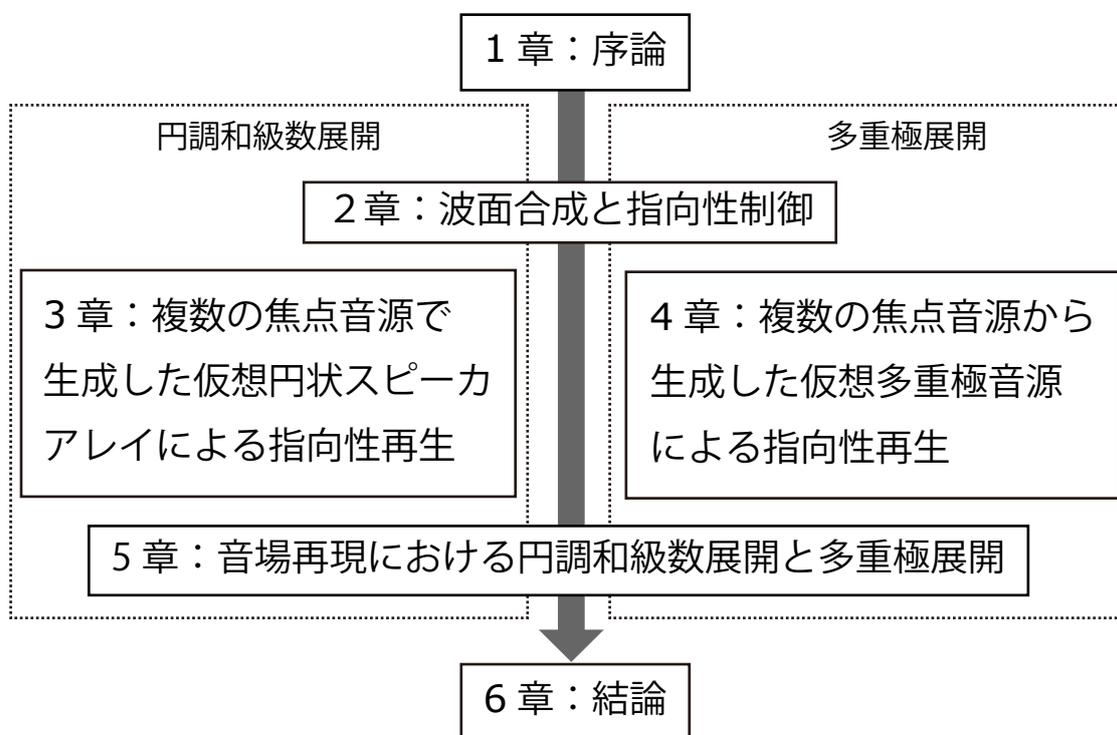


図 1.8: 本論文の構成

第2章 波面合成と指向性制御

本章では、複数のスピーカからなるスピーカアレイで所望の音場を再現する波面合成技術と、スピーカアレイで指向性を制御する手法について述べる。いずれの技術も音場をキルヒホッフ-ヘルムホルツの積分方程式でモデル化し、そこから駆動関数と呼ばれる各音源の重み係数を導出する。この重み係数をフィルタとして各スピーカに適用することにより、所望の波面を再現したり、スピーカの指向特性を制御することができる。本章では、これら波面合成技術と指向性制御技術について、概要を述べる。

2.1 波面合成技術

波面合成技術は、多数のスピーカを用いて高精度に聴覚情景を再現する空間音響再生技術である。波面合成技術は、60年以上前に Snow により初めて提案され [71]、その後、Berkout をはじめとするデルフト大学の研究チームにより波面合成理論の検討と、実機を用いた検証が進められた [5] [45]。当初の研究では、水平面上にスピーカを直線上に並べた直線状スピーカアレイを用いて水平面の再生エリアを再現するに止まっていたが、曲線、平面など任意の形状のスピーカアレイを用いて3次元空間の再生エリアの音場を再現するよう拡張が進められ、同時に直線アレイを用いた平面の音場再現手法との理論統一が行われた [6]。近年、直線アレイを用いた波面合成技術をライブイベントなどに応用する取り組みが行われている [72] [73]。ステレオ再生やサラウンド再生でも仮想的な音源の移動は表現できるが、上下左右といった方向しか制御することができない。一方、波面合成技術を用いると、音源の方向に加えて奥行き方向の制御も可能となり、客席まで仮想的な音源が飛び出して聞こえるといった効果を実現することができる。このような技術には様々な実装形態が考えられるが、一例として、筆者らが開催したライブイベントにおける波面合成技術の適用事例について、付録 A にて紹介する。

波面合成技術は、以下で示すキルヒホッフ-ヘルムホルツの積分方程式を基礎としてまとめられる。

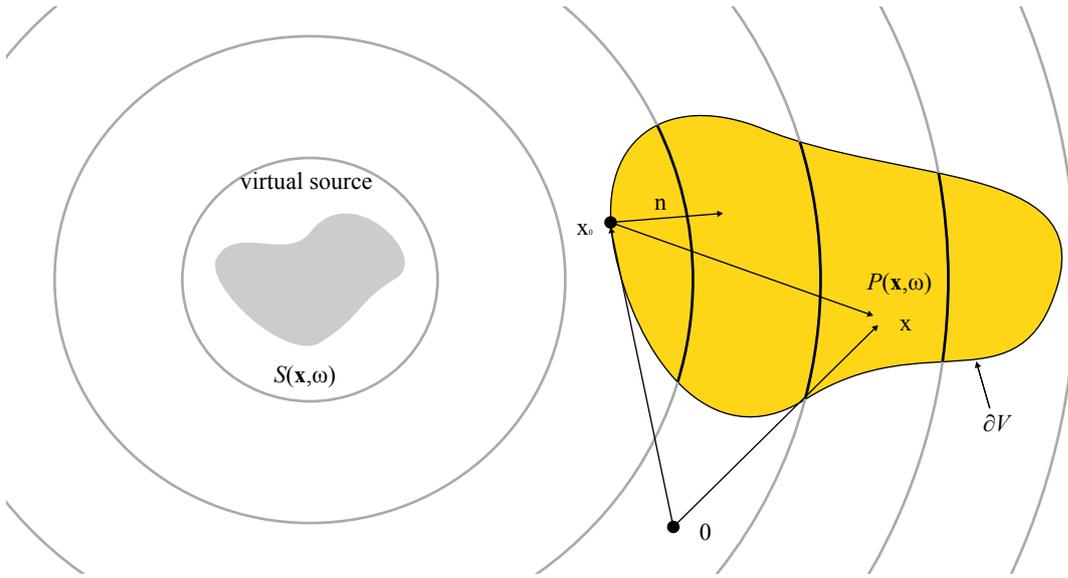


図 2.1: キルヒホッフ-ヘルムホルツ積分方程式における閉領域と音源の位置関係.

2.1.1 キルヒホッフ-ヘルムホルツの積分方程式

今, 音場を 2 次元平面に限定して話を進める. 受聴者を囲むスピーカアレイは, 波動方程式の非斉次境界条件とみなすことができる [54]. 閉領域と音源の位置関係を図 2.1 に示す. 非斉次境界条件を満たす曲面で囲まれた閉領域 V に対する斉次波動方程式の解は, 次に示すキルヒホッフ-ヘルムホルツの積分方程式で表される.

$$P(\mathbf{x}, \omega) = - \oint_{\partial V} \left(\mathbf{G}(\mathbf{x}|\mathbf{x}_0, \omega) \frac{\partial}{\partial \mathbf{n}} P(\mathbf{x}_0, \omega) - P(\mathbf{x}_0, \omega) \frac{\partial}{\partial \mathbf{n}} \mathbf{G}(\mathbf{x}|\mathbf{x}_0, \omega) \right) d\mathbf{S}_0 \quad (2.1)$$

ここで, ω は角周波数であり, $P(\mathbf{x}, \omega)$ は閉曲面 ∂V で囲まれた領域 V 内部の点 \mathbf{x} の音圧を表し, $G(\mathbf{x}|\mathbf{x}_0, \omega)$ は位置 \mathbf{x}_0 におけるディラックパルス駆動源とする非斉次波動方程式のグリーン関数を表し, $P(\mathbf{x}_0, \omega)$ は, 閉曲面 ∂V 上の音圧分布を表す. また, \mathbf{n} は閉曲面 ∂V と直交し領域 V の内部を向いた単位ベクトルを表し, $\frac{\partial}{\partial \mathbf{n}}$ は \mathbf{n} 方向の方向微分を表す. すなわち,

$$\frac{\partial}{\partial \mathbf{n}} P(\mathbf{x}_0, \omega) = \langle \nabla P(\mathbf{x}, \omega), \mathbf{n}(\mathbf{x}_0) \rangle \quad (2.2)$$

ここで, $\langle \cdot, \cdot \rangle$ は二つの関数の内積を表す. なお, 閉領域 V の内部には音源を含まないものとし, 閉領域 V の外部の音圧分布 $P(\mathbf{x}_0, \omega)$ は考慮しない. なお, 2 次元平面における非斉次波動方程式のグリーン関数 $G(\mathbf{x}|\mathbf{x}_0, \omega)$ は, 次式で与えられる [54].

$$G(\mathbf{x}|\mathbf{x}_0, \omega) = \frac{j}{4} H_0^{(2)}(k|\mathbf{x} - \mathbf{x}_0|) \quad (2.3)$$

なお, $j = \sqrt{-1}$ は虚数単位, $k = \omega/c$ は波数, c は音速を表す. $H_0^{(2)}$ は 0 次の第 2 種ハンケル関数である.

2.1.2 2次音源と駆動関数

目標の音場が与えられたとき、直線上あるいは円上に密に並べた各スピーカに適切なフィルタを設定して目標の音場を再現するのが波面合成技術である [6]. 目標音場を再現するために再生場で用いるスピーカのことを2次音源、これらに与えるフィルタを駆動関数と呼ぶ. ここでは、直線状に並べたスピーカを2次音源として用いた場合の波面合成について説明する. 今、スピーカを $\mathbf{x}_0 = (x_0, 0)$ と x 軸上に連続に配置したとすると、直線状スピーカアレイが $\mathbf{x} = (x, y)$ に作る音圧は次式で与えられる. なお、 $\mathbf{x} = (x, y)$ は $y > 0$ における観測位置を表し、 $\mathbf{x}_0 = (x_0, 0)$ は2次音源の位置を表す.

$$P(\mathbf{x}, \omega) = - \int_{-\infty}^{\infty} D_{2D}(\mathbf{x}_0, \omega) G_{2D}(\mathbf{x} - \mathbf{x}_0, \omega) d\mathbf{x}_0 \quad (2.4)$$

D_{2D} は \mathbf{x}_0 における駆動関数を表し、 $G_{2D}(\mathbf{x}, \omega)$ は式 (2.3) に示す2次元グリーン関数を表す. また、 ω は角周波数を表す. 所望の音場が $S(\mathbf{x}, \omega)$ で表されるとき、第一種レイリー積分を用いると駆動関数は以下の式で表せることが知られている.

$$D_{2D}(\mathbf{x}_0, \omega) = 2 \frac{\partial}{\partial y} S(\mathbf{x}, \omega) \Big|_{\mathbf{x}=\mathbf{x}_0} \quad (2.5)$$

2.1.3 焦点音源法

波面合成技術を用いて、スピーカアレイと客席との間に飛び出す仮想的な音源を作り出すことができる. このような仮想音源は焦点音源として知られている. 波面合成技術を応用して焦点音源を作り出す焦点音源法では、仮想音源の位置に吸込み型音源を仮定して駆動関数を導出する [74]. 音響信号は仮想音源の位置に向かって2次音源から放射され、仮想音源に向かって収束する音場と、仮想音源から発散される音場が生成される. その結果、仮想音源の位置に無指向性音源を配置した際にできる音場が生成される (図 2.2). 焦点音源法では、所望の音場を以下の式で定義される吸い込み型の音源でモデル化する.

$$S_{2D}(\mathbf{x}, \omega) = \frac{j}{4} \sqrt{jk} H_0^{(1)}(k|\mathbf{x} - \mathbf{x}_s|) \quad (2.6)$$

ここで、 $H_0^{(1)}$ は、0次の第1種ハンケル関数を表し、 $\mathbf{x}_s = (x_s, y_s)$ は、 $y_s > 0$ における焦点の位置を表す. 式 (2.6) を式 (2.5) に代入すると、駆動関数は次式で定義できる.

$$D_{2D}(\mathbf{x}_0, \mathbf{x}_s, \omega) = \frac{(jk)^{3/2}}{2} \frac{y_0 - y_s}{|\mathbf{x}_0 - \mathbf{x}_s|} H_1^{(1)}(k|\mathbf{x}_0 - \mathbf{x}_s|) \quad (2.7)$$

ここで、 $H_1^{(1)}$ は1次の第1種ハンケル関数を表す.

理論上、波面合成には連続な2次音源を用いる必要があるが、連続な音源を実現することは難しいためスピーカを密に並べることで連続な音源を近似的に実装する. 連

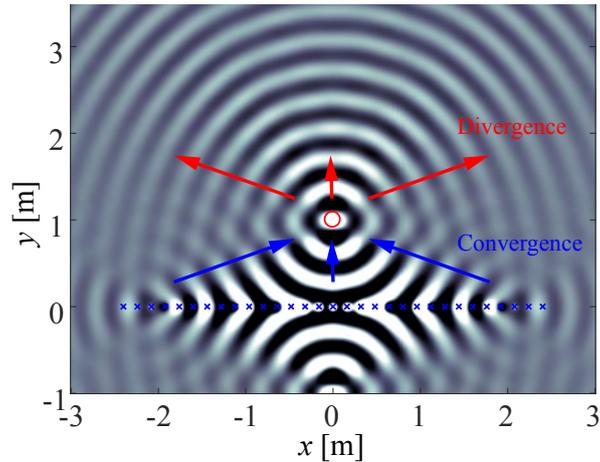


図 2.2: 直線状スピーカアレイを用いて生成した焦点音源が作る音場 (1 kHz の単一周波数). 色の濃淡は音圧の大小を表す. 'o' は焦点音源, 'x' は 2 次音源を表す.

連続な音源の代わりに等間隔にスピーカを配置して駆動することは、駆動信号の離散サンプリング、および各スピーカから発生した音場の重ね合わせによる補間プロセスに相当する。一般的に、デジタル信号処理ではサンプリング対象となる信号が含む周波数の 2 倍以上の周期でサンプリングしなければ元の連続信号を正確に復元できずに歪みが生じるエイリアシングが発生するが、本来連続な音源を前提としている駆動信号を、離散的に配置されたスピーカアレイで駆動する波面合成でもスピーカの間隔と再現対象の信号周波数によって空間的なエイリアシングが発生する。デジタル信号処理におけるエイリアシング周波数と同様に、駆動関数の空間的なサンプリング間隔も以下の空間エイリアシング周波数にしたがうことが示されている [47].

$$f_{al} \leq \frac{c}{2\Delta x} \quad (2.8)$$

ここで、 c は音速、 Δx はスピーカ間隔を表す。例えば音速 c が 343.46 m/s、スピーカ間隔が 0.05 m のとき、空間エイリアシング周波数はおよそ 3400 Hz と計算できる。

駆動関数をスピーカアレイを用いて実装すると、空間エイリアシングとは別の信号歪みも発生する。理論上、2 次音源は無限長である必要があるが、実際に無限に長いスピーカアレイを用いることはできないため有限長で打ち切ってスピーカアレイを実装する必要がある。そのため、打ち切り誤差と呼ばれる信号歪みが生じる。2 次音源を有限長で打ち切ることによる再現音場への影響については、様々な検討が行われてきた [47] [75] [76] [77]。数学的に、2 次音源の打ち切りは空間的に並んだ駆動関数列に適切な窓関数をかけることに相当し [75]、受聴者に対して意図した音場を呈示可能な受聴エリア (以下、スイートスポット) の限定をもたらす [76]。有限長のスピーカアレイで

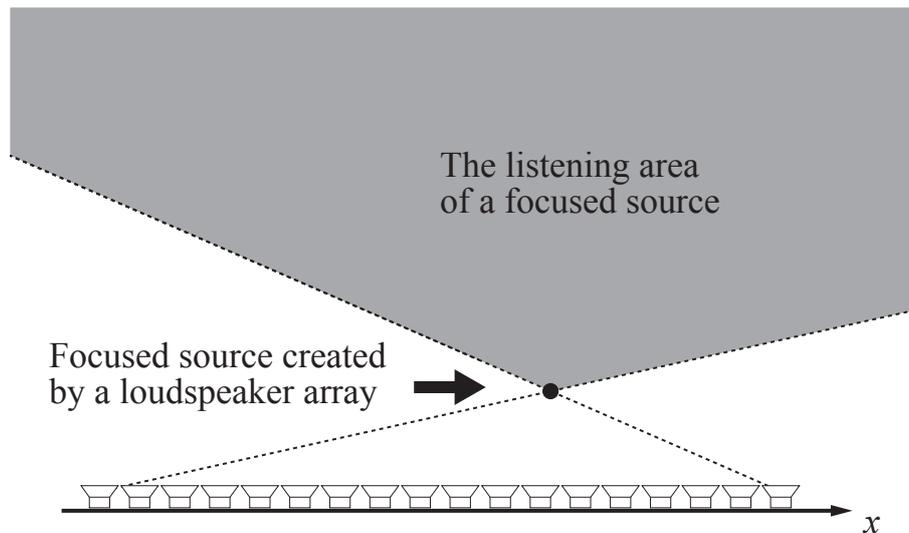


図 2.3: 有限な長さの 2 次音源を用いた場合にできる焦点音源のスイートスポット.

生成した焦点音源のスイートスポットは、図 2.3 に示す範囲で近似的に表すことができる。図からもわかるとおり、焦点音源が 2 次音源から遠ざかるほどスイートスポットも狭くなることがわかる。また、2 次音源を有限長で打ち切ることで、2 次音源の端点からの散乱による音場の歪みが生じることが確認されている。このような歪みは、端点に追加の音源を配置することでモデル化できることも明らかにされている [77]。このような端点からの散乱効果は、駆動関数に平滑化の窓関数 (tapering window) をかけることで抑圧できることも確認されている。さらに、有限長の 2 次音源を用いて焦点音源を生成する際の音場の歪みとして焦点音源のサイズの増加が確認されている [78]。物理学あるいは光学の分野では、このような焦点の増大は散乱 (diffraction) として知られており、例えば光学システムの解像度の低下などをもたらす。同様に、焦点音源のサイズが増大することで音像がぼやけるといった効果が現れる。

2.1.4 焦点音源法の時間領域実装

周波数領域で定義した式(2.7)で示した駆動関数は、2次音源毎に異なる周波数領域の関数で定義されている。そのため、実際にある音響信号を焦点音源として所望の位置に定位させるためには、2次音源と同数の異なるフィルタを音響信号に適用する必要がある。焦点音源の位置が動くような場合には、音響信号へのフィルタリングに加えて、フィルタ（駆動関数）の再計算も必要となるため、再生信号の計算に必要な演算量の増加が深刻となる。本項では、2.1.3項で求めた駆動関数に近似を加えた上で逆フーリエ変換を適用することにより、駆動関数を時間領域のフィルタに変換する。時間領域のフィルタに変換することにより、周波数領域の駆動関数において必要であった2次音源毎のフィルタリングを2次音源毎のゲイン乗算と遅延処理に変換して低演算量化することが可能となることを示す。

時間領域の駆動関数は式(2.7)で示した周波数領域の駆動関数から導出を行うことができる。 ν 次の第一種ハンケル関数は、漸近展開により次式のように複素指数関数で近似できる [79].

$$H_\nu^{(1)}(x) \approx \sqrt{\frac{2}{\pi x}} e^{j(x - \frac{\pi}{2}\nu - \frac{\pi}{4})} \quad (2.9)$$

なお、 ν は次数を表す整数のインデクスである。この関係式を式(2.7)に代入すると周波数領域の駆動関数を次式のように変形できる。

$$\begin{aligned} D_{2D}(\mathbf{x}_0, \mathbf{x}_s, \omega) &= \frac{(jk)^{3/2}}{2} \frac{y_0 - y_s}{|\mathbf{x}_0 - \mathbf{x}_s|} H_1^{(1)}(k|\mathbf{x}_0 - \mathbf{x}_s|) \\ &\approx \frac{(jk)^{3/2}}{2} \frac{y_0 - y_s}{|\mathbf{x}_0 - \mathbf{x}_s|} \sqrt{\frac{2}{\pi k|\mathbf{x}_0 - \mathbf{x}_s|}} e^{j(k|\mathbf{x}_0 - \mathbf{x}_s| - \frac{\pi}{2} - \frac{\pi}{4})} \\ &= \frac{(jk)^{3/2}}{2} \frac{y_0 - y_s}{|\mathbf{x}_0 - \mathbf{x}_s|} \sqrt{\frac{2}{\pi k|\mathbf{x}_0 - \mathbf{x}_s|}} e^{j(k|\mathbf{x}_0 - \mathbf{x}_s|)} \cdot j^{-\frac{3}{2}} \\ &= \frac{k}{2} \frac{y_0 - y_s}{|\mathbf{x}_0 - \mathbf{x}_s|} \sqrt{\frac{2}{\pi|\mathbf{x}_0 - \mathbf{x}_s|}} e^{j(k|\mathbf{x}_0 - \mathbf{x}_s|)} \\ &= \frac{k}{\sqrt{2\pi}} \frac{y_0 - y_s}{|\mathbf{x}_0 - \mathbf{x}_s|^{3/2}} e^{j(k|\mathbf{x}_0 - \mathbf{x}_s|)} \end{aligned} \quad (2.10)$$

なお、第2式から第3式への変形においては $e^{-j\frac{3\pi}{4}} = (e^{j\frac{\pi}{2}})^{-\frac{3}{2}} = j^{-\frac{3}{2}}$ を用いた。これを時間と周波数に関する逆フーリエ変換することで、次式のとおり時間領域の駆動関数に変換できる。なお、2次音源は離散的に配置された複数のスピーカに対応するので、2次音源の位置 \mathbf{x}_0 を非負整数のインデクス η を用いて $\mathbf{x}_\eta = (x_\eta, y_\eta)$ と置き換えるこ

とに注意する.

$$\begin{aligned} u(\mathbf{x}_\eta, \mathbf{x}_s, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\omega/c}{\sqrt{2\pi} |\mathbf{x}_\eta - \mathbf{x}_s|^{3/2}} e^{j\frac{\omega}{c} |\mathbf{x}_\eta - \mathbf{x}_s|} \cdot e^{j\omega t} d\omega \\ &= h(t) * \left\{ \frac{y_\eta - y_s}{(2\pi |\mathbf{x}_\eta - \mathbf{x}_s|)^{3/2}} \delta \left(t + \frac{|\mathbf{x}_\eta - \mathbf{x}_s|}{c} \right) \right\}. \end{aligned} \quad (2.11)$$

ここで, $*$ は畳込み演算, t は時間, $\mathbf{x}_s = (x_s, y_s)$ は焦点の位置を表す. $h(t)$ は次式で定義されるフィルタである.

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{\omega}{c} e^{j\omega t} d\omega \quad (2.12)$$

このフィルタは波面合成プレフィルタと呼ばれ, 2次音源の位置に対して独立に定義できる. したがって, 時間領域の焦点音源法は, 波面合成プレフィルタを入力信号に適用した上で, 遅延処理とゲイン乗算を実施することで実現できるため, 周波数領域でのフィルタリングと逆フーリエ変換が必要な周波数領域での実装と比べて低演算量で実現することができる.

一般に, デジタル信号として離散化された信号に対する遅延処理は, デジタル信号のサンプル単位で行うため, これより短い精度の遅延処理 (例えば, 48000 Hz の場合, およそ 0.021 ms 以下の精度) を実現することができない. 単一の焦点音源を生成するだけであれば大きな問題は起らないが, 複数の焦点音源を空間的に密に生成して処理を行うような場合には, 遅延処理の精度低下が問題となる. このような遅延処理の精度改善には非整数遅延制御フィルタを適用する必要があるが, このフィルタの詳細については, 2.1.5 項で説明する.

2.1.5 時間領域実装の音場再現精度改善

焦点音源法は, 周波数領域の駆動関数では個々の実スピーカに対するフィルタ演算が必要になる. これに対し, 時間領域の実装では実スピーカごとに異なるゲイン乗算と遅延処理で代用できるため, 演算量を劇的に削減できる. しかし, 周波数領域で定義されたフィルタでは可能であった任意の精度での非整数サンプルの遅延処理が, 時間領域での実装ではデジタル信号のサンプリング周期に限られるため, 焦点音源が作る音場の精度が低下する. 例えば, ナイキスト周波数 48kHz のシステムを例にあげると, 再現される音場は最大 7.1 mm ($\approx 343.36 \div 48000$) の誤差を含む (音速を 340m/s とした). 特に, 3章や4章で述べるように, 複数の焦点音源を空間的に密に生成して仮想的な指向性音源を生成する場合には, ある焦点音源を作るための駆動関数が別の焦点音源に対しても影響を与えるため, 本来生成する音場を高精度に再現することが困

難となる．そこで本項では，サンプリング周期より短い遅延を適用可能な非整数遅延フィルタを時間領域実装に組み込むことで，再現音場の精度改善が行えることを示す．

非整数遅延フィルタ

非整数遅延フィルタは，サンプリング周期より短い遅延処理を実現する．理想的な非整数遅延フィルタは sinc 関数を用いて実現される [80]．このフィルタは，理論的には無限長の非因果性フィルタである．現実には，フィルタは因果性フィルタとして利用できるよう有限長で打ち切った上で利用する．このフィルタは全域通過型の周波数特性と直線位相を有するものの，少ないタップ長のフィルタとして実装するのが難しいため，フィルタリングに多くの演算量を必要とする．

少ない演算量で非整数サンプルの遅延処理を行うために，様々な実装形態のフィルタが提案されている．FIR 型フィルタの代表例は，次式に基づく Lagrange 補間に基づく方法である [81]．

$$p_{LI}(t) = \prod_{l=0, l \neq t}^M \frac{\tau_{\text{frac}} - l}{t - l}, \quad (2.13)$$

ここで， τ_{frac} は非整数遅延のサンプル数， M はフィルタ長を表す．Lagrange 補間によれば，フィルタ係数は閉じた式で定義できるため遅延量が時間的に変化するような場合でも効率よく実装することができる．Lagrange 補間で定義されたフィルタは，特に低次の場合，群遅延が平坦になるものの低域通過型の周波数特性となるため音質が劣化する．例として，次数 $M = 1$ の Lagrange 補間で求めた非整数サンプル遅延フィルタの群遅延特性と周波数応答を図 2.4，図 2.5 に示す．群遅延特性が低域において比較的フラットなのに対して，群遅延特性が目標となる非整数サンプル遅延の値に応じて異なり，強い低域通過特性を持つことがわかる．また，フィルタ次数の増加に比べて周波数特性の変化が小さいため，平坦な周波数特性を得るためには長いフィルタを用いなければならない，演算量が増大する [82]．

非整数遅延フィルタの別の例は，IIR 型全域通過フィルタとして知られる Thiran フィルタである [83]．この IIR 型フィルタは，FIR 型フィルタに比べて短いフィルタ長で全周波数帯域に渡ってフラットな周波数特性を実現する．フィルタ係数は次式で定義される．

$$p_{\text{Th}}(t, \tau_{\text{frac}}) = (-1)^t \binom{M}{t} \prod_{l=0, l \neq t}^M \frac{\tau_{\text{frac}} - M - l}{\tau_{\text{frac}} - M + l + t}, \quad (2.14)$$

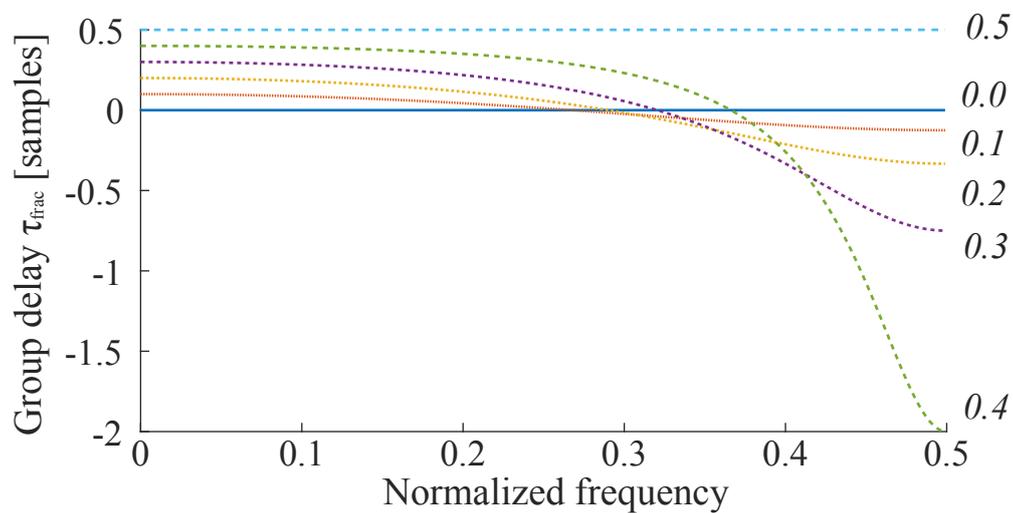


図 2.4: Lagrange 補間で求めた非整数サンプル遅延フィルタの群遅延特性 (サンプリング周波数が 48 kHz の場合, 正規化周波数 0.5 が 24 kHz に相当).

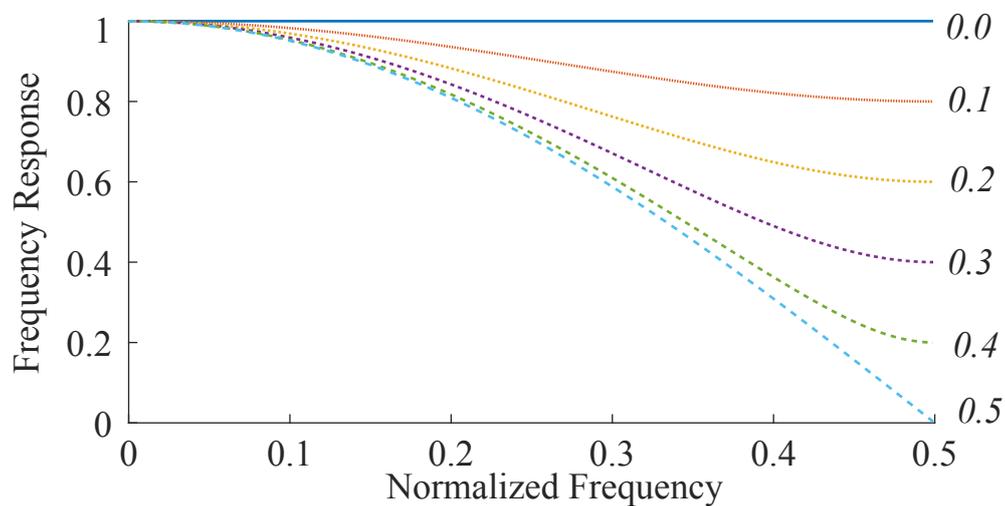


図 2.5: Lagrange 補間で求めた非整数サンプル遅延フィルタの周波数特性 (サンプリング周波数が 48 kHz の場合, 正規化周波数 0.5 が 24 kHz に相当). τ_{frac} は目標となる群遅延.

ここで, τ_{frac} は非整数遅延のサンプル数, M はフィルタ長を表す. Thiran フィルタは平坦な周波数特性を与えるものの, 図 2.6 に示すとおり, 群遅延特性が平坦にならない. 特に高域の群遅延特性は, Lagrange 補間で実現できる群遅延特性よりもはるかに大きな誤差が生じる. しかし, 図 2.7 に示すとおり, 4.8 kHz 以下の周波数帯域に限定してみると線形な位相特性が実現できていることがわかる. 直線状スピーカアレイを用いて波面合成を実装する際には, 有限の大きさを持つスピーカを用いるため, 空間エイリアシングが発生する. 音速を 340 m/s とした時, スピーカ間隔を 0.07 m 以下とすれば, 4.8 kHz 以下の周波数帯域で空間エイリアシングが発生しないので, 実用上, 1 次の Thiran フィルタは十分な性能を有すると言える. 3 章および 4 章では, Thiran フィルタを用いた時間領域実装の精度改善について検討するが, その際には 1 次のフィルタを用いることとする.

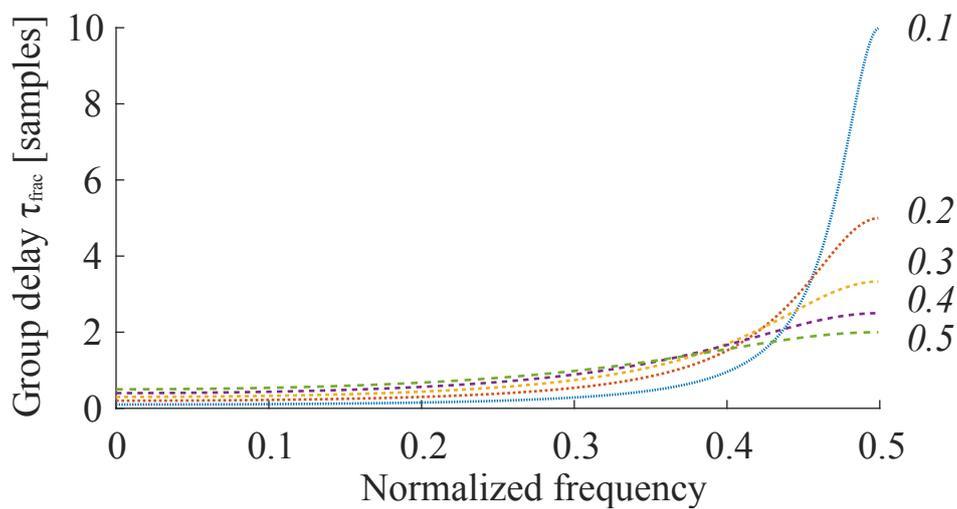


図 2.6: 1次 Thiran フィルタの群遅延特性 (サンプリング周波数が 48 kHz の場合, 正規化周波数 0.5 が 24 kHz に相当). τ_{frac} は目標となる群遅延.

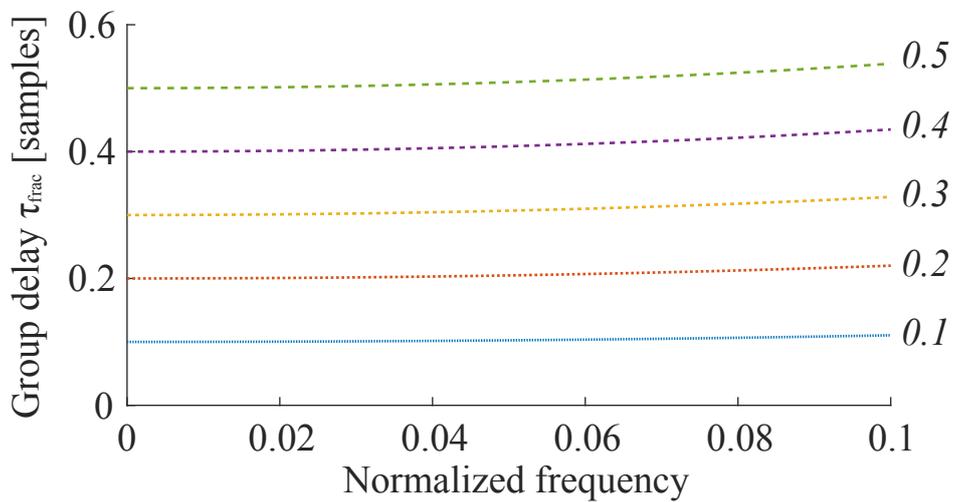


図 2.7: 1次 Thiran フィルタの群遅延特性 (サンプリング周波数が 48 kHz の場合, 正規化周波数 0.1 が 4.8 kHz に相当).

2.2 指向性制御技術

実世界では楽器の他にも頭部・体を含めた人間の発声器官など様々な音源が存在する。その多くには、音が強く伝搬する方向や弱く伝搬する方向のパターン（指向特性）がある。これらを忠実に再現することで、単なる無指向性音源を使って達成できる以上の臨場感を実現することが期待できる。また、音声を明瞭に呈示する場所と、逆に聞き取りにくくする場所を、指向性制御技術で作り返すことも可能となる。このように、音の指向性を制御する技術には幅広い応用の可能性があり大変重要である。

指向性制御技術には様々な手法が提案されているが、本節では2.1節で述べた波面合成技術と関わりの深い空間フーリエ変換をベースとした指向性制御技術と、複数のモノポール音源を密に並べて制御する多重極音源をベースとした指向性制御技術について説明する。

2.2.1 空間フーリエ変換に基づく指向性制御

音響信号は空間を伝搬する波動であり、その挙動は波動方程式に従う [54]。 x 軸に沿って進行する 1 次元の音波 $p(x, t)$ が従う波動方程式は次式のとおりである。

$$\frac{\partial^2}{\partial x^2} p(x, t) = \frac{1}{c^2} \frac{\partial^2 p(x, t)}{\partial t^2} \quad (2.15)$$

ここで、 c は音速、 t は時間を表すインデックスである。この波動方程式の一般解には次式がある。

$$p(x, t) = P(\omega) e^{-jkx} \cdot e^{j\omega t} \quad (2.16)$$

さらに、波動方程式の線型性を考慮すると異なる周波数の独立な解の重ね合わせも式 (2.15) の解となる。

$$p(x, t) = \sum_{\omega} P(\omega) e^{-jkx} \cdot e^{j\omega t} \quad (2.17)$$

ここで、位置 x を原点に固定して、式 (2.17) の和を $-\infty < \omega < \infty$ での積分に置き換えると以下の時間に関する逆フーリエ変換の式を得る。

$$p(t) = \int_{-\infty}^{\infty} P(\omega) e^{j\omega t} d\omega \quad (2.18)$$

一方、時刻を $t = 0$ に固定した場合、波動方程式の解である式 (2.17) は次式で与えられる。

$$p(x) = \sum_k \tilde{P}(k) e^{-jkx} \quad (2.19)$$

ここで、 $\tilde{P}(k)$ は波数 k の波の複素振幅である。この式の両辺に e^{jkx} をかけた上で、右辺の和を $-\infty < x < \infty$ での積分に変換することで次の空間フーリエ変換の式を得る。

$$\tilde{P}(k) = \int_{-\infty}^{\infty} p(x) e^{jkx} dx \quad (2.20)$$

このように、フーリエ変換を用いて時間と周波数を関係付けられるのと同様に、空間フーリエ変換を用いて空間と波数を関係づけられることがわかる。ここまで1次元の波動方程式とその一般解を基底とするフーリエ変換について見てきたが、これを2次元あるいはそれ以上に拡張することが可能である。2次元平面上での波動方程式は次式のとおりである。

$$\nabla^2 p(x, y, t) = \frac{1}{c^2} \frac{\partial^2 p(x, y, t)}{\partial t^2} \quad (2.21)$$

この方程式の一般解は変数分離の方法を用いて以下のように書ける。

$$p(x, y, t) = A(k_x, \omega) e^{-jk_x x} e^{-j\sqrt{k^2 - k_x^2} y} e^{j\omega t} \quad (2.22)$$

さらに、波動方程式の解の線型性を考慮すると、この波動方程式の解は次式のような異なる独立な解の重ね合わせで表すことができる。

$$p(x, y, t) = \sum_{k_x, \omega} A(k_x, \omega) e^{-jk_x x} e^{-j\sqrt{k^2 - k_x^2} y} e^{j\omega t} \quad (2.23)$$

この式は、任意の場所の音圧はあらゆる方向から到来する平面波の和として表現できることを表す。これを音場の平面波展開と呼ぶ。

円座標系と空間フーリエ変換

2.2.1項で、任意の場所の音圧はあらゆる方向から到来する平面波の和として表現できることを確認した。直線状スピーカアレイや平面状スピーカアレイを用いて音場を再現する場合には音場の平面波展開が有用だが、例えば円状にスピーカを配置したスピーカアレイから音場を合成する場合には、デカルト座標系で定義された音場の平面波展開を利用するのは効率的ではない。そこで、デカルト座標系で定義された波動方程式を円筒座標系に変換し、その解を基底とした展開（ヘリカル波展開）について本項で述べる。

最初に図2.8に示す円座標系を導入する。デカルト座標系における任意の位置 $\mathbf{x} = (x, y)$ は次式にしたがい円座標系 $\mathbf{x} = (r, \phi)$ に変換できる。

$$\begin{aligned} x &= r \cos \phi \\ y &= r \sin \phi \end{aligned} \quad (2.24)$$

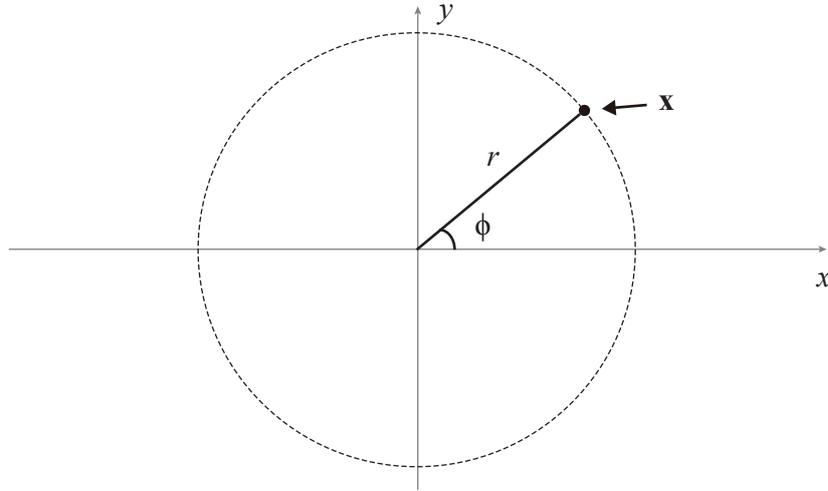


図 2.8: 円座標系

今, 2次元デカルト座標系における波動方程式は, 次式で表される.

$$\nabla^2 p(x, y, t) = \frac{1}{c^2} \frac{\partial^2}{\partial t^2} p(x, y, t) \quad (2.25)$$

ここで, 2次の空間微分 ∇^2 はデカルト座標系では次式で表される.

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad (2.26)$$

これを円座標系に変換すると, 次式が得られる.

$$\nabla^2 = \frac{1}{r} \frac{\partial}{\partial r} + \frac{\partial^2}{\partial r^2} + \frac{1}{r^2} \frac{\partial^2}{\partial \phi^2} \quad (2.27)$$

この波動方程式もデカルト座標系のと様と同様, 変数分離の方法で解を求めることができる. 円座標系における波動方程式の解は, ある円を境界として内部の音場を求める内部問題と, 円の外部の音場を求める外部問題があるが, 本論文では外部問題を対象とするため, 以下に外部問題の解を示す.

$$p(r, \phi, \omega) = \sum_{\nu=-\infty}^{\infty} A_{\nu}(\omega) H_{\nu}^{(2)}(kr) e^{j\nu\phi} \quad (2.28)$$

ここで, ν は展開次数, $A_{\nu}(\omega)$ は展開級数, k は波数, $H_{\nu}^{(2)}$ は ν 次の第2種ハンケル関数, $e^{j\nu\phi}$ は円調和関数である. したがって, 円座標系における外部問題の解の導出は, 音場を円調和級数展開した際の展開係数 $A_{\nu}(\omega)$ を求めることに等しい.

次に, $e^{j\nu\phi}$ に着目し, 半径 r の円周上で観測された音圧 $p(r, \phi, \omega)$ を円周方向に空間フーリエ変換すると,

$$\tilde{P}_{\nu}(r, \omega) = \frac{1}{2\pi} \int_0^{2\pi} p(r, \phi, \omega) e^{-j\nu\phi} d\phi \quad (2.29)$$

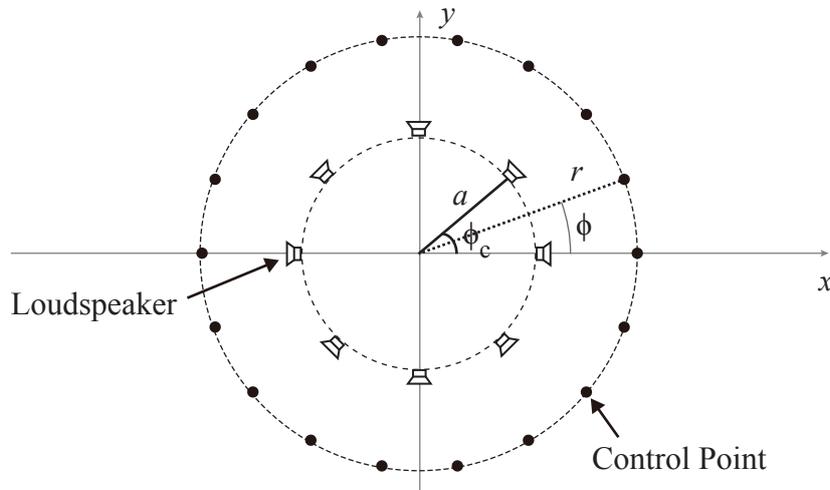


図 2.9: 円状アレイと制御点の位置関係

となる。また、次式に示す空間逆フーリエ変換は円調和展開である。

$$p(r, \phi, \omega) = \sum_{\nu=-\infty}^{\infty} \tilde{P}_{\nu}(r, \omega) e^{j\nu\phi} \quad (2.30)$$

$\tilde{P}_{\nu}(r, \omega)$ は円調和スペクトルと呼ばれており、次数 ν は円周上に入る波の数を表す。すなわち $k_{\phi} = \nu/r$ の関係が成り立つ。円周方向のフーリエ変換は、円周上の音圧分布を $e^{j\nu\phi}$ の成分に分解していることに相当する。これは、モノポールやダイポールの指向特性の成分に分解することと等価ともみなすことができる。モノポールやダイポールの成分への展開は多重極展開と呼ばれ、円調和展開とは密接な関係があるが詳細は5章にて議論する。

円状アレイによる指向性再生

直線や円などの幾何形状にスピーカを並べてできるスピーカアレイを制御することで、指向性音源が生成する音場を再現することができる。直線状スピーカアレイを用いて指向性を再現する手法には2.3節に述べるような手法があるが、ここでは円状にスピーカを並べてできる円状アレイによる指向性音源の音場再現について述べる。図2.9に示すとおり、座標の原点に目標となる音場を生成する指向性音源を配置し、指向性音源を囲む形で制御点を円状に配置する。解くべき課題としては、指向性音源が制御点において生成する音圧分布を、半径 a ($< r$) の円状にスピーカを並べてできるスピーカアレイを制御することで再現することである。このような課題の解法として、以下に述べる Pressure-matching method と、 Mode-matching method が知られている

[51]. 以下に順に説明する.

Pressure-matching method

今, 半径 a の円状に \mathcal{L} 個のスピーカを位置 $\mathbf{a}_l = (a, \phi_l)$ に配置する ($0 \leq l \leq \mathcal{L} - 1$). また, 半径 a より十分大きい半径 r の円上に \mathcal{M} 個の制御点を位置 $\mathbf{r}_\eta = (r, \phi_\eta)$ に配置する ($0 \leq \eta \leq \mathcal{M} - 1$). η 番目の制御点 \mathbf{r}_η と, l 番目のスピーカを \mathbf{a}_l の間の伝達関数を $G(\mathbf{r}_\eta, \mathbf{a}_l, \omega)$, l 番目のスピーカに与える駆動関数を $D_l(\omega)$ とすると, この円状アレイが η 番目の制御点 \mathbf{r}_η に生成する音圧は次式で表される.

$$\tilde{p}(\mathbf{r}_\eta, \omega) = \sum_{l=0}^{\mathcal{L}-1} G(\mathbf{r}_\eta, \mathbf{a}_l, \omega) D_l(\omega). \quad (2.31)$$

所望の音場について, すべての制御点での音圧分布を一行に並べて以下のベクトルを構成する.

$$\mathbf{p}_\omega = [p(\mathbf{r}_0, \omega) \ p(\mathbf{r}_1, \omega) \ \cdots \ p(\mathbf{r}_{\mathcal{M}-1}, \omega)]^T \quad (2.32)$$

これを, 式 (2.31) で表せる円状スピーカアレイが制御点で再現する音圧分布 $\tilde{p}(\mathbf{r}_\eta, \omega)$ と一致させるためには, 以下の損失関数 \mathcal{J} を最小化するような駆動信号 $D_l(\omega)$ を求める必要がある.

$$\mathcal{J}(\omega) = \|\mathbf{p}_\omega - \mathbf{G}_\omega \mathbf{D}_\omega\|_2^2 \quad (2.33)$$

ここで, $\mathbf{D}_\omega = [D_0(\omega) \ D_1(\omega) \ \cdots \ D_{\mathcal{L}-1}(\omega)]^T$,

$$\mathbf{G}_\omega = \begin{pmatrix} G(\mathbf{r}_0, \mathbf{a}_0, \omega) & G(\mathbf{r}_0, \mathbf{a}_1, \omega) & \cdots & G(\mathbf{r}_0, \mathbf{a}_{\mathcal{L}-1}, \omega) \\ G(\mathbf{r}_1, \mathbf{a}_0, \omega) & G(\mathbf{r}_1, \mathbf{a}_1, \omega) & \cdots & G(\mathbf{r}_1, \mathbf{a}_{\mathcal{L}-1}, \omega) \\ \vdots & \vdots & \ddots & \vdots \\ G(\mathbf{r}_{\mathcal{M}-1}, \mathbf{a}_0, \omega) & G(\mathbf{r}_{\mathcal{M}-1}, \mathbf{a}_1, \omega) & \cdots & G(\mathbf{r}_{\mathcal{M}-1}, \mathbf{a}_{\mathcal{L}-1}, \omega) \end{pmatrix} \quad (2.34)$$

この問題は, \mathbf{G}_ω の擬似逆行列を用いることで, 以下のとおりに解くことができる.

$$\mathbf{D}_\omega = (\mathbf{G}_\omega^H \mathbf{G}_\omega + \lambda \mathbf{I})^{-1} \mathbf{G}_\omega^H \mathbf{p}_\omega, \quad (2.35)$$

なお, $[\cdot]^H$ は行列の共役転置を表す. また λ は, 擬似逆行列が発散するのを防ぐために導入する正則化パラメータである. このようにスピーカから再生した音響信号の重ね合わせが目標音場の音圧分布と一致するような駆動関数を求める手法を Pressure-matching method と呼ぶ [84] [85]. Pressure-matching method では周波数ごとに逆行列を求めて駆動信号を得るため, 駆動信号は周波数ごとに異なるゲインを持つ. そのため駆動関数の周波数特性はフラットにならず, 帯域により強調や抑圧が起こって音

質が劣化する [52]. なるべくフラットな周波数特性の駆動関数を得るためには, 正則化パラメータのチューニングなどが必要となるが, チューニングを行っても完全にフラットな周波数特性を得ることは難しく, Pressure-matching method の課題の一つとなっている.

Mode-matching method

ここでは, Pressure-matching method とは異なり, 円調和級数展開を用いて駆動関数を導出する方法について述べる. 半径 a 上に連続的に配置されたスピーカが再現する音場は, 位置 $\mathbf{a}_l = (a_l, \phi_l)$ の駆動関数を $D_l(\omega)$ とすると次式で表される.

$$\tilde{p}(r, \phi, \omega) = \int_0^{2\pi} G(r, \phi | a, \phi_l, \omega) D_l(\omega) a d\phi_l \quad (2.36)$$

伝達関数 $G(r, \phi | a, \phi_l, \omega)$ と, 駆動関数 $D_l(\omega)$ をそれぞれ円調和展開すると,

$$G(r, \phi | a, \phi_l, \omega) = \sum_{\nu=-\infty}^{\infty} \check{G}_\nu(r | a, k) e^{j\nu(\phi - \phi_l)} \quad (2.37)$$

$$D_l(k) = \sum_{\nu=-\infty}^{\infty} \check{D}_\nu(k) e^{j\nu\phi_l} \quad (2.38)$$

となる. ここで, $\check{G}_\nu(r | a, k)$, $\check{D}_\nu(k)$ は, それぞれ $G(r, \phi | a, \phi_l, \omega)$, $D_l(\omega)$ の円調和展開係数である. これら円調和展開係数を式 (2.36) に代入すると, 円状アレイが再現する音場は次式のとおりになる.

$$\tilde{p}(r, \phi, \omega) = 2\pi a \sum_{\nu=-\infty}^{\infty} \check{G}_\nu\left(r | a, \frac{\omega}{c}\right) \check{D}_\nu\left(\frac{\omega}{c}\right) e^{j\nu\phi} \quad (2.39)$$

一方, 任意の音場 $p(r, \phi, \omega)$ は以下のとおり円調和展開できる.

$$p(r, \phi, \omega) = \sum_{\nu=-\infty}^{\infty} \check{p}_\nu\left(r, \frac{\omega}{c}\right) e^{j\nu\phi} \quad (2.40)$$

ここで, $\check{p}_\nu(r, \frac{\omega}{c})$ は, 音場 $p(r, \phi, \omega)$ の円調和展開係数である.

上の2式のうち ν 次の項のみを取り出して比較すると次式を得る.

$$\check{p}_\nu\left(r, \frac{\omega}{c}\right) = 2\pi a \check{G}_\nu\left(r | a, \frac{\omega}{c}\right) \check{D}_\nu\left(\frac{\omega}{c}\right) \quad (2.41)$$

これより駆動関数の円調和展開級数は以下のとおり導出できる.

$$\check{D}_\nu\left(\frac{\omega}{c}\right) = \frac{\check{p}_\nu\left(r, \frac{\omega}{c}\right)}{2\pi a \check{G}_\nu\left(r | a, \frac{\omega}{c}\right)} \quad (2.42)$$

このように、目標音場の円調和展開係数と円状アレイの再現音場の円調和展開係数の係数比較により駆動関数を導出する手法を Mode-matching method と呼ぶ。Mode-matching method では Pressure-matching method のように逆行列を計算することなく駆動関数を得られるため、高速に駆動関数を算出することができる。周波数領域での駆動信号は、式 (2.42) の円調和展開係数を逆フーリエ変換して次式のとおり算出できる。

$$\begin{aligned} D_l(\omega) &= \sum_{\nu=-\infty}^{\infty} \check{D}_\nu \left(\frac{\omega}{c} \right) e^{j\nu\phi_l} \\ &= \sum_{\nu=-\infty}^{\infty} \frac{\check{p}_\nu(r, \frac{\omega}{c})}{2\pi a \check{G}_\nu(r|a, \frac{\omega}{c})} e^{j\nu\phi_l} \end{aligned} \quad (2.43)$$

なお、ここまで連続な 2 次音源を仮定して導出を進めてきたが、実際には離散的に配置された 2 次音源（スピーカアレイ）を用いて音場を再現する。離散的に配置された 2 次音源についても上記と同様な導出を経て次式のとおり、周波数領域の駆動関数を導出することができる。

$$D_l(\omega) = \sum_{\nu=-\infty}^{\infty} \frac{\check{p}_\nu(r, \frac{\omega}{c})}{\mathcal{L} \check{G}_\nu(r|a, \frac{\omega}{c})} e^{j\nu\phi_l} \quad (2.44)$$

ここまで、円調和展開係数は次数 ν を無限に利用できる前提で議論してきたが、計算機上で無限級数を計算することは困難であるので、十分大きい次数 N で駆動関数の計算を打ち切る。

$$D_l(\omega) = \sum_{\nu=-N}^N \frac{\check{p}_\nu(r, \frac{\omega}{c})}{\mathcal{L} \check{G}_\nu(r|a, \frac{\omega}{c})} e^{j\nu\phi_l} \quad (2.45)$$

本来、無限個のスピーカが必要なところを離散的に配置した有限個のスピーカで近似するため、表現可能な次数 N はスピーカの数 \mathcal{L} に制約を受ける。円調和展開では N 次の音源を $2N + 1$ 個の直交基底で表現する。逆にいえば、 \mathcal{L} 個のスピーカでは、 $N \leq (\mathcal{L} - 1)/2$ 次までの音場しか再現できない。したがって、最大打ち切り次数は次式で決定できる。

$$N_{\mathcal{L}} = \left\lfloor \frac{\mathcal{L} - 1}{2} \right\rfloor \quad (2.46)$$

なお、 $\lfloor \cdot \rfloor$ は床関数 $\lfloor z \rfloor = \max\{n \in \mathbb{Z} | n \leq z\}$ である。今、円調和展開の上限の次数について考察したが、再現対象の周波数によっては別の制約が必要となる場合がある。

再現対象の周波数に応じた次数打ち切りが必要となる場合について考察するために、目標音場が以下のような円調和級数で表せる場合を考える。

$$p(r, \phi, \omega) = \sum_{\nu=-\infty}^{\infty} \check{S}_\nu(r, k) H_\nu^{(2)}(kr) e^{j\nu\phi} \quad (2.47)$$

この式において、円調和展開係数は以下の式となる。

$$\check{p}_\nu(r, k) = \check{S}_\nu(r, k) H_\nu^{(2)}(kr) \quad (2.48)$$

ここで、2次元グリーン関数の円調和展開は次式で与えられる。

$$\begin{aligned} G(r, \phi | a, \phi_l, \omega) &= \frac{j}{4} H_0^{(2)} \left(\frac{\omega}{c} |\mathbf{r} - \mathbf{a}| \right) \\ &= \frac{j}{4} \sum_{\nu=-\infty}^{\infty} J_\nu \left(\frac{\omega}{c} a \right) H_\nu^{(2)} \left(\frac{\omega}{c} r \right) e^{j\nu(\phi - \phi_l)} \end{aligned} \quad (2.49)$$

これを式(2.37)と比較することで、グリーン関数の円調和展開係数は次式のように求められる。

$$\check{G}_\nu \left(r | a, \frac{\omega}{c} \right) = \frac{j}{4} J_\nu \left(\frac{\omega}{c} a \right) H_\nu^{(2)} \left(\frac{\omega}{c} r \right) \quad (2.50)$$

式(2.48)と式(2.50)を式(2.43)に代入すると、駆動関数は次式で与えられる。

$$D_l(\omega) = \sum_{\nu=-N}^N \frac{4\check{S}_\nu(r, \frac{\omega}{c})}{j\mathcal{L}J_\nu(\frac{\omega}{c}a)} e^{j\nu\phi_l} \quad (2.51)$$

ここで、式(2.51)の分母に現れる ν 次のベッセル関数 $J_\nu(x)$ に着目する。図2.10に、 $a = 0.2$ とした場合の ν 次のベッセル関数 $J_\nu(x)$ を示す。ベッセル関数 $J_\nu(x)$ は、0次のときのみ $x = 0$ で1の値をもち、他の次数では0の値を持つ。1000 Hz以下の領域に着目すると、 $\nu > 0$ の次数のベッセル関数は高次になるほど広い区間で0に近い値を持つ。ベッセル関数は駆動関数である式(2.51)の分母に現れるので、高次のベッセル関数まで駆動関数の計算に用いると、1000 Hz以下の領域で駆動関数が非常に大きな値となり音質を劣化させる。したがって、再現対象の周波数に応じて駆動関数の算出に用いる次数を制限する必要がある。打ち切り次数の決定には様々な手法が提案されている [86][87][88][89]。

そのうちのひとつに、任意の音場が平面波の重ね合わせで実現できる性質を利用して、所望の平面波が事前に設定した精度で再現できる次数まででベッセル関数の計算を打ち切る方法がある。 ϕ' 方向から到来する平面波が作る音場は円調和展開を用いて次式で与えられる。

$$\begin{aligned} p(r, \phi, \omega) &= e^{jkr \cos(\phi - \phi')} \\ &= \sum_{\nu=-\infty}^{\infty} j^\nu J_\nu(kr) e^{j\nu(\phi - \phi')} \end{aligned} \quad (2.52)$$

この音場の平面波展開を有限の次数で打ち切った際の誤差から、次式にしたがって次数を打ち切ることで平面波の音場再現誤差をおよそ4%以下にできることが示されて

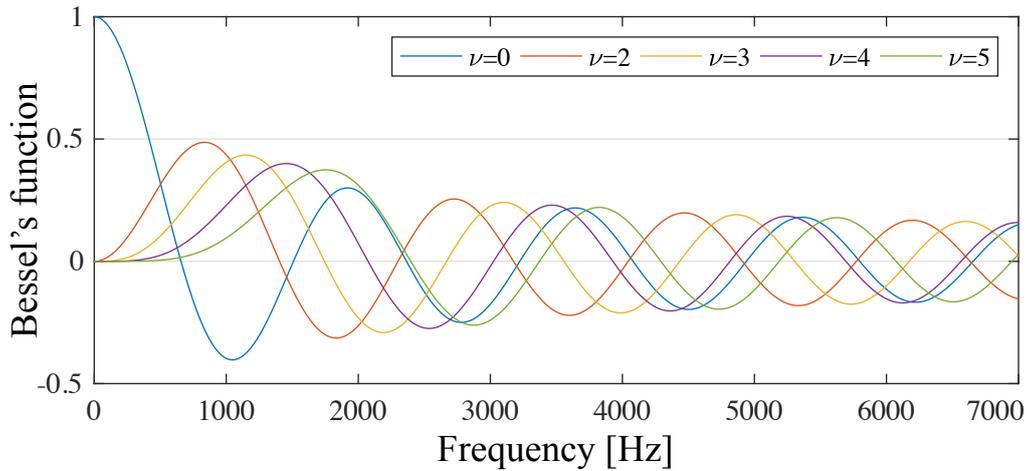


図 2.10: ベッセル関数の挙動. 円状アレイの半径を 0.2 m とした場合.

いる.

$$N = \lceil kr \rceil \quad (2.53)$$

ここで, $\lceil \cdot \rceil$ は天井関数 $\lceil z \rceil = \min\{n \in \mathbb{Z} | n \geq z\}$ である. 以上述べたように, 駆動関数が極端に大きな値をとることがないように, 特に低域では式 (2.46) 及び式 (2.53) を用いて円調和展開の最大次数を制御する方法がとられる.

2.2.2 多重極音源に基づく指向性制御

同じ振幅で極性が反対の 2 個の音源が空間的に密に存在しているとき, この音源の組は 8 の字の指向特性を持つ. このような音源はダイポール音源と呼ばれる. 複数の音源が固有の極性を持ち空間的に密に配置されて固有の指向特性を生成するとき, これらの音源の集合を多重極音源と呼ぶ. ダイポール音源の他にも 2 次の多重極音源であるクアドラポールや, それ以上の次数の多重極音源も存在する. これら多重極音源は波動方程式のグリーン関数の高次微分に相当する音場を再現することが知られており, 多重極音源に重み係数をかけて重ね合わせることは音場をテイラー展開することに相当する [90]. このように所望の音場を高次多重極音源の重ね合わせで表現し, 各多重極音源に与える重み係数を算出することを多重極展開と呼ぶ. 2.2.1 項では, 音場の再現に空間フーリエ変換を用いたが, 多重極展開を用いて同じ音場を再現できることが知られており, これを用いた音場制御や指向性制御の検討が進められてきた [55] [57] [58] [59] [60] [91]. また, 電磁気学などの分野において, 複雑な電荷分布が作る電場を効率良く計算する手法としても一般的に利用されている [92] [93] [94] [95]. 以下, 多重極音源について概要を説明した上で, 多重極展開の指向性制御への応用について述

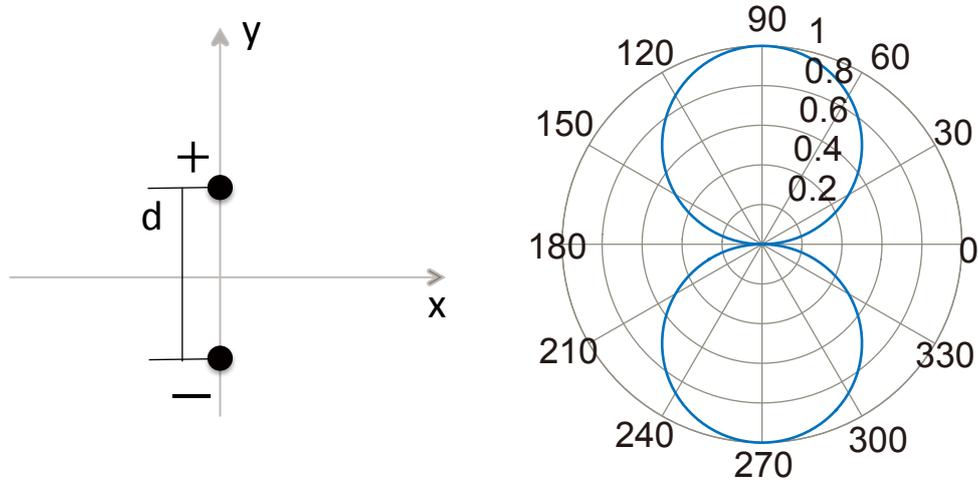


図 2.11: 左図: y 軸方向のダイポール音源とダイポール音源を構成するモノポール音源の位置関係. 右図: y 軸方向のダイポール音源が持つ指向特性.

べる. まずは, もっとも単純なダイポール音源と多重極展開の関係から説明を始める.

ダイポール音源

最初にダイポール音源が作る音場について考える. 図 2.11 に示すとおり, 2 個のモノポール音源を距離 d 離れたところに配置し, 片方の極性を反転させるとダイポール音源が実現できる [54]. ダイポール音源が位置 $\mathbf{r} = (r, \phi)$ に作る音場は以下の式で表現できる.

$$\begin{aligned}
 S_{\text{dipole}}(\mathbf{r}, k) &= \frac{j}{4} \{ H_0^{(2)}(k|\mathbf{r} - \Delta|) - H_0^{(2)}(k|\mathbf{r} + \Delta|) \} \\
 &= \frac{j}{4} \Delta \sum_{n=-\infty}^{\infty} \left\{ \frac{J_n(k\Delta)}{\Delta} - (-1)^n \frac{J_n(-k\Delta)}{\Delta} \right\} H_n^{(2)}(kr) e^{jn\phi} \quad (2.54)
 \end{aligned}$$

なお, $d = 2\Delta$ とする.

ここで, ベッセル関数についてロピタルの定理を用いると, 次の関係式が成り立つ.

$$\lim_{\Delta \rightarrow 0} \frac{J_n(k\Delta)}{\Delta} = \lim_{\Delta \rightarrow 0} k \frac{dJ_n(z)}{dz} \Big|_{z=k\Delta} = k \frac{J_{n-1}(0) - J_{n+1}(0)}{2} \quad (2.55)$$

また, ベッセル関数について次式が成り立つ.

$$J_\nu(x) = 0 \quad (\nu \neq 0) \quad (2.56)$$

よって, 式 (2.54) の最後の式は次のとおり変形できる.

$$\begin{aligned}
S_{\text{dipole}}(\mathbf{r}, k) &= \frac{j}{4} \Delta k \left\{ H_1^{(2)}(kr) e^{j\phi} - H_{-1}^{(2)}(kr) e^{-j\phi} \right\} \\
&= \frac{j}{4} \Delta k \left\{ H_1^{(2)}(kr) (e^{j\phi} + e^{-j\phi}) \right\} \\
&\approx \frac{j}{4} \Delta k \left\{ 2H_1^{(2)}(kr) \cos \phi \right\} \\
&\approx \frac{j}{4} \left\{ H_0^{(2)}(kr) \cdot jdk \cos \phi \right\} \\
&\approx G_{2D}(kr) \{ jdk \cos \phi \}
\end{aligned} \tag{2.57}$$

ここで、グリーン関数 G_{2D} は式 (2.3) の 2 次元平面で定義された非斉次波動方程式のグリーン関数である。また、式の変形には $H_{-n}^{(2)}(z) = (-1)^n H_n^{(2)}(z)$ と $H_n^{(2)}(z) \approx j^n H_0^{(2)}(z)$ の関係式 [79] [96] を用いた。

一方、グリーン関数 G_{2D} を x 軸方向に微分して次式を得る。

$$\begin{aligned}
\frac{\partial G_{2D}(kr)}{\partial x} &= \frac{j}{4} \left(k \frac{dr}{dx} \right) \frac{dH_0^{(2)}(z)}{dz} \Big|_{z=kr} \\
&= \frac{j}{4} (k \cos \phi) \frac{H_{-1}^{(2)}(z) - H_1^{(2)}(z)}{2} \Big|_{z=kr} \\
&= \frac{j}{4} (k \cos \phi) \left(-H_1^{(2)}(z) \right) \Big|_{z=kr} \\
&= \frac{j}{4} (k \cos \phi) \left(-jH_0^{(2)}(z) \right) \Big|_{z=kr} \\
&= \frac{j}{4} H_0^{(2)}(kr) (-jk \cos \phi) \\
&= G_{2D}(kr) (-jk \cos \phi)
\end{aligned} \tag{2.58}$$

ここで、 ϕ は観測点の偏角で $\cos \phi = x/r$ を満たす。式 (2.57) と式 (2.58) を比較すると、ダイポール音源が作る音場は 2 次元グリーン関数 G_{2D} の 1 次微分にスピーカ間隔から決まる係数 $-d$ が掛かったものと一致することがわかる [59]。ここでは、 x 軸方向に微分することで 2 次元グリーン関数の 1 次微分を求めたが、これを y 軸方向に微分した場合には、 $\cos \phi$ の代わりに $\sin \phi$ が現れることは容易に確認できるため、導出は省略する。

高次の多重極音源

ダイポール音源の例から、自由空間グリーン関数を空間微分することで多重極音源が得られることがわかった。ダイポール音源同様、自由空間グリーン関数に対して空間微分を繰り返すことで高次の多重極音源が得られることが知られている [59]。本項

では, 2次元の場合にグリーン関数の高次微分が作る音場が高次の多重極音源が作る音場に比例することを示す.

ダイポール音源を間隔 $d (= 2\Delta)$ で配置して, 片方の極性を反転させた場合にできる音場 $S_{\text{quad}}(\mathbf{r}, k)$ は次式で与えられる.

$$\begin{aligned} S_{\text{quad}}(\mathbf{r}, k) &= S_{\text{dipole}}(\mathbf{r} + \Delta, k) - S_{\text{dipole}}(\mathbf{r} - \Delta, k) \\ &= (jdk \cos \phi) \left\{ G_{2D}(k|\mathbf{r} - \Delta|) - G_{2D}(k|\mathbf{r} + \Delta|) \right\} \end{aligned} \quad (2.59)$$

上式は式 (2.54) を用いて以下のとおり変形できる.

$$\begin{aligned} S_{\text{quad}}(\mathbf{r}, k) &= (jdk \cos \phi) \left\{ G_{2D}(k|\mathbf{r} - \Delta|) - G_{2D}(k|\mathbf{r} + \Delta|) \right\} \\ &= (jdk \cos \phi) S_{\text{dipole}}(kr) \\ &= (jdk \cos \phi)^2 G_{2D}(kr) \end{aligned} \quad (2.60)$$

ダイポール音源を間隔 d 離して配置することで形成した 2 次の多重極音源 (クアドラポール音源) が作る音場は, 元の 2 次元グリーン関数 G_{2D} に対して $(jdk \cos \phi)^2$ を掛けたものと一致することが確認できた. 同様にして, x 軸方向に m 次, y 軸方向に n 次の多重極となる (m, n) 次多重極音源が作る音場は, 次式で与えられることがわかる.

$$S_{m,n}(kr) = G_{2D}(kr) (jdk)^{m+n} \cos^m \phi \sin^n \phi \quad (2.61)$$

一方, グリーン関数 G_{2D} の 1 次微分に対する x 軸方向の 1 回微分は次式で与えられる.

$$\begin{aligned} \frac{\partial^2 G_{2D}(kr)}{\partial x^2} &= (-jk \cos \phi) \left. \frac{dG_{2D}(z)}{dz} \right|_{z=kr} \\ &= G_{2D}(kr) (-jk \cos \phi)^2 \end{aligned} \quad (2.62)$$

同様の計算を繰り返すことで, 2次元グリーン関数 $G_{2D}(kr)$ に対して x 軸方向に m 回, y 軸方向に n 回微分したものが作る音場は次式で与えられる.

$$\frac{\partial^{m+n} G_{2D}(kr)}{\partial x^m \partial y^n} = G_{2D}(kr) (-jk)^{m+n} \cos^m \phi \sin^n \phi \quad (2.63)$$

式 (2.61) と式 (2.63) を比較すると, (m, n) 次の多重極音源が作る音場は 2 次元グリーン関数 G_{2D} の (m, n) 次微分が作る音場に $(-d)^{m+n}$ を掛けたものと一致することがわかる.

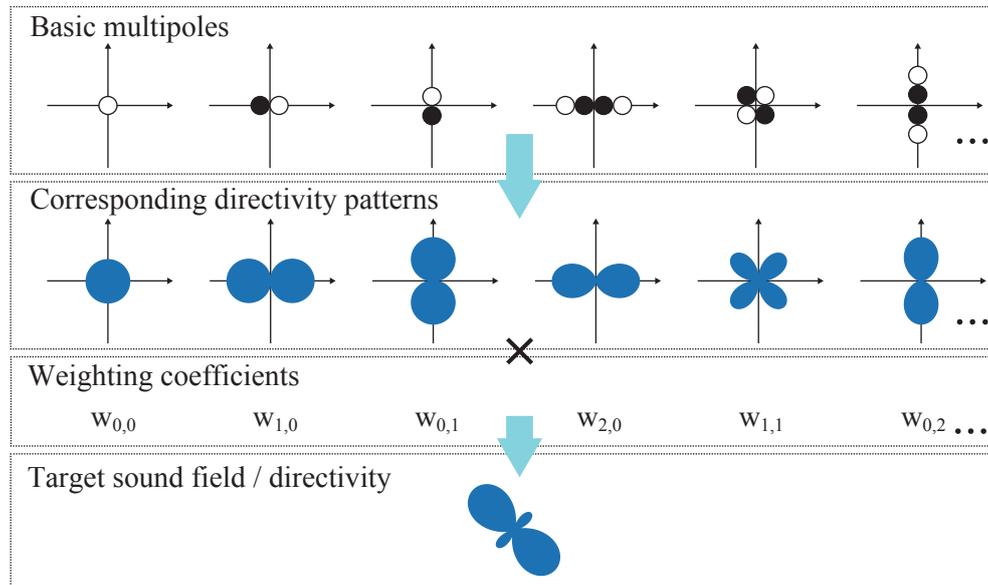


図 2.12: 多重極展開の概念図.

2.2.3 多重極展開を用いた指向性制御

2.2.1 項で、音場が円調和級数展開により表現できることを確認した。これは、音場を円調和級数展開のように直交基底 $e^{j\nu\phi}$ の線形和で表現する他に、多重極音源の重み付き和により表現できることを示している。ここで、所望の音場を多重極音源の重ね合わせで定義し、各多重極音源に乗ずる重み係数を算出して音場を表現する手法を多重極展開と呼ぶ。このとき、所望の音場は 2.2.2 項の議論から次式により定義できる。

$$S(\mathbf{r}, \omega) = \sum_{m,n} w_{m,n} \cdot \frac{\partial^{m+n}}{\partial x^m \partial y^n} G_{2D}(kr) \quad (2.64)$$

ここで、 m, n はそれぞれ x 軸方向と y 軸方向への微分の回数を表す非負整数のインデックスであり、 $w_{m,n}$ は (m, n) 次の多重極音源に対する重み係数であり、 $G_{2D}(kr)$ は 2 次元平面のグリーン関数である。多重極展開の概念図を図 2.12 に示す。

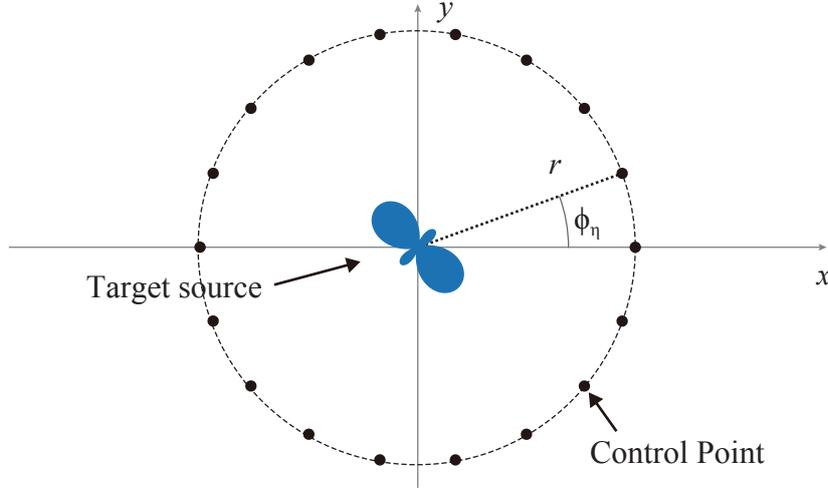


図 2.13: 多重極展開における指向性音源と制御点の位置関係.

次に所望の音場が与えられたとき, 各多重極音源に乗ずる重み係数を算出する手法について述べる. ここでは 2.2.1 項で述べた Pressure-matching method に基づき, 重み係数を算出する. 今, 座標系の中央に音場あるいは指向特性の再現対象となる指向性音源があるとして, これを囲む形で図 2.13 のように M 個の制御点 \mathbf{r}_η を半径 r の円状に等間隔に配置する. このとき, 多重極音源の重ね合わせにより制御点 \mathbf{r}_η における音圧分布は式 (2.64) で与えられる.

さて, 所望の音場について, すべての制御点での音圧分布を一行に並べて以下のベクトルを構成する.

$$\mathbf{p}_\omega = [p(\mathbf{r}_0, \omega) \ p(\mathbf{r}_1, \omega) \ \cdots \ p(\mathbf{r}_{M-1}, \omega)]^T \quad (2.65)$$

また, 重み係数 $w_{m,n}$ についても一行に並べて以下のベクトルを構成する. ここで, N は多重極の最大次数を表す.

$$\mathbf{w}_\omega = [w_{0,0} \ w_{0,1} \ w_{1,0} \ w_{0,2} \ w_{1,1} \ w_{2,0} \ \cdots \ w_{N-1,1} \ w_{N,0}]^T \quad (2.66)$$

これを, 式 (2.64) に従い多重極音源の重ね合わせが制御点で再現する音圧分布 $\tilde{p}(\mathbf{r}_\eta, \omega)$ と一致させるためには, 以下の損失関数 \mathcal{J} を最小化するような重み係数 $w_{m,n}(\omega)$ を求める必要がある.

$$\mathcal{J}(\omega) = \|\mathbf{p}_\omega - \mathbf{G}_\omega \mathbf{w}_\omega\|_2^2 \quad (2.67)$$

ここで、伝達関数行列 \mathbf{G}_ω は次式で定義できる.

$$\mathbf{G}_\omega = \begin{pmatrix} \dot{g}_{0,0}(r, \phi_0) & \dot{g}_{0,1}(r, \phi_0) & \cdots & \dot{g}_{N,0}(r, \phi_0) \\ \dot{g}_{0,0}(r, \phi_1) & \dot{g}_{0,1}(r, \phi_1) & \cdots & \dot{g}_{N,0}(r, \phi_1) \\ \vdots & \vdots & \ddots & \vdots \\ \dot{g}_{0,0}(r, \phi_{M-1}) & \dot{g}_{0,1}(r, \phi_{M-1}) & \cdots & \dot{g}_{N,0}(r, \phi_{M-1}) \end{pmatrix} \quad (2.68)$$

なお、 $\dot{g}_{m,n}(r, \phi)$ は式 (2.63) で表される (m, n) 次の多重極音源の伝達関数であり、以下の式にしたがう.

$$\dot{g}_{m,n}(r, \phi) = \frac{\partial^{m+n} G_{2D}(kr)}{\partial x^m \partial y^n} = G_{2D}(kr) (-jk)^{m+n} \cos^m \phi \sin^n \phi \quad (2.69)$$

この問題は、 \mathbf{G}_ω の擬似逆行列を用いることで、以下のとおりに解くことができる.

$$\mathbf{w}_\omega = (\mathbf{G}_\omega^H \mathbf{G}_\omega + \lambda \mathbf{I})^{-1} \mathbf{G}_\omega^H \mathbf{p}_\omega \quad (2.70)$$

なお、 $[\cdot]^H$ は行列の共役転置を表す. また λ は、擬似逆行列が発散するのを防ぐために導入する正則化パラメータである.

2.3 直線状アレイを用いた指向性音源の音場再現

指向性を有する音源が生成する音場を、直線状スピーカアレイを用いて再現する波面合成技術も提案されている. 式 (2.5) において、目標音場 $S(\mathbf{x}, \omega)$ として指向性音源が作る音場を用いることで、指向性音源が生成する音場を再現するための駆動関数を導出することができる. まずは、スピーカアレイよりも音源が後方にある場合について駆動関数の導出 [65] を示し、これを拡張してスピーカよりも飛び出す仮想音源を生成する際の駆動関数の導出 [68] を示す.

さて、2次元平面において指向性音源が作る音場は、次式に示すとおり、円調和級数展開によりモデル化できる [65].

$$S(\mathbf{r}, \omega) = \sum_{\nu=-\infty}^{\infty} \check{S}_\nu^{(2)}(\omega) H_\nu^{(2)}(kr) e^{j\nu\alpha} \quad (2.71)$$

ここで、 $\check{S}_\nu^{(2)}$ は ν 次の円調和級数を表し、 $H_\nu^{(2)}$ は ν 次の第2種ハンケル関数を表す. r および α は任意の位置 \mathbf{x} を極座標で表示したときの動径および偏角を表す. この手法で用いる座標系を図 2.14 に示す. 式 (2.71) で定義される指向性音源の音場に対して、 y 軸方向に空間微分を行うことで、位置 \mathbf{x}_0 に配置したスピーカの駆動関数を以下の式のとおり導出することができる.

$$D_{\text{ref}}(\mathbf{x}_0) = jk \sum_{\nu=-\infty}^{\infty} \check{S}_\nu^{(2)}(\omega) \{ H_{\nu-1}^{(2)}(kr_0) e^{j(\nu-1)\alpha_0} + H_{\nu+1}^{(2)}(kr_0) e^{j(\nu+1)\alpha_0} \} \quad (2.72)$$

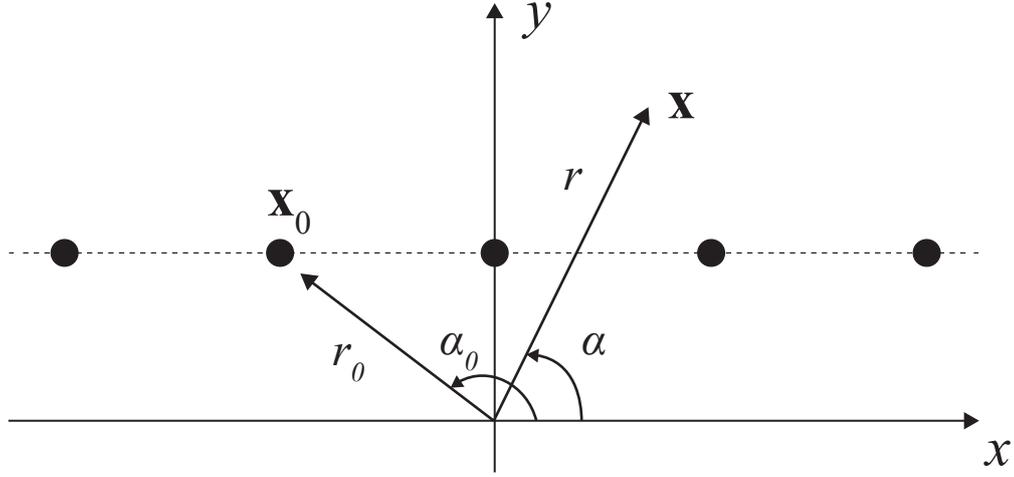


図 2.14: 2.3 節の手法で用いる座標系 [65].

ここで, r_0 は原点にある仮想音源から位置 \mathbf{x}_0 のスピーカまでの距離, α_0 はその偏角である.

スピーカより前面に飛び出す仮想音源を作るには, 目標となる指向特性を有する吸い込み型の仮想音源が作る音場に対して式 (2.5) を適用して駆動関数を導出する. ここで, 焦点音源の位置に向かって収束した音が焦点を過ぎた後, 発散して指向性音源の音場を作ることと考えると, 目標となる音場をあらかじめ 180° 回転させておかなければならないことに注意する [68]. 以上をふまえて, 円調和展開級数 $\check{S}_\nu^{(1)}(\omega)$ でモデル化された指向特性を持つ焦点音源を生成するための駆動関数は次式で定義できる.

$$D_{\text{ref}}(\mathbf{x}_0) = jk \sum_{\nu=-\infty}^{\infty} \check{S}_\nu^{(1)}(\omega) \{ H_{\nu-1}^{(1)}(kr_0) e^{j(\nu-1)\alpha_0} + H_{\nu+1}^{(2)}(kr_0) e^{j(\nu+1)\alpha_0} \} \quad (2.73)$$

焦点音源法では, 逆フーリエ変換を用いて低演算量な時間領域の実装を導くことができた. この手法も同様に時間領域の実装に変換することが可能である [68]. 周波数領域の駆動関数の式 (2.73) に現れるハンケル関数に式 (2.9) を適用した後, 逆フーリエ変換して次式の時間領域での実装を導くことができる.

$$u_{\text{ref}}(\mathbf{x}_0, t) = 2\sqrt{\frac{2}{\pi r_0^3}} y_0 \cdot \delta\left(t - \frac{r_0}{c}\right) * \check{h}_{\text{ref}}(t) * \check{s}^{(2)}(\alpha_0, t) \quad (2.74)$$

なお, $h_{\text{ref}}(t)$ は次式で定義される時間領域のフィルタである.

$$\check{h}_{\text{ref}}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \sqrt{\frac{\omega}{jc}} e^{j\omega t} d\omega \quad (2.75)$$

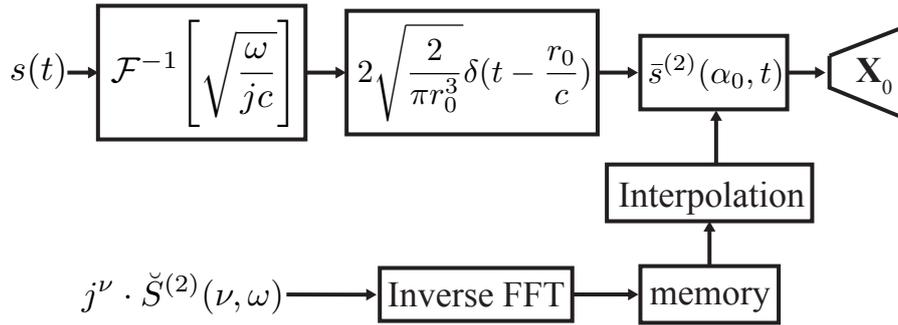


図 2.15: 時間領域実装のブロック図 [65].

ここで, $\bar{s}^{(2)}(\alpha_0, t)$ は指向性音源が作る音場の平面波展開係数を逆フーリエ変換により時間領域に変換した係数に相当し, 次式で表される.

$$\bar{s}^{(2)}(\alpha_0, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\sum_{\nu=-\infty}^{\infty} j^{\nu} \cdot \check{S}^{(2)}(\nu, \omega) \cdot e^{j\nu\alpha_0} \right] e^{j\omega t} d\omega \quad (2.76)$$

このフィルタは, 円調和級数 $\check{S}^{(2)}(\nu, \omega)$ の逆フーリエ変換を用いて効率的に実装できる [65].

一般に, フーリエ変換は等間隔のサンプリングを前提としているが, 図 2.14 からわかるとおり, 2 次音源のサンプリング点 α_0 は角度の間隔が等間隔にならない. そこで, 平面波に対し等間隔サンプリングを前提とした逆円調和変換を適用してフィルタを算出し, このフィルタを用いて補間処理を行うことで, 2 次音源の位置に一致する時間領域の駆動関数を求める.

ここで紹介した手法を時間領域で実装した際の処理フローを図 2.15 に示す. 指向特性の形状そのものを変えず, 単に音源の位置を平行移動したり, 音が強く放射される方向であるビームの向きを回転させる場合には, 円調和級数 $\check{S}_{\nu}^{(2)}$ を円状にシフトさせれば良いだけなので, 追加の演算量なく効率的に処理することができる.

2.4 音場再現と客観評価尺度

提案手法の性能を評価するため、本研究において以下の2種類の評価尺度を用いる。最初の評価尺度は、音場の再現精度を評価する誤差評価尺度（以下、音場の誤差評価尺度）である。

$$Err_{SF}(x, y) = 10 \log 10 \left(\frac{|s_{\text{org}}(x, y) - s(x, y)|^2}{|s_{\text{org}}(x, y)|^2} \right), \quad (2.77)$$

ここで、 $s_{\text{org}}(x, y)$ は式 (2.71) で計算できる位置 (x, y) の目標音場を表し、 $s(x, y)$ は再現場において位置 (x, y) で観測される音場を表す。音場の誤差評価尺度 $Err_{SF}(x, y)$ は、各実験ごとに定める領域において、一定間隔の格子状に配置した観測点で計算する。

第2の評価尺度は、再現される指向特性の精度を評価する誤差評価尺度（以下、指向特性の誤差評価尺度）である。アプリケーションによっては、再現される位相が目標となる指向性音源が生成する音場の位相とずれていても、音響信号が強く伝搬する方向と弱く伝搬する方向のパターンが再現されていれば十分な場合も多いため、音場の再現精度以外に、指向特性の再現精度についても評価を行う。

$$Err_{Dir}(x, y) = 10 \log 10 \left(\frac{||s_{\text{org}}(x, y)| - |s(x, y)||^2}{|s_{\text{org}}(x, y)|^2} \right), \quad (2.78)$$

ここで $s_{\text{org}}(x, y)$ は、 (x_c, y_c) を音源の中心とする、単位円上の制御点 $(x, y) = (x_c + \cos \phi, y_c + \sin \phi)$ で観測される音圧を示す。評価においては、実験毎に観測点を配置する単位円の半径と、隣あう観測点間の間隔を決定して評価尺度を計算する。

2.5 音源が持つ指向特性

人間が発話するとき、頭部や体の形状の違いによる直接音の放射方向や反射音の発生により、音が強く伝搬する方向と弱く伝搬する方向（指向特性）が生じる。本節では、人物の発話や楽器の持つ指向特性について述べる。また、円調和級数展開などの手法を用いて指向性を合成する際に、円調和展開級数を有限次元で打ち切ることで音場再現において生じる誤差について議論する。

2.5.1 人物発話や楽器における水平面上の指向特性

人物発話や楽器により生じる指向特性について、音響特性のモデリング手法の評価や建築物の音響設計などを目的として多くの検討がなされてきた [97, 98, 99, 100, 101]。録音技術の制約や、ターゲットとなるアプリケーションで利用する音声帯域幅から、多くは 4 kHz までに限定して指向特性の測定が行われてきた。しかし、近年、4 kHz より高

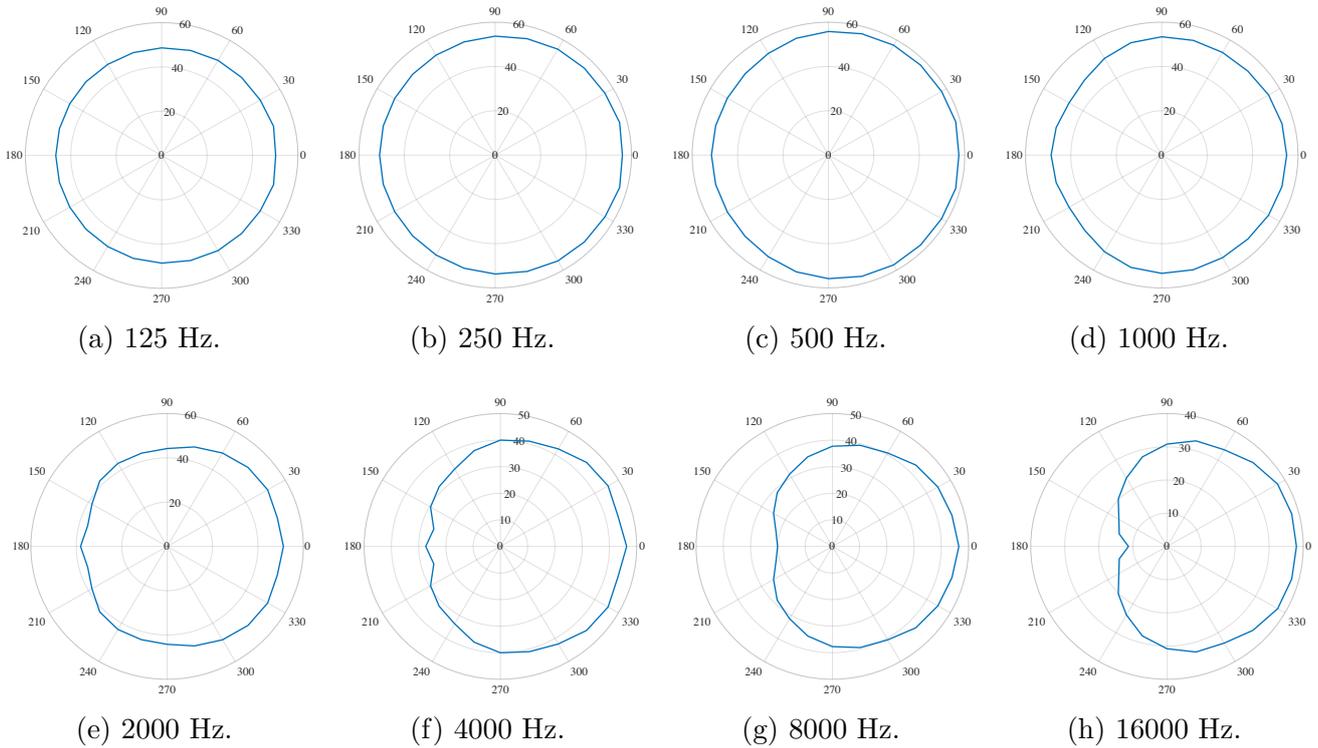


図 2.16: 周波数毎の指向性の変化 [104]. 表示の単位は dB(SPL).

域の周波数成分が人間の音声知覚に重要な影響を及ぼすことが指摘されていることから [102, 103], 4 kHz より高域を含む広帯域音声の指向特性についても検討が進められている [104]. Monson らは, 15 名の被験者の録音音声に対するスペクトル分析により, 発話の指向性を調査した. 15 名の被験者は 2 年以上の歌声発声トレーニングを受けた経験を持つ. 録音においては, 被験者を中心とする半径 0.6 m の半円上に 15° 間隔でマイクを設置して同時に録音を行った. 発話の指向特性として左右対象を仮定すると, 周波数毎の指向特性は図 2.16 のようになる.

スピーカで音を再生する際, 高周波の音ほど指向性が鋭くなるが音声についても同様の傾向が確認できる. Monson らは, さらに男性と女性での指向性の差, および通常の発話と歌声の指向性の差について調査したが, 有意な差は確認されなかった. 一方, 音圧の大きさや発話する音素の違いによって指向性に有意な差が現れることを確認している [104].

Otondo らは, 45° 間隔に配置した 8 チャンネルのマイクを用いてクラリネットの音を同時録音することによって, 楽器の指向特性を調査している [101]. その結果, 図 2.17 に示すとおり, 同じオクターブバンドの音であっても指向特性が異なることが示されている. 図より明らかなおとおり, 楽器の指向特性は, 低域であっても, 音声発話より複

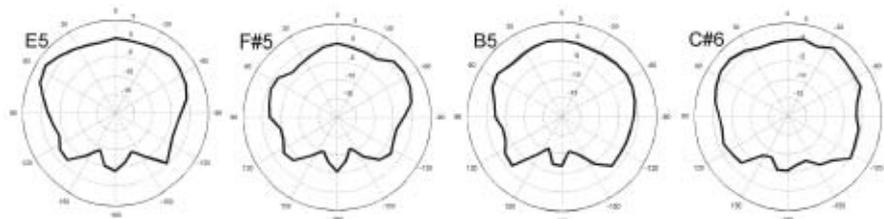


図 2.17: 楽器 (クラリネット) の指向特性. 判例は音階を表し周波数と以下のとおり対応. 'E5': 659.2 Hz. 'F#5': 740 Hz. 'B5': 987.8 Hz. 'C#6': 1108.7 Hz. 指向特性のデータには Otondo らの測定データを用いた [101].

雑な指向特性を有することがわかる.

2.5.2 有限次元の円調和級数展開による打ち切り誤差

音声発話など, 指向性を持った音源から生成された音響信号をスピーカアレイで再現するには, 複数のマイクで同時に録音を行った上で, スピーカアレイに適切な駆動信号を与えて音響信号を再生する. 2.2.1 項で述べたとおり, 音源の指向特性が十分高い次数の円調和展開級数でモデル化されているとき, 円状に並べた各スピーカに式 (2.45) で算出される駆動関数を与えることで音源が持つ指向性を再現できる. しかし, 録音の際に使うことができるマイクの数にも限界があるため, 指向性再現のために利用できる円調和展開級数も利用可能な次数までで打ち切ることとなる. 本項では, 音声発話の指向性を再現するのに用いる円調和展開級数を有限で打ち切った際に生じる誤差について評価する. なお, 録音データとして, Monson らの評価データを用いる [104].

2.5.1 項で述べたとおり, 15° 間隔においた 13 本のマイクを用いて音声発話を録音し, 左右対象な指向性を仮定して円上に観測される音圧レベルをプロットすると, 周波数毎に図 2.16 のような指向性が得られる. これらの測定データを用いて円調和展開係数を以下のとおり算出する.

$$\check{S}_\nu^{(2)}(\omega) = \tilde{S}_\nu(\omega)/H_\nu^{(2)}(kr) \quad (2.79)$$

ここで, ν 次のフーリエ係数 $\tilde{S}_\nu(\omega)$ は次式で算出できる.

$$\tilde{S}_\nu(\omega) = \int_{-\infty}^{\infty} S(\mathbf{x}) e^{-j\nu\alpha} d\alpha \quad (2.80)$$

マイクアレイの半径 ($r = 0.6\text{m}$), マイク数 ($M = 24$, 半円での測定において 12 本のマイクを利用しているため全円においては 24 本となる), および当該マイク数で表

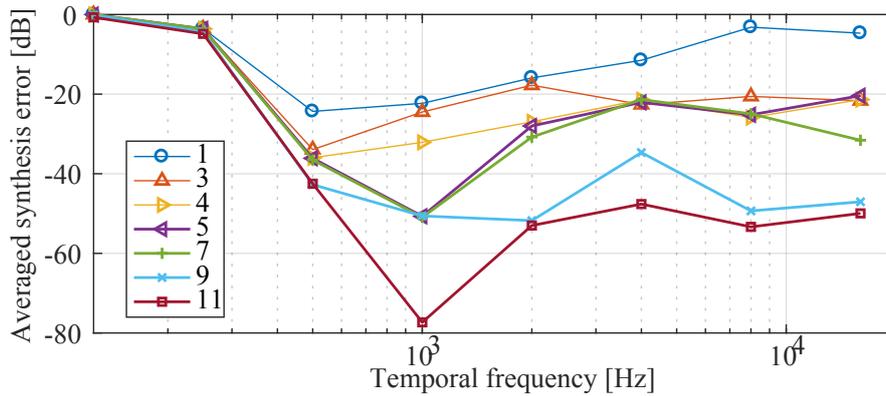


図 2.18: 円調和展開級数の最大次数と周波数毎の指向特性の再現精度の関係. 凡例の数値は打ち切りの最大次数を表す.

現可能な最大波数 k_{max} から, 利用可能な円調和展開級数の次数 \mathcal{N} は次の関係式で算出できる [105].

$$\mathcal{N} \approx \frac{2\pi f_{max} r}{c} \quad (2.81)$$

ここで, $f_{max} = cM/4\pi r$ である. 以上を用いて, $\mathcal{N} \approx 12$ となる.

最大次数 $\mathcal{N} \approx 12$ のとき, 式 (2.71) で計算される音場を再現対象の音場 $s_{org}(x, y)$ とし, 打ち切る円調和級数展開の次数を変化させながら再現される音場 $s(x, y)$ を式 (2.71) を用いて算出した. なお, 音源を中心とする半径 1 m の円上に 2° 間隔で設定した制御点において式 (2.71) の音場を計算し, $\mathcal{N} \approx 12$ の時の音場との誤差を式 (2.78) を用いて算出した. 算出した再現精度の誤差の平均値を図 2.18 に示す. なお, 表示を簡潔にするために 5 次以上の高次については奇数次のみを表示した. 図 2.18 より, いずれの周波数においても打ち切り次数が上がるに連れて再現精度も向上することがわかる. 周波数毎に再現精度が異なるのは, 図 2.16 に示すとおり, 指向性が周波数毎に異なるためであると考えられる.

さらに, 周波数毎に算出した再現誤差の平均値について, 利用する最大次数毎に計算した平均値を図 2.19 に示す. 結果のグラフより, 次数が増加するにつれて概ね線形に再現精度が改善することがわかる. 最大次数が 4 のとき再現精度の平均値は -20.90dB となる. したがって, 円調和級数展開の最大次数を 4 とすれば振幅レベルでの平均誤差を約 1% 以下に抑えられることがわかる.

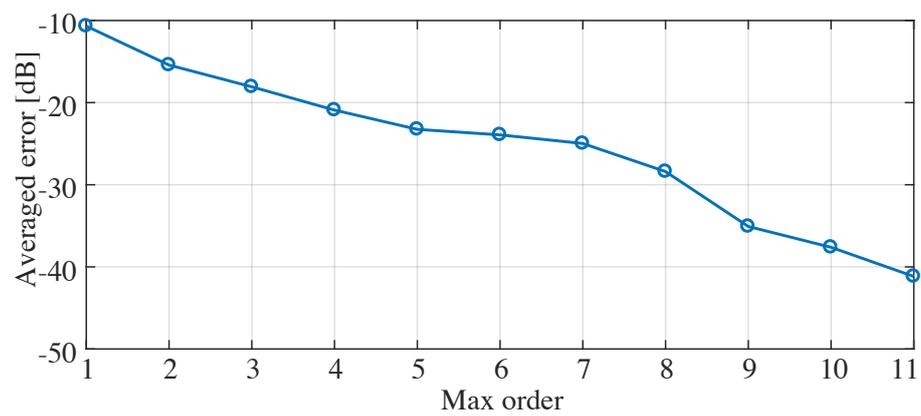


図 2.19: 円調和展開級数の最大次数と指向特性の再現精度の平均値の関係を, 次数毎にプロットした結果.

第3章 複数の焦点音源で生成した 仮想円状スピーカアレイによる 指向性再生

本章では指向性を有する音源が生成する音場を、スピーカよりも前方の受聴者近くに再現する手法について検討する。2.2節で述べたとおり、指向性音源が作る音場や指向特性は円状スピーカアレイでも再現することができる。そこで、本章ではまず、直線状に密に並べた一連のスピーカからなる直線状スピーカアレイから、直線状スピーカアレイの前方に複数の焦点音源を生成し、これらに適切な重みを与えることで仮想の円状スピーカアレイを形成する。次に直線状スピーカアレイ前方に形成した仮想の円状スピーカアレイを用いて指向性を合成することで、スピーカよりも前方に所望の指向性を持つ指向性仮想音源を生成する。提案手法では、各焦点音源の位置を制御することにより、指向性を持つ仮想音源の位置や回転を実現できる。

なお、指向性を有する音源が生成する音場を再現する手法には、2.3節で述べた従来技術があり、スピーカより前面に仮想音源が飛び出すよう拡張した手法も提案されている [68]。本章では、これら従来手法に対する提案手法の優位性についても論じる。

さて、直線状スピーカアレイを用いて仮想の円状スピーカアレイを生成して制御するためには、2つの駆動関数を直列に接続した駆動関数を求めて実スピーカを駆動すれば良い。2つの駆動関数のうち、1つめの駆動関数は仮想の円状スピーカアレイを制御して指向特性を再現するための駆動関数である。この駆動関数は、2.2.1項にて述べたように、事前に所望の指向性音源が作る音場から算出した円調和展開係数を解析的に変換することにより求められる。2つめの駆動関数は複数の焦点音源を生成して仮想の円状スピーカアレイを生成するための駆動関数である。この駆動関数として式(2.7)を用いて、指向性音源を生成したい位置を中心に円状に必要な仮想スピーカの数だけ焦点音源を生成すればよい。なお、式(2.11)で定義した時間領域の駆動関数を用いれば、式(2.12)で定義された波面合成プレフィルタを適用した後の音響信号に対して、焦点音源の位置に応じた遅延処理とゲイン乗算処理を実施するだけで所望の位置に焦点音源を生成することができる。指向特性そのものが変わらず音源の平行移動や回転のみを表現したい場合に必要な追加の演算は、遅延量やゲインの再計算のよう

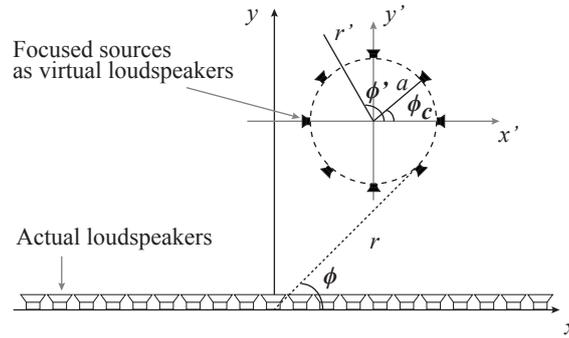


図 3.1: 実スピーカアレイと仮想の円状アレイの位置関係, および絶対座標系と相対座標系.

に無視できるくらい小さなものだけですむため, 非常に効率よく処理を行うことができる.

3.1 円状スピーカアレイを用いた指向特性の制御

2.2.1 項でも述べたとおり, 円状スピーカアレイを用いて指向性制御をする手法には様々なものが提案されてきた. Pressure-matching method に基づくものは容易に定式化できるため幅広く用いられているものの [84], 小さな半径の円状スピーカアレイを用いる場合に駆動関数の低周波数領域で生じる過剰なゲインが深刻な音質劣化をもたらすことが指摘されてきた [52].

2.2.1 項で述べた mode-matching method を含む円調和級数展開に基づくものも提案されている [53] [106] [107]. 2.2.1 項に述べたとおり低周波数領域では円調和級数の打ち切り次数に注意する必要があるが, pressure-matching method に基づく手法より駆動関数が持つゲインを柔軟に制御可能となるため, 本提案手法では円調和級数展開に基づく手法を用いて指向特性の制御を行う. 図 3.1 に実スピーカからなる直線状スピーカアレイと, 焦点音源で生成した仮想の円状スピーカアレイの位置関係を示す. 仮想の円状スピーカアレイは, 当該円状スピーカアレイの中心を原点とする相対座標 (x' - y' 平面) で位置を指定できることに注意する.

仮想の円状スピーカアレイが位置 (r', ϕ') に作る音圧は, 図 3.1 の相対座標系を用いて次式で与えられる.

$$p(r', \phi') = \int_0^{2\pi} G(r', \phi' | a, \phi_c) D(a, \phi_c) d\phi_c \quad (3.1)$$

ここで, $G(r', \phi' | a, \phi_c)$ は, 位置 $\mathbf{x}_c = (a, \phi_c)$ の音源による伝達関数を表し, $D(a, \phi_c)$ は

位置 $\mathbf{x}_c = (a, \phi_c)$ の仮想スピーカに与える駆動関数を表す. ここで, 表記の簡単のため角周波数 ω を省略した. これら伝達関数と駆動関数を円調和級数展開すると, 上式の音場は次のとおり変形できる.

$$p(r', \phi') = \int_0^{2\pi} \sum_{\nu=-\infty}^{\infty} G_{\nu}^{(2D)} e^{j\nu(\phi' - \phi_c)} \cdot \sum_{\nu'=-\infty}^{\infty} D_{\nu'} e^{j\nu'\phi_c} d\phi_c \quad (3.2)$$

ここで, ν と ν' は円調和級数の次数を表す. なお, 円調和展開においては, 次式の直交性が成り立つ.

$$\frac{1}{2\pi} \int_0^{2\pi} e^{-j\nu\phi_c} \cdot e^{j\nu'\phi_c} d\phi_c = \begin{cases} 1 & \nu = \nu' \\ 0 & \nu \neq \nu' \end{cases} \quad (3.3)$$

この関係式を式 (3.2) に代入して次式のように変形する.

$$p(r', \phi') = 2\pi \sum_{\nu=-\infty}^{\infty} G_{\nu}^{(2D)} D_{\nu} e^{j\nu\phi'}. \quad (3.4)$$

原点に配置した任意の指向性音源が作る音場は, 円調和級数展開を用いて次式で定義できる.

$$p(r', \phi') = \sum_{\nu=-\infty}^{\infty} C_{\nu} H_{\nu}^{(2)}(kr') e^{j\nu\phi'}. \quad (3.5)$$

ここで, C_{ν} は ν 次の円調和級数を表し, $H_{\nu}^{(2)}$ は ν 次の第2種ハンケル関数を表す. これら2式の係数比較より, 駆動関数を円調和展開して得られる展開係数 D_{ν} は次式で与えられる.

$$D_{\nu}(\omega) = \frac{C_{\nu} H_{\nu}^{(2)}(kr')}{2\pi G_{\nu}^{(2D)}} \quad (3.6)$$

ここで, 伝達関数の円調和級数展開は次式で定義できる.

$$G(r', \phi' | a, \phi_c) = \frac{j}{4} \sum_{\nu=-\infty}^{\infty} J_{\nu}(ka) H_{\nu}^{(2)}(kr') e^{j\nu(\phi' - \phi_c)}, \quad (3.7)$$

ここで, $J_{\nu}(ka)$ は ν 次のベッセル関数を表す. これら2式より駆動関数を円調和級数展開した展開係数は次式で与えられる.

$$D_{\nu}(\omega) = \frac{2}{j\pi} \cdot \frac{C_{\nu}}{J_{\nu}(ka)}. \quad (3.8)$$

この円調和級数を用いて, 仮想の円状スピーカアレイから所望の音源の指向特性を再現することができる.

3.2 直線状スピーカアレイを用いた仮想円状スピーカアレイの生成

直線状アレイで生成した焦点音源を仮想のスピーカとして用いる検討が進められている [108] [109] [110]. Spors らは, 実スピーカアレイで生成した多数の焦点音源を, 大きさを持つ実スピーカでは困難なほど密な間隔で並べて円状スピーカアレイを構成することで, 実スピーカを用いて作った円状スピーカアレイよりもはるかに高い空間エイリアシング周波数が実現できることを示した. このように大きさを持たない焦点音源が仮想のスピーカとして使える点を利用し, 本章では, 焦点音源で作った仮想の円状スピーカアレイを制御することで指向性音源が作る音場を再現する手法について検討する. 直線状アレイを用いて式 (2.7) の焦点音源を生成する駆動関数と, 3.1 節の円状アレイを制御して所望の指向特性を再現する式 (3.8) の展開係数を組み合わせることで, 直線状アレイの前面に所望の指向特性を持つ仮想的な音源を作ることができる.

\mathcal{L} 個の焦点を位置 $\mathbf{x}_l = (a, \frac{2\pi l}{\mathcal{L}})$ に生成する際 (ただし $0 \leq l < \mathcal{L}$), η 番めの 2 次音源に与える駆動関数 $D_{CA}(\mathbf{x}_\eta, \mathbf{x}_l, \omega)$ は式 (2.7) に式 (3.8) を乗じて次式で表される.

$$\begin{aligned}
 D_{CA}(\mathbf{x}_\eta, \mathbf{x}_l, \omega) &= \sum_{l=0}^{\mathcal{L}-1} D_{2D}(\mathbf{x}_\eta, \mathbf{x}_l, \omega) \cdot \left(\sum_{\nu=-N}^N D_\nu(\mathbf{x}_\eta) \cdot e^{j\nu \frac{2\pi l}{\mathcal{L}}} \right) \\
 &= \sum_{l=0}^{\mathcal{L}-1} -\frac{(jk)^{3/2}}{2} \frac{y_\eta - y_l}{|\mathbf{x}_\eta - \mathbf{x}_l|} H_1^{(1)}(k|\mathbf{x}_\eta - \mathbf{x}_l|) \cdot \left(\sum_{\nu=-N}^N D_\nu(\mathbf{x}_\eta) \cdot e^{j\nu \frac{2\pi l}{\mathcal{L}}} \right) \\
 &= \sum_{l=0}^{\mathcal{L}-1} -\frac{(jk)^{3/2}}{2} \frac{y_\eta - y_l}{|\mathbf{x}_\eta - \mathbf{x}_l|} H_1^{(1)}(k|\mathbf{x}_\eta - \mathbf{x}_l|) \cdot \left(\sum_{\nu=-N}^N \frac{2}{j\pi} \cdot \frac{C_\nu}{J_\nu(ka)} \cdot e^{j\nu \frac{2\pi l}{\mathcal{L}}} \right)
 \end{aligned} \tag{3.9}$$

ここで, N は円調和級数展開の打ち切り次数である.

3.3 時間領域での実装

2.1.4 項で述べたとおり, 焦点音源法を時間領域で実装すると演算量を削減することができる. 本手法も円状アレイを構成する個々の焦点音源を時間領域での実装により生成することで演算量の低減が可能となる. 本節では, 提案手法の時間領域実装について述べる.

式 (3.9) で定義した周波数領域の駆動関数を, 時間周波数について逆フーリエ変換

することで、次式で定義される時間領域の駆動関数を得る.

$$\begin{aligned} u_{\text{CA}}(\mathbf{x}_\eta, t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} D_{\text{CA}}(\mathbf{x}_\eta, \omega) e^{j\omega t} d\omega \\ &= \sum_{l=0}^{\mathcal{L}-1} w(\mathbf{x}_l, t) * u(\mathbf{x}_\eta, \mathbf{x}_l, t) \end{aligned} \quad (3.10)$$

ここで、 $*$ は畳み込み演算を表し、 $u(\mathbf{x}_\eta, \mathbf{x}_l, t)$ は式(2.11)で定義される時間領域で実装された焦点音源法の駆動関数を表し、 $w(\mathbf{x}_l, t)$ は仮想の円状スピーカアレイに含まれる l 番めの焦点音源に与える重み係数を表す. この重み係数 $w(\mathbf{x}_l, t)$ は次式で表される.

$$w(\mathbf{x}_l, t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\sum_{\nu=-N}^N \frac{2}{j\pi} \cdot \frac{C_\nu}{J_\nu(ka)} \cdot e^{j\nu \frac{2\pi l}{\mathcal{L}}} \right] e^{j\omega t} d\omega. \quad (3.11)$$

式(3.10), 式(2.11), および式(3.11)より, 入力信号に対して式(2.12)の波面合成プレフィルタを適用して出力した信号に, 式(3.11)で提案した重み係数に対応するフィルタを畳み込んだ上で, 式(2.11)のうち波面合成プレフィルタを除いたゲイン乗算・遅延処理を適用する. これを焦点音源の数 \mathcal{L} だけ繰り返して足し合わせることで位置 \mathbf{x}_η に位置する実スピーカに与える駆動関数を求めることができる.

3.4 時間領域での実装における音場再現の精度改善

焦点音源法は, 周波数領域の駆動関数では個々の実スピーカに対するフィルタ演算が必要になる. これに対し, 時間領域の実装では実スピーカごとに異なるゲイン乗算と遅延処理で代用できるため, 演算量を劇的に削減できる. しかし, 周波数領域で定義されたフィルタでは可能であった任意の精度での非整数サンプルの遅延処理が, 時間領域での実装ではデジタル信号のサンプリング周期に限られるため, 焦点音源が作る音場の精度が低下する. そこで本節では, 2.1.5項で述べた非整数遅延フィルタを時間領域実装に組み込むことで, 再現音場の精度改善を試みた検討について述べる.

3.4.1 Thiran フィルタを用いた時間領域実装による精度改善

一般にスピーカアレイを構成するスピーカの間隔で決まるエイリアシング周波数より上の周波数帯域では, 空間エイリアシングが発生するため, 音場を正確に再現できない. さらにフィルタ長が短い Lagrange 補間では, 低域通過特性により高音域のパワーが抑圧されるため音質が劣化する. 全域通過型である Thiran フィルタは全帯域に渡りフラットな周波数特性を持ち, 実スピーカアレイで一般的なスピーカ配置間隔におけるエイリアシング周波数以下ではフラットな群遅延特性を有する. そこで, Thiran

フィルタで定義される非整数遅延フィルタを時間領域に適用して音場の再現精度改善を検討する. 新たな時間領域の駆動関数は, 式 (3.10) の駆動関数の後段に, Thiran フィルタを従属接続することで次式に示すとおり得られる.

$$u_{\text{CA}}(\mathbf{x}_\eta, t) = \sum_{l=0}^{\mathcal{L}-1} w(\mathbf{x}_l, t) * u(\mathbf{x}_\eta, \mathbf{x}_l, t) * p_{\text{Th}}(t, \tau_{\text{frac}}) \quad (3.12)$$

ここで, $n = N - m$ とした. また, 非整数サンプル遅延 τ_{frac} は, 次式で算出できる.

$$\tau_{\text{frac}} = \frac{|\mathbf{x}_\zeta - \mathbf{x}_{\mu,\nu}^{m,n}|}{c} - \left\lfloor \frac{|\mathbf{x}_\zeta - \mathbf{x}_{\mu,\nu}^{m,n}|}{c} \right\rfloor \quad (3.13)$$

ここで, $\lfloor x \rfloor$ は, 小数 x の整数部分を取り出す操作に相当する.

3.5 性能評価実験

まず最初に, 異なる重みを与えた複数の焦点音源が仮想の円状アレイを形成することを確認する. また, 非整数遅延フィルタを組み合わせた時間領域での手法について, 再現される音場, および指向特性の精度を確認する. 最後に, 再現される音場や指向特性の精度と演算量の観点から, 提案手法と 2.3 節に述べた従来手法の性能を評価する. なお, 時間領域での実装の方が周波数領域での実装よりも演算量が少ないのは明らかなため, 演算量の評価には 2.1.5 項で述べた Thiran フィルタにより音場再現精度を改善した時間領域実装のみを用いた. また, Thiran フィルタを用いない時間領域での実装については, 指向特性の再現精度が著しく劣化することが明らかなため, 評価には加えなかった.

3.5.1 実験条件

本章で実施するシミュレーションの実験条件を本項にまとめる. 例えば, 仮想の円状アレイの半径について性能比較する場合に, 仮想の円状アレイの半径について記載するように, 各実験においては本項で述べる条件からの差分のみを陽に記載する.

シミュレーションは MATLAB により実装した. 実験の構成を図 3.2 に示す. 音速は, 343.36 m/s とした. また, x 軸上の $-1.6 \leq x \leq 1.6$ の範囲に 0.05 m 間隔で 65 個のスピーカを直線状に並べてスピーカアレイを構成した. 焦点音源で構成する仮想の円状アレイの半径は 0.01 m とした. 目標となる音場は 2.5.2 項で述べたとおり, 音声発話の再現誤差の平均値が約 1 % 以下となる最大次数 $N = 4$ の円調和級数 $\check{S}_\nu^{(2)}(\omega)$ でモデル化することとし, $2N + 1 = 9$ 個の乱数で生成した. この乱数は, 振幅 1.0 以下の複素数で生成した. なお, 目標音場は式 (2.71) で算出した. また, 仮想の円状アレイ

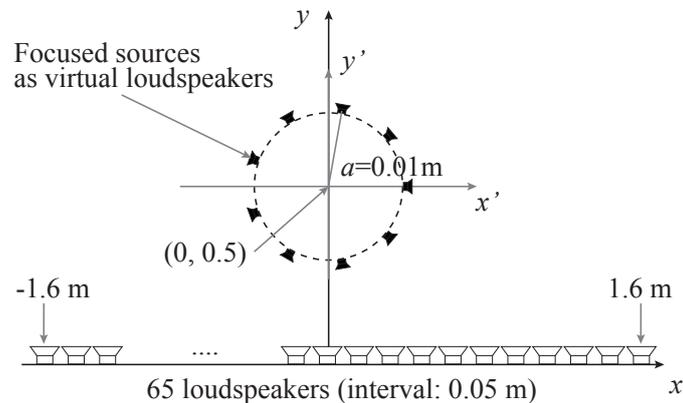


図 3.2: 実験のセットアップ.

の中心は, $(x_c, y_c) = (0.0, 0.5)$ とした. 仮想の円状アレイに含まれる焦点音源の数は 9 個とし, 円周上に等間隔となるように生成した. なお, これら焦点音源の座標は円状アレイの中心座標と図 3.1 に記載の相対座標から容易に計算できる.

3.5.2 客観評価尺度

提案手法の性能を評価するため, 2.4 節で述べた, 式 (2.77) で表される音場の誤差評価尺度を用いた. 音場の誤差評価尺度 $Err_{SF}(x, y)$ は, $-1.5 \text{ m} \leq x \leq 1.5 \text{ m}$, $1.0 \text{ m} \leq y \leq 3.0 \text{ m}$ の領域において, 0.01 m 間隔の格子状に配置した観測点で計算した.

3.5.3 円状アレイの半径が再現精度に与える影響

円状アレイを構成する焦点音源同士の間隔は, 理論上, 無限小の値を取りうる. しかし, 現実のスピーカにはそれぞれ大きさがあるため, 一定の間隔を開けてスピーカアレイを実装せざるを得ない. したがって, 焦点音源同士の間隔が音場の再現誤差に与える影響を把握しておくことは重要である. 本節では, 焦点音源同士の間隔を変化させて誤差評価尺度がどのように変化するかを明らかにする. 実験条件は 3.5.1 項に記載のとおりであるが, 本項では仮想円状アレイの半径のみ, 0.001 m, 0.01 m, 0.05 m, 0.1 m の 4 種類に変化させて音場の誤差評価尺度を評価した. 3.5.1 項に記載のとおり, $-1.5 \text{ m} \leq x \leq 1.5 \text{ m}$, $1.0 \text{ m} \leq y \leq 3.0 \text{ m}$ の領域において, 0.01 m 間隔の格子状に配置した観測点で誤差評価尺度を計算し, 100 Hz から 3500 Hz までの範囲でそれぞれ平均値を算出した.

音場の誤差評価尺度のグラフを図 3.3 に示す. アレイ半径を 0.001 m, 0.01 m, 0.05 m としたときには, 従来手法と同等の再現精度を達成しているが, アレイ半径 0.1 m

とすると 1800 Hz 以降, 高周波になるにしたがって指向特性の再現精度が低下した. アレイ半径 0.1 m のとき, アレイの直径は 0.2 m であるが, これを波長と考えると対応する周波数は約 1700 Hz ($343.36 \text{ m/s} \div 0.2 \text{ m} \approx 1700 \text{ Hz}$) であり, 再現精度が低下し始める周波数と概ね一致する. よって, アレイの直径が再現したい最大周波数に対応する波長以下となるようアレイ半径を選択すればよいことがわかる. 実スピーカで構成した直線状スピーカアレイのスピーカ間隔は 0.05 m であり, これは空間エイリアシング周波数およそ 3400 Hz に対応する. 3400 Hz をカバーするためにはアレイ半径を 0.05 m とすれば十分ではあるが, 本章ではアレイ半径を 0.01 m として以降の実験を行う.

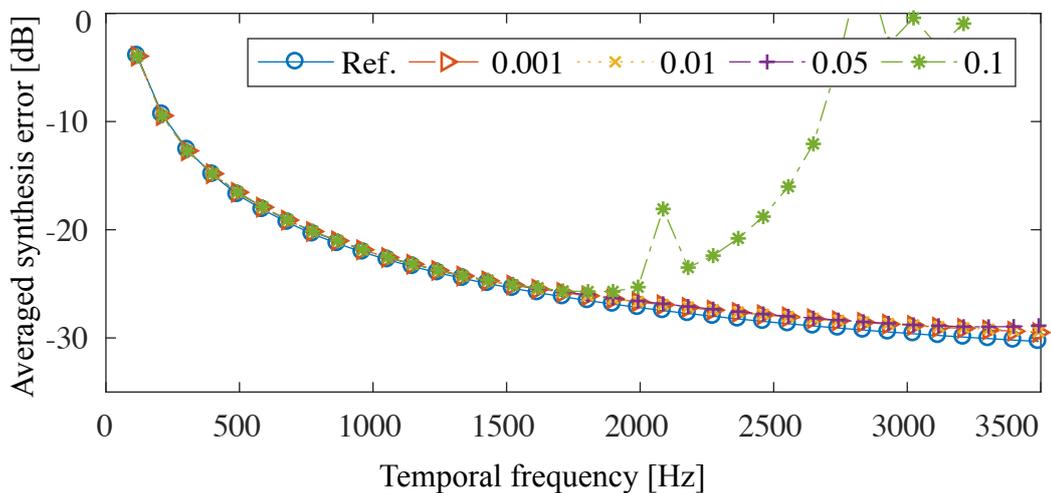


図 3.3: 音場に関する誤差評価尺度の平均値. 凡例については次のとおり. 'Ref.': 従来手法 (周波数領域での実装). 凡例の数値は, 対応する値を円状アレイの半径としたときの提案手法 (周波数領域での実装). 例) '0.01': 円状アレイ半径 0.01 m のときの提案手法 (周波数領域での実装).

3.5.4 音場の再現精度

音場の再現精度を評価するため単一周波数に対する音場の再現実験を行った。シミュレーションには、1992.2 Hz の正弦波を用いた。なお、この周波数は1024点の離散フーリエ変換を行う際に、2000 Hz にもっとも近い周波数であったため採用した。指向性音源が回転した際の再現精度を比較するため、指向性音源が反時計方向に 10° 回転した際の音場についても生成して比較を行った。音源に回転を加えなかった際の音場の再現結果を図3.4に示す。図3.4 (a) は目標となる指向性音源が作る音場を表し、(b), (c) は周波数領域で実装した提案手法が作る音場と、目標音場と比較したときの音場の誤差評価尺度を示す。特に (c) からわかるとおり、焦点音源の受聴エリア内で高精度に音場を再現できていることがわかる。また、図3.4 (d), (e) は、時間領域で実装した提案手法が生成する音場と、目標音場からの音場の誤差評価尺度を表す。縞の濃淡が誤差の dB 表示を表し、黒いほど誤差が小さい。なお、Thiran フィルタにより非整数サンプル遅延の補間を行っている。精度は、周波数領域での実装に及ばないが、指向性音源の音場が再現できていることがわかる。図3.4 (f), (g) は、周波数領域で実装した従来手法が生成する音場と、目標音場からの音場の誤差評価尺度を表す。特に (g) から明らかなおおり、提案手法と同様、高精度に音場を再現可能なことがわかる。次に、音源を反時計方向に 10° 回転した際の再現結果を図3.5に示す。図より明らかのように、音源を反時計方向に 10° 回転した場合も同様の傾向が確認できる。

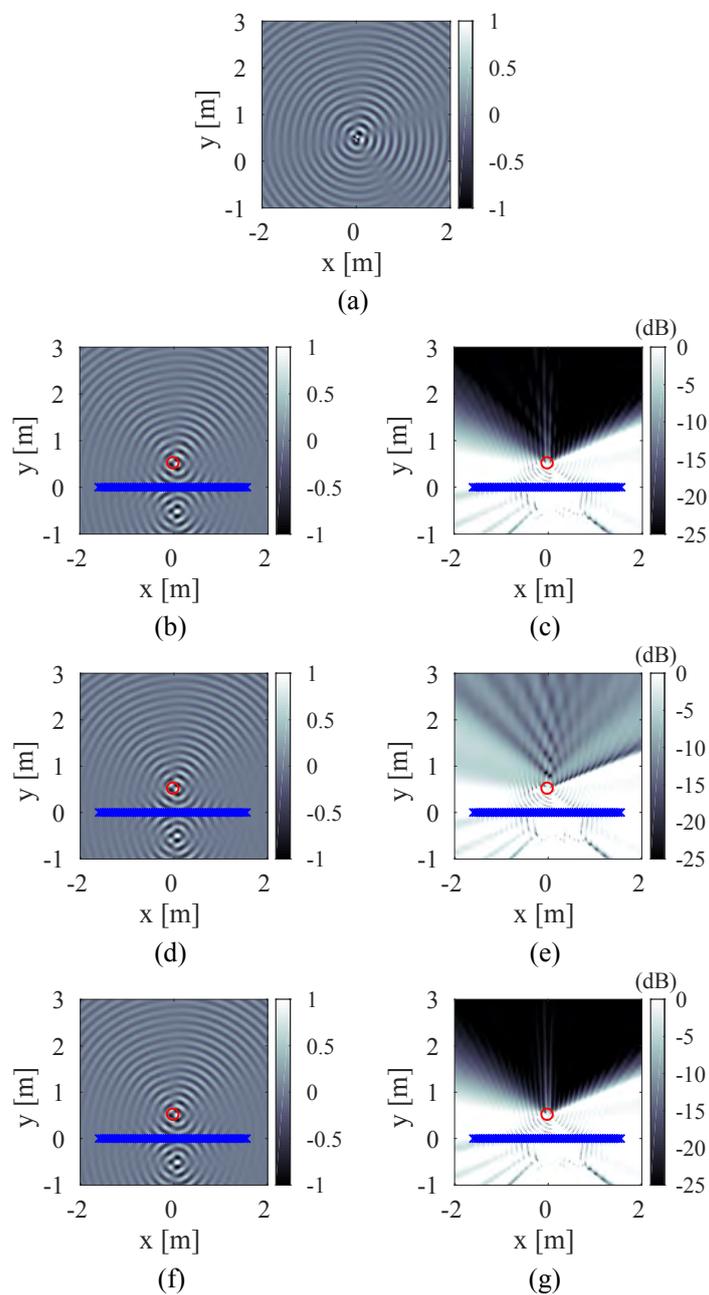


図 3.4: 音場の誤差評価尺度. (a): 目標音場. (b): 周波数領域で実装した提案手法の再現音場. (c): (b) の音場の誤差評価尺度. (d): Thiran フィルタで再現精度を改善した提案手法の時間領域実装による再現音場. (e): (d) の音場の誤差評価尺度. (f): 周波数領域で実装した従来手法の再現音場. (g): (f) の音場の誤差評価尺度.

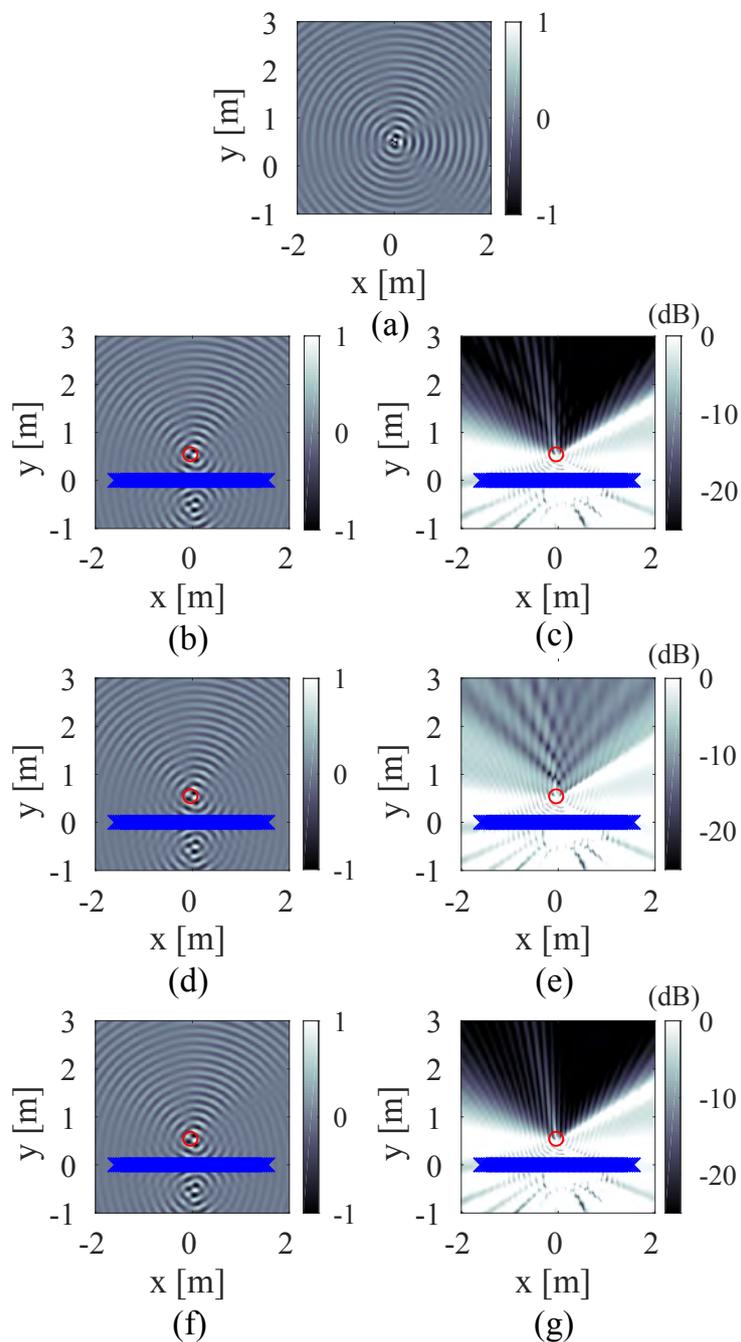


図 3.5: 指向性音源を反時計回りに回転 ($\phi = 10^\circ$) したときに再現された音場と、音場の誤差評価尺度. (a): 目標音場. (b): 周波数領域で実装した提案手法の再現音場. (c): (b) の音場の誤差評価尺度. (d): Thiran フィルタで再現精度を改善した提案手法の時間領域実装による再現音場. (e): (d) の音場の誤差評価尺度. (f): 周波数領域で実装した従来手法の再現音場. (g): (f) の音場の誤差評価尺度.

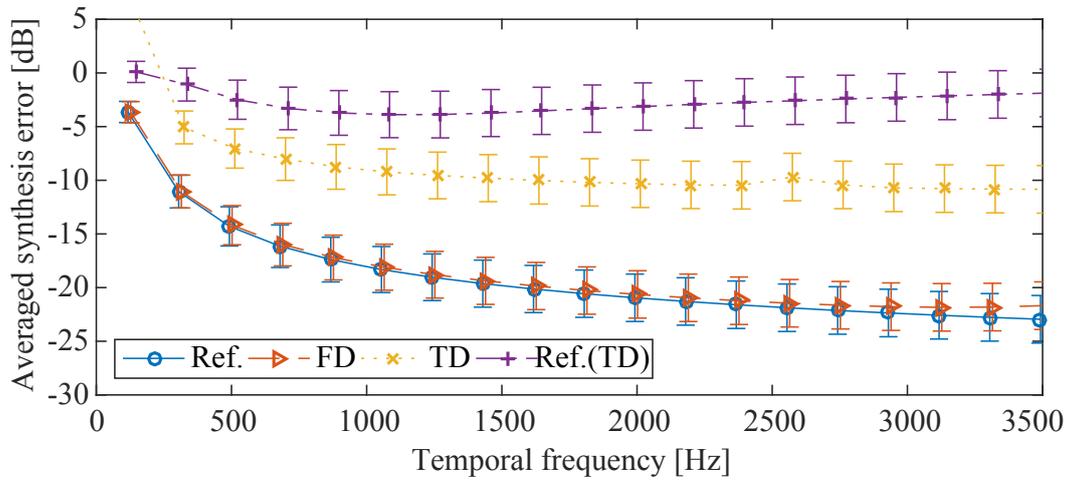


図 3.6: 音場の評価尺度について周波数ごとに算出した平均値. 'Ref.': 従来手法の周波数領域実装 [65, 68]. 'FD': 提案手法 (周波数領域での実装). 'TD': Thiran フィルタで補間した提案手法の時間領域実装. 'Ref.(TD)': 従来手法の時間領域実装 [65, 68].

次に、音場に関する誤差評価尺度の平均値が周波数に応じてどのように変化するかを確認する. $-1.5 \text{ m} \leq x \leq 1.5 \text{ m}$, $1.0 \text{ m} \leq y \leq 3.0 \text{ m}$ の領域において 0.01 m 間隔の格子状に配置した観測点で算出した音場に関する誤差評価尺度の平均値を、 100 Hz から空間エイリアシング周波数である 3500 Hz まで計算した. なお、評価に用いる円調和展開級数による依存性を排するため、 $2N + 1 = 9$ 個の乱数を 100 パターン生成し、最大次数 $N = 4$ の円調和級数 $\check{S}_p^{(2)}(\omega)$ として用いた. 当該乱数は、振幅 1.0 以下の複素数で生成した. また、音源の平行移動が音場の再現精度におよぼす影響について確認するため、仮想音源の位置を $-0.5 \text{ m} \leq x \leq 0.5 \text{ m}$, $0.25 \text{ m} \leq y \leq 0.75 \text{ m}$ の範囲でランダムに選択して、上記 100 パターンの円調和展開級数のそれぞれに対応させた. さらに、音源の回転が音場の再現精度におよぼす影響について確認するため、仮想音源の向きを $-30^\circ \leq \phi \leq 30^\circ$ の範囲でランダムに選択して、上記 100 パターンの円調和展開級数のそれぞれに対応させた. 100 パターンの計算結果の平均値を 95% 信頼度区間とともに図 3.6 に示す. なお、 95% 信頼度区間の算出においては student-t 分布を仮定した.

結果のグラフから、周波数領域での実装では音場の再現精度に有意な差を確認できない. 一方、時間領域の実装では、提案手法の再現精度が従来手法の再現精度を有意に上回ることを確認した. また、時間領域の実装と周波数領域の実装を比較すると、時間領域の実装の再現精度は 10 dB 以上低下することが確認できる. 2.3 節に述べた手法

の時間領域実装については, 2.3節に示すように, 駆動関数の計算過程で用いる補間処理で生じる誤差が原因で, 音場の再現精度が低下したと考えられる.

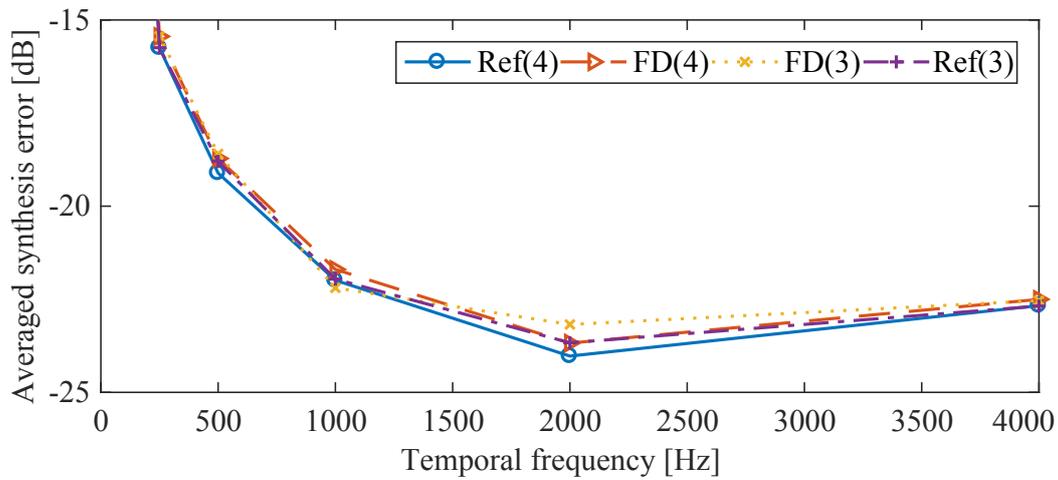


図 3.7: 周波数ごとの音場の誤差評価尺度. 凡例は以下のとおり. 'Ref(4)': 最大次数 $N = 4$ の従来手法 (周波数領域での実装). 'FD(4)': 最大次数 $N = 4$ の提案手法 (周波数領域での実装). 'FD(3)': 最大次数 $N = 3$ の提案手法 (周波数領域での実装). 'Ref(3)': 最大次数 $N = 3$ の従来手法 (周波数領域での実装).

次に, 2.5.2 項で用いた Monson らによる指向性のデータ [104] を用いて, 円調和展開級数の最大次数打ち切りがおよぼす誤差について評価する. 円調和展開級数の最大次数 $N = 3$ と $N = 4$ の場合について, 周波数領域実装の提案手法と従来手法の音場の誤差評価尺度を算出した. 結果を図 3.7 に示す. なお, 音場再現に用いる直線状スピーカアレイの間隔から, 3400 Hz 以上の音声帯域で空間エイリアジングが発生するため, 3400 Hz にもっとも近い周波数である 4000 Hz まで, 音場の再現精度を算出した.

音声発話の指向性を再現する場合, 最大次数 $N = 4$ の場合と $N = 3$ の場合を比較すると, 従来手法も提案手法も最大次数を $N = 3$ としても, 1 dB 程度しか再現精度が低下しないことが確認できた.

3.5.5 演算量の評価

演算量を評価するため、従来法と提案法を C 言語により実装し、ITU-T G.192 にて提供される浮動小数点版の演算量カウンタ [111] を各実装に組み込んだ。周波数領域での実装に比べて時間領域での実装に必要な演算量が非常に少ないことは明らかであるので、時間領域での実装のみを C 言語で実装した。

本シミュレーションでは、指向特性そのものは変化せず、その向きと位置のみが 20ms 毎に時間的に変化すると仮定した。シミュレーションにおけるスピーカの数 は 65 とし、波面合成プレフィルタの長さを 513 サンプルとした。円調和級数展開の最大次数は $N = 4$ とし、提案手法の仮想アレイに含まれる焦点音源の数は $2N + 1$ 個とした。提案法における時空間フィルタ（焦点音源ごとに畳み込むフィルタ）は、事前に計算してメモリに格納した。従来手法については、長さ 129 サンプルの 256 個の時空間フィルタを等間隔にサンプリングして算出してメモリに格納しておき、目的方向に対応する時空間フィルタは、メモリに格納した時空間フィルタの Lagrange 補間により算出した。

計測した演算量を表 3.1 に示す。無指向性の焦点音源を周波数領域実装で生成する際に必要な演算量を参考として記載する。提案手法の演算量は、円調和級数展開の最大次数が増加するにつれて増加したが、 $N = 4$ までであれば無指向性の焦点音源を周波数領域実装で生成するのに必要な演算量よりも大幅に少ないことがわかる。一方、従来手法の演算量は、円調和級数の次数が上がっても変化しなかった。これは、従来手法の時空間フィルタの長さが円調和級数の次数と独立であり、長さが変化しないためである。提案手法で必要な演算量は仮想の円状アレイを構成するための焦点音源の数に依存するため、円調和級数の最大次数が上がっても演算量に直ちに影響することはないが、円状アレイでの指向性再現には式 (2.46) で示した個数のスピーカが必要となるため、次数に比例して演算量を増加させる必要があることに注意が必要である。時間領域で実装した提案手法によれば、 $N = 6$ すなわち焦点音源の数が 13 個までであれば、従来手法よりも高い再現精度かつ低演算量での音場制御が可能であることがわかる。

表 3.1: 各手法について測定した演算量の推定結果 (weighted millions of operations per seconds. 以下, WMOPS). 'Reference': 時間領域で実装した従来手法 [65]. 'Proposed': 時間領域で実装した提案手法. 括弧の中の数字は円調和級数展開の最大次数を表す. 'Mono': 焦点音源法. (FD) は周波数領域での実装を表し, (TD) は時間領域の実装を表す [74].

Method	(WMOPS)
Reference in TD[68]	2089.37
Proposed in TD ($N = 4$)	1425.61
Proposed in TD ($N = 6$)	1965.28
Monopole in FD[48]	8130.42

3.6 本章のまとめ

本章では、直線状に密に並べた一連のスピーカからなる直線状スピーカアレイを用いて生成した焦点音源で仮想の円状スピーカアレイを生成し、この円状アレイの指向性制御を行うことにより、直線状スピーカアレイ前方に指向性音源を配置したときに生成される音場を生成する手法を提案した。直線状スピーカアレイを用いて仮想の円状スピーカアレイを生成して制御するためには、2つの駆動関数を直列に接続した駆動関数を求めて実スピーカを駆動すれば良い。2つの駆動関数のうち、1つめの駆動関数は仮想の円状スピーカアレイを制御して指向特性を再現するための駆動関数である。この駆動関数は、2.2.1項にて述べたように、事前に所望の指向性音源が作る音場から算出した円調和展開係数を解析的に変換することにより求められる。2つめの駆動関数は複数の焦点音源を生成して仮想の円状スピーカアレイを生成するための駆動関数である。この駆動関数として式(2.7)を用いて、指向性音源を生成したい位置を中心に円状に必要な仮想スピーカの数だけ焦点音源を生成すればよい。周波数領域で実装した場合、提案手法は従来法と同等の音場再現精度を達成できることを確認した。

さらに、ライブイベントなどリアルタイム性が要求されるようなアプリケーションに向けて時間領域での実装を導出し、従来手法よりも低演算量で高精度な音場再現ができることを確認した。式(2.11)で定義した時間領域の駆動関数を用いることで、式(2.12)で定義された波面合成プレフィルタを適用した後の音響信号に対して、焦点音源の位置に応じた遅延処理とゲイン乗算処理を実施するだけで所望の位置に焦点音源を生成することができる。指向特性そのものが変わらず音源の平行移動や回転のみを表現したい場合に必要な追加の演算は、遅延量やゲインの再計算のように無視できるくらい小さなものだけですむため、非常に効率よく処理を行うことができる。時間領域の実装について、音場の再現精度と演算量を評価し、時間領域で実装された従来法よりも低演算量かつ高精度な音場再現が可能であることを確認した。

3.6.1 本章の貢献と関連研究

本章の内容は、研究業績リスト [C-1] の内容をまとめたものである。この研究の貢献は、焦点音源で生成した仮想の円状スピーカアレイを制御して指向性音源が作る音場を再現し、特に時間領域で従来法より低演算量で高精度な音場再現ができることを示した点にある。

指向性を有する音源が生成する音場を再現する手法には、2.3節で述べた従来技術があり、スピーカより前面に仮想音源が飛び出すよう拡張した手法も既に提案されてい

た [68]. しかし, 指向特性を保ったまま音源が平行移動したり回転するような場合にも駆動信号の再計算が必要になるなど, 演算量の面で冗長な部分が残されていた. これに対して提案手法は, 指向性の再現と音源の移動を独立に制御可能な駆動関数を提案しており, 特に時間領域で実装した場合に演算量・再現精度の両面で優位性を実現している.

本章の提案手法は, 客席近くに迫る仮想音源に対して低演算量かつ高精度に指向特性を付与することができるため, ライブイベントなどのリアルタイムアプリケーションの臨場感を高めることに貢献すると言える.

第4章 複数の焦点音源から生成した 仮想多重極音源による指向性再生

前章に引き続き、本章でも指向性を有する音源が生成する音場を、スピーカよりも前方の受聴者近くに再現する手法を検討する。3章では直線状スピーカアレイにより複数の焦点音源を生成し、それらの位置や与える重み係数を制御することにより、仮想の円状スピーカアレイを生成して所望の指向性音源が作る音場を再現した。一方、2.2節でも述べたとおり、指向性音源が作る音場や指向特性は多重極音源でも再現することができる。本章では、直線状にスピーカが等間隔に並んだ直線状スピーカアレイを用いて、スピーカアレイの前方に複数の焦点音源を生成し、これらに適切な重みを与えることで仮想の多重極音源を形成する。直線状スピーカアレイ前方に形成した仮想の多重極音源を用いて指向性を合成することで、スピーカよりも前方に所望の指向性を持つ指向性仮想音源を生成する。図 4.1 に提案手法の概念図を示す。

当初のモチベーションは、例えばダイポール音源を再現する際に、仮想円状アレイでは例えば 8 個の焦点音源が必要なのに対し、仮想多重極音源では 2 個の焦点音源で十分であるという点がある。ただし、複雑な指向特性を再現するためには高次の多重極音源を用いる必要があり、一般に、高次の多重極は多数のモノポール音源により生成されるため、多重極音源の重ね合わせを用いた指向性制御には多くの演算量が必要となる。そこで、隣り合う多重極を重ね合わせることで得られる効率的な実現方法も導く。さらに演算量を削減するため、時間領域での実装についても述べる。

4.1 焦点音源からなる多重極音源の生成

本節では、スピーカアレイによって生成される焦点音源をモノポール音源と仮定し、複数の焦点音源を組み合わせることで、多重極音源が生成する音場を再現する。まずは 1 次元の場合を考える。 m 次の多重極は、2 個の $(m - 1)$ 次の多重極を微小な間隔で並べた上で、片方の多重極の極性を反転させることで生成できる。このような単純な構造で多重極を実装すると、図 4.2 に示すように m 次の多重極を作るために 2^m 個のモノポール音源が必要となる。したがって、再現したい多重極の次数が上がるにつれて、必要なモノポール音源の数は指数関数的に増加する。このような課題を解決するために、多重極音源の効率的な実装方法が提案されている [58]。図 4.3 に示すよう

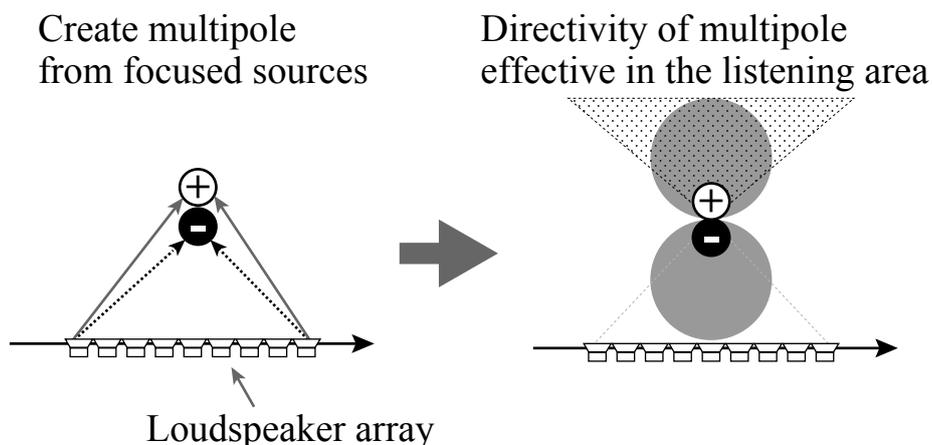


図 4.1: 複数の焦点音源で生成した多重極音源の概念図.

に, m 次多重極は, 隣り合う $m-1$ 次多重極を重ね合わせることで $m+1$ 個のモノポール音源だけで作ることができる. したがって, m 次多重極に含まれる μ 番目のモノポール音源の座標 x_μ^m は次式で定義できる.

$$x_\mu^m = x_c + (m - 2\mu)\Delta \quad (0 \leq \mu \leq m) \quad (4.1)$$

ここで, x_c は多重極の中心の位置を表す. $\Delta = d/2$ はダイポール音源を形成するモノポール音源間の距離 d の半分の距離を表す.

各モノポール音源に与える重みは, 通常 ± 1 であるが, 隣り合う $m-1$ 次多重極を重ね合わせて m 次多重極を構成し, 重なりあうモノポール音源の重みについて和をとる場合は, 次式のとおり定義できる.

$$g_\mu^m = \begin{cases} 1 & (\mu = 0) \\ g_\mu^{m-1} - g_{\mu-1}^{m-1} & (0 < \mu < m) \\ -g_{\mu-1}^{m-1} & (\mu = m) \end{cases} \quad (4.2)$$

これを 2 次元に拡張する. x 軸方向に m 次, y 軸方向に n 次の極性を持つ (m, n) 次多重極に含まれる (μ, ν) 番目のモノポール音源の座標 $\mathbf{x}_{\mu, \nu}^{m, n} = (x_\mu^m, y_\nu^n)$ と重み $g_{\mu, \nu}^{m, n}$ は次式で定義できる.

$$\begin{aligned} x_\mu^m &= x_c + (m - 2\mu)\Delta & (0 \leq \mu \leq m) \\ y_\nu^n &= y_c + (n - 2\nu)\Delta & (0 \leq \nu \leq n) \end{aligned} \quad (4.3)$$

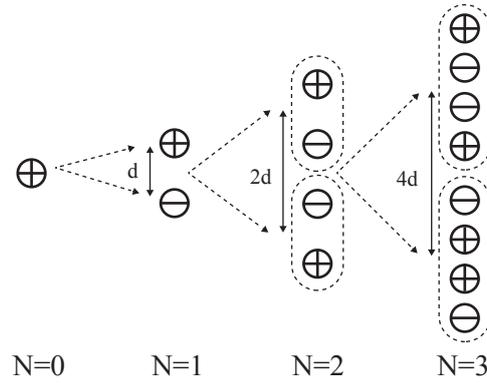


図 4.2: 多重極音源に含まれるモノポール音源の位置 (3次まで) [58].

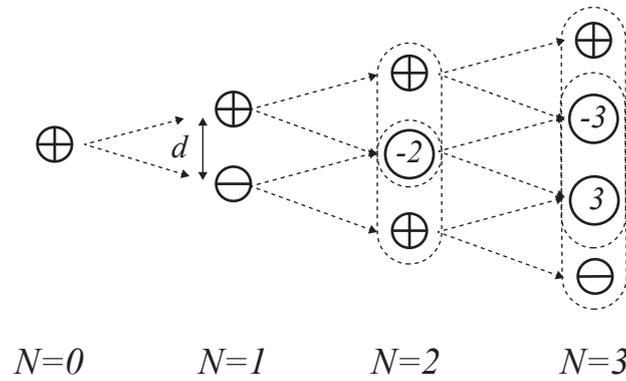


図 4.3: 隣り合う低次の多重極を距離 d ずらして重ね合わせて実現した多重極音源 [58].

$$g_{\mu,\nu}^{m,n} = g_{\mu}^m \cdot g_{\nu}^n \quad (4.4)$$

以上の各モノポール音源の座標とその重み係数から, (m, n) 次多重極を生成するために直線スピーカアレイのうちの位置 x_0 のスピーカに与える駆動関数は次式で定義できる.

$$D(\mathbf{x}_0) = \sum_{\mu=0}^m \sum_{\nu=0}^n g_{\mu,\nu}^{m,n} \cdot D(\mathbf{x}_0, \mathbf{x}_{\mu,\nu}^{m,n}) \quad (4.5)$$

ここで, $D(\mathbf{x}_0, \mathbf{x}_{\mu,\nu}^{m,n})$ は, 式 (2.7) で定義した焦点音源を位置 \mathbf{x}_0 に生成するための駆動関数である.

4.2 多重極音源の効率的表現方法

重みをつけた多重極音源を重ね合わせて指向性音源の音場を表現する際, 4.1 節で述べた多重極音源の構成方法を用いると, 偶数次と奇数次の多重極音源に含まれるモノ

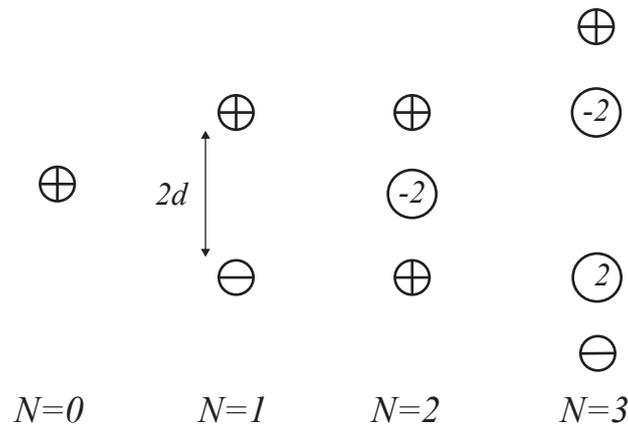


図 4.4: 少ない数のモノポール音源で作った多重極音源の効率的な実装 [58].

ポール音源の位置が重ならないため N 次までの多重極音源の重ね合わせに $2N + 1$ 個のモノポール音源が必要となる。羽田らは、図 4.4 に示すように、多重極に含まれるモノポール音源同士の間隔を等間隔としないことで、さらに多重極に含まれるモノポール音源の数を減らす手法を提案し、4 次までの 1 次元多重極音源の重ね合わせによる指向性制御が可能となることを示した [58]。この手法により、例えば、ダイポールを作る際のモノポール音源間の間隔を 2 次の多重極に含まれるモノポール音源間の距離の倍に設定し、これを重ね合わせることで 2 次までの多重極音源の重ね合わせを 3 個のモノポール音源のみで実現できる。本節では、これまで陽に定式化されていなかった羽田らによる手法を 2 次元に拡張し、かつ一般の次数の多重極音源の重ね合わせについて多重極音源に含まれるモノポール音源の座標とそれらに与える重み係数を定式化する。

まずは、多重極音源に含まれるモノポール音源の座標の定式化から始める。図 4.3 において、偶数次の多重極と奇数次の多重極を比較すると、偶数次の多重極に現れるモノポール音源は、奇数次に現れるモノポール音源の間に位置していることがわかる。これを、奇数次の多重極に現れるモノポール音源のうち、中心付近の 2 個の距離のみ他の 2 倍の距離を開けることで、図 4.4 に示すように偶数次の多重極に現れるモノポール音源で奇数次の多重極音源を構成できる。すなわち、偶数次の多重極に含まれるモノポール音源の座標は式 (4.3) で表され、奇数次の多重極に含まれるモノポー

ル音源の座標は、次式で表される。

$$\begin{aligned} x_\mu^m &= x_c + (m - 2\mu)\Delta & \left(0 \leq \mu \leq m, \quad \mu \neq \frac{m+1}{2}\right) \\ y_\nu^n &= y_c + (n - 2\nu)\Delta & \left(0 \leq \nu \leq n, \quad \nu \neq \frac{n+1}{2}\right) \end{aligned} \quad (4.6)$$

式(4.3)で表される偶数次の場合とは、多重極音源の中心近傍のモノポール音源に対してのみ、他の2倍の距離を設定している点が異なる。

次に、各多重極に含まれるモノポール音源に与える重み係数を求める。重み係数についてもモノポール音源の位置と同様に、奇数次の多重極のみ補正が必要となる。奇数次の多重極に含まれるモノポール音源に与える重み係数 g_μ^m は、式(4.2)に示す偶数次の多重極音源に含まれるモノポール音源に与える重み係数 g_μ^{m-1} を用いて、次式で定義できる。

$$g_\mu^m = \begin{cases} g_{\mu-1}^{m-1} & (\mu = 0, 1) \\ g_\mu^{m-1} - g_{\mu-2}^{m-1} & (1 < \mu < m-1) \\ -g_{\mu-1}^{m-1} & (\mu = m-1, m) \end{cases} \quad (4.7)$$

2次元に拡張する場合のモノポール音源の重み係数は、式(4.2)、(4.7)を、式(4.4)に代入することで算出できる。例として、図4.5に $N=4$ までの多重極音源に含まれるモノポール音源に与える重み係数を示す。図中の記号の違いが、対応する位置のモノポール音源の重みを表す。各記号と重みの値は図4.5に示すとおりである。

本手法により、例えば、129個のモノポール音源が必要な4次までの多重極音源の重ね合わせを、21個のモノポール音源で実現できる。

4.3 時間領域での駆動関数

焦点音源は、式(2.11)に示す時間領域での駆動関数を用いて効率よく生成される。時間領域での表現を用いることで、複数の焦点音源で生成される多重極音源も効率よく実現できる。式(2.11)、および式(4.5)から、 (m, n) 次の多重極音源を作るために $\mathbf{x}_\zeta = (x_\zeta, y_\zeta)$ に位置する ζ 番目の2次音源(直線状に並べた実スピーカ)に与える駆動関数は次式で与えられる。

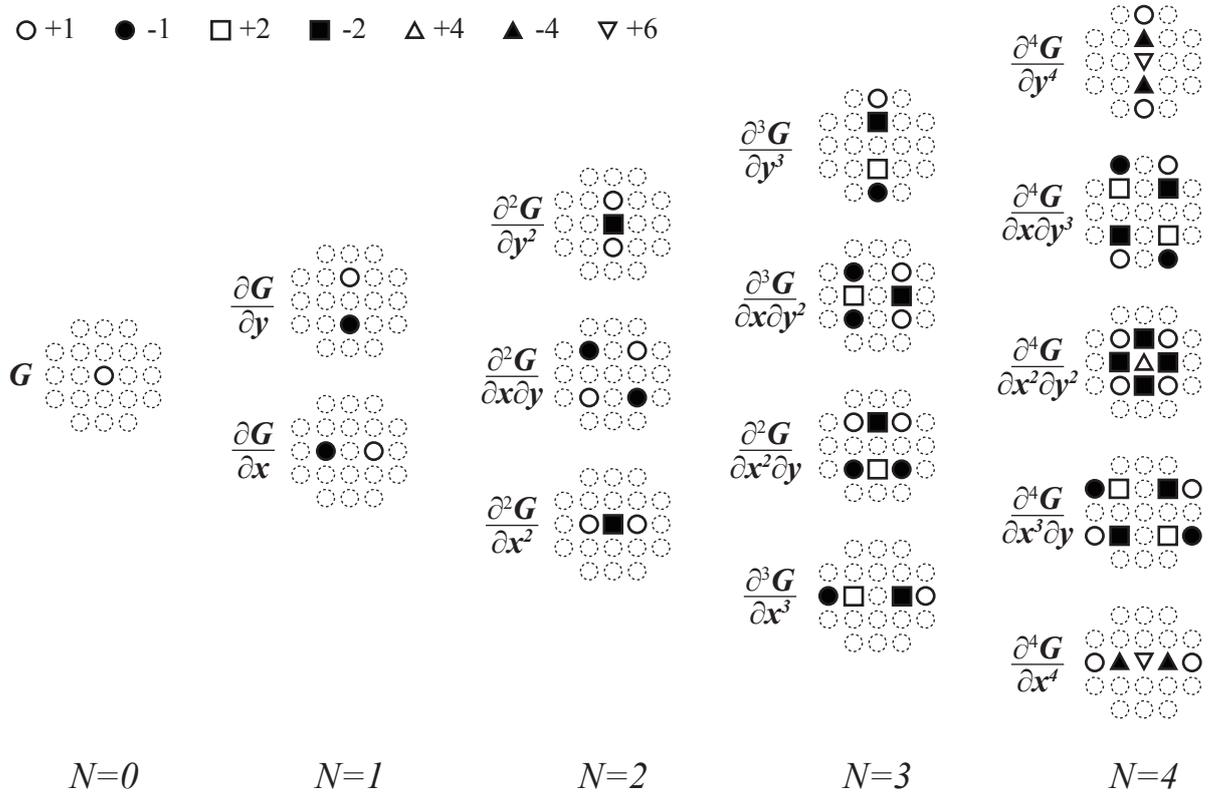


図 4.5: 4次までの多重極とそれに含まれるモノポール音源の位置と重み係数. 図中の記号の違いは, 対応する位置のモノポール音源の重みを表す. 各記号と重みの対応は図中に示すとおり.

$$\begin{aligned}
 u_{\zeta}(t) &= h(t) * \sum_{\mu=0}^m \sum_{\nu=0}^n \frac{g_{\mu,\nu}^{m,n} \cdot |y_{\nu}^n - y_{\zeta}|}{|\mathbf{x}_{\zeta} - \mathbf{x}_{\mu,\nu}^{m,n}|^3} \delta\left(t + \frac{|\mathbf{x}_{\zeta} - \mathbf{x}_{\mu,\nu}^{m,n}|}{c}\right) \\
 &= \sum_{\mu=0}^m \sum_{\nu=0}^n \frac{g_{\mu,\nu}^{m,n} \cdot |y_{\nu}^n - y_{\zeta}|}{|\mathbf{x}_{\zeta} - \mathbf{x}_{\mu,\nu}^{m,n}|^3} h\left(t + \frac{|\mathbf{x}_{\zeta} - \mathbf{x}_{\mu,\nu}^{m,n}|}{c}\right)
 \end{aligned} \tag{4.8}$$

ここで, $*$ は畳み込み演算を表す. $\mathbf{x}_{\mu,\nu}^{m,n} = (x_{\mu}^m, y_{\nu}^n)$ は (m, n) 次の多重極音源に含まれる (μ, ν) 番目のモノポール音源の位置を表し, $g_{\mu,\nu}^{m,n}$ は式 (4.4) で計算できる (μ, ν) 番目のモノポール音源に与える重み係数を表す.

以上, (m, n) 次の多重極音源を生成するための時間領域の駆動関数を導出したが, 複雑な指向特性を持つ指向性音源の音場は, これらの多重極音源に重みをつけて重ね合わせることで再現する. このような指向性音源を生成するための $\mathbf{x}_{\zeta} = (x_{\zeta}, y_{\zeta})$ に位置する ζ 番目の実スピーカに与える時間領域の駆動関数は次式で定義できる.

$$\begin{aligned}
u_{\zeta}(t) &= h(t) * \sum_{m=0}^N w_{m,N-m} \left(\sum_{\mu=0}^m \sum_{\nu=0}^{N-m} \frac{g_{\mu,\nu}^{m,N-m} \cdot |y_{\nu}^{N-m} - y_{\zeta}|}{|\mathbf{x}_{\zeta} - \mathbf{x}_{\mu,\nu}^{m,N-m}|^3} \delta\left(t + \frac{|\mathbf{x}_{\zeta} - \mathbf{x}_{\mu,\nu}^{m,N-m}|}{c}\right) \right) \\
&= \sum_{m=0}^N w_{m,N-m} \left(\sum_{\mu=0}^m \sum_{\nu=0}^{N-m} \frac{g_{\mu,\nu}^{m,N-m} \cdot |y_{\nu}^{N-m} - y_{\zeta}|}{|\mathbf{x}_{\zeta} - \mathbf{x}_{\mu,\nu}^{m,N-m}|^3} h\left(t + \frac{|\mathbf{x}_{\zeta} - \mathbf{x}_{\mu,\nu}^{m,N-m}|}{c}\right) \right) \quad (4.9)
\end{aligned}$$

ここで、 N は多重極音源の重ね合わせに用いる多重極音源の最高次数を表す。また、 $w_{m,n}$ は式 (2.70) で得られる (m, n) 次の多重極音源の展開係数を表す。

4.4 時間領域での実装における音場再現の精度改善

焦点音源法は、周波数領域の駆動関数では個々の実スピーカに対するフィルタ演算が必要になる。これに対し、時間領域の実装では実スピーカごとに異なるゲイン乗算と遅延処理で代用できるため、演算量を劇的に削減できる。しかし、周波数領域で定義されたフィルタでは可能であった任意の精度での非整数サンプルの遅延処理が、時間領域での実装ではデジタル信号のサンプリング周期に限られるため、焦点音源が作る音場の精度が低下する。特に、多重極音源のように複数の焦点音源を空間的に密に生成する場合に、ある焦点音源を作るための駆動関数が別の焦点音源に対しても寄与するため、多重極音源が本来生成する音場を高精度に再現することが困難となる。そこで本節では、3章同様に、2.1.5項で述べた非整数遅延フィルタを時間領域実装に組み込むことで、再現音場の精度改善を試みた検討について述べる。

4.4.1 Thiran フィルタを用いた多重極音源の時間領域実装

一般にスピーカアレイを構成するスピーカの間隔で決まるエイリアシング周波数より上の周波数帯域では、空間エイリアシングが発生するため、音場を正確に再現できない。さらにフィルタ長が短い Lagrange 補間では、低域通過特性により高音域のパワーが抑圧されるため音質が劣化する。全域通過型である Thiran フィルタは全帯域に渡りフラットな周波数特性を持ち、実スピーカアレイで一般的なスピーカ配置間隔におけるエイリアシング周波数以下ではフラットな群遅延特性を有する。そこで、Thiran フィルタで定義される非整数遅延フィルタを時間領域の多重極音源に適用して音場の再現精度改善を検討する。新たな時間領域の駆動関数は、式 (4.9) の駆動関数の後段に、Thiran フィルタを従属接続することで次式に示すとおり得られる。Thiran フィルタを組み合わせた新たな時間領域の駆動関数を図 4.6 に示す。

$$u_{\zeta}(t) = \sum_{m=0}^N w_{m,n} \left(\sum_{\mu=0}^m \sum_{\nu=0}^n \frac{g_{\mu,\nu}^{m,n} \cdot |y_{\nu}^n - y_{\zeta}|}{|\mathbf{x}_{\zeta} - \mathbf{x}_{\mu,\nu}^{m,n}|^3} h\left(t + \left\lfloor \frac{|\mathbf{x}_{\zeta} - \mathbf{x}_{\mu,\nu}^{m,n}|}{c} \right\rfloor\right) * p_{\text{Th}}(t, \tau_{\text{frac}}) \right) \quad (4.10)$$

ここで, $n = N - m$ とした. また, 非整数サンプル遅延 τ_{frac} は, 次式で算出できる.

$$\tau_{\text{frac}} = \frac{|\mathbf{x}_{\zeta} - \mathbf{x}_{\mu,\nu}^{m,n}|}{c} - \left\lfloor \frac{|\mathbf{x}_{\zeta} - \mathbf{x}_{\mu,\nu}^{m,n}|}{c} \right\rfloor \quad (4.11)$$

ここで, $\lfloor x \rfloor$ は, 小数 x の整数部分を取り出す操作に相当する.

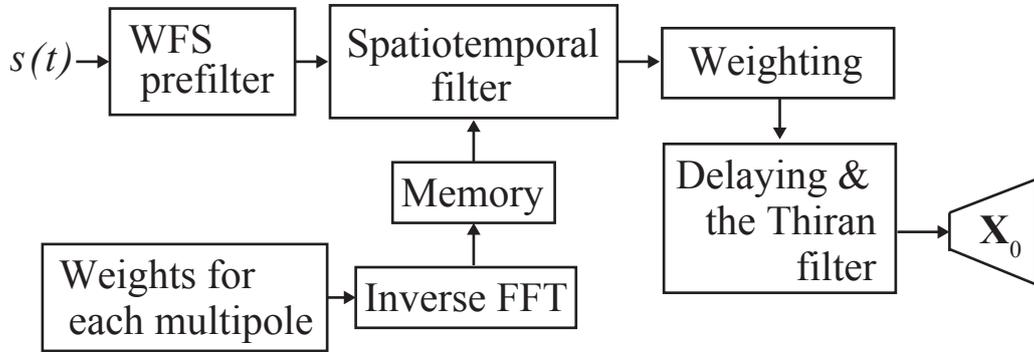


図 4.6: Thiran フィルタで精度改善した多重極音源の時間領域実装の処理図.

4.5 性能評価

まず最初に, 異なる重みを与えた複数の焦点音源が多重極音源を形成することを確認する. 次に, 隣り合う多重極の一部を重ね合わせて必要な焦点音源の数を削減する効率的な手法で, 同様の指向性制御ができることを確認する. さらに, 非整数遅延フィルタを組み合わせた時間領域での手法について, 再現される指向特性の精度を確認する. 最後に, 再現される指向特性の精度と演算量の観点から, 提案手法と従来手法の性能を評価する.

なお, 演算量の評価に当たっては, 4.2 節で述べた多重極の効率的な手法, 4.4.1 項で述べた Thiran フィルタにより音場再現精度を改善した時間領域の手法のみを用いた. 4.2 節で述べたような効率的な手法を用いない場合, 当然, 演算量は増加するためである. また, Thiran フィルタを用いない場合についても, 指向特性の再現精度が著しく劣化することが明らかのため, 評価には加えなかった.

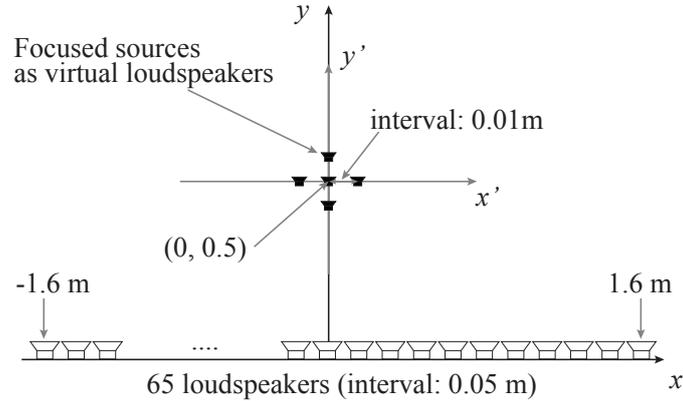


図 4.7: 実験のセットアップ。

4.5.1 実験条件

本章で実施するシミュレーションは、本節にまとめた実験条件で行う。例えば、多重極音源に含まれるモノポール音源同士の距離について性能比較する場合に、モノポール音源同士の間隔について記載するように、各実験の項においては本節で述べる条件からの差分のみを陽に記載する。なお、直線アレイの配置などの条件は、3.5.1項と同じであるが、再度記載する。

シミュレーションは MATLAB を用いて行った。実験の構成を図 4.7 に示す。音速は、343.36 m/s とした。また、 x 軸上の $-1.6 \leq x \leq 1.6$ の範囲に 0.05 m 間隔で 65 個のスピーカを直線状に並べてスピーカアレイを構成した。多重極音源に含まれるモノポール音源同士の距離は 0.01 m とした。多重極音源の中心は、 $(x_c, y_c) = (0.0, 0.5)$ とした。多重極音源に含まれる各モノポール音源の座標は、この中心座標と図 4.5 に記載の相対座標から計算できる。例えば、 $\frac{\partial G}{\partial y}$ に対応する多重極音源に含まれるモノポール音源の座標は、 $(0, 0.51)$ と $(0, 0.49)$ のように計算できる。また、多重極音源の重ね合わせの最大次数は $N = 4$ とした。目標となる音場及び指向特性を算出するため、目標音場の円調和級数 $\check{S}_\nu^{(2)}(\omega)$ を $2N + 1 = 9$ 個の乱数で生成した。当該乱数は、振幅 1.0 以下の複素数で生成した。目標音場は、次式で算出した。

$$S(\mathbf{x}, \omega) = \sum_{\nu=-N}^N \check{S}_\nu^{(2)}(\omega) H_\nu^{(2)}(kr) e^{j\nu\alpha} \quad (4.12)$$

また、多重極音源の重ね合わせで指向性音源の音場を再現する際に、各多重極音源に与える重み係数は次式で与えられる最小二乗法で算出した [51]。

$$\mathbf{w} = (\mathbf{G}^H \mathbf{G} + \lambda \mathbf{I})^{-1} \mathbf{G}^H \mathbf{s} \quad (4.13)$$

ここで、 \mathbf{w} は各多重極音源に与える重み係数を縦に並べたベクトル $\mathbf{w} = [w_{0,0}, \dots, w_{0,N}]^T$ である。また、 $[\cdot]^H$ は行列の共役転置を表す。 λ は行列 \mathbf{G} の逆行列の算出において解が不定になるのを防ぐための正則化項である [50]。本章では実験的に求めた値 $\lambda = 0.001$ を用いた。 \mathbf{s} は、目標となる指向性音源を中心に持つ半径 1 m の単位円上に等間隔（本章の実験では 2° 間隔）に設定した 180 点の制御点における音圧を縦に並べてできるベクトルである。伝達関数行列 $\mathbf{G} = [v_{0,0}, \dots, v_{0,N}]$ の各列 $v_{m,n}$ は、上記単位円の中心に (m, n) 次の多重極音源を配置した際に、上記制御点で観測される音圧を縦に並べてできるベクトルである。これは、単位円の中心に配置した (m, n) 次の多重極音源と単位円上の制御点の間の伝達関数と等価である [112]。

$$v_{m,n}(\mathbf{r}, k) = G_{2D}(kr)(jdk)^{m+n} \cos^m \phi \sin^n \phi \quad (4.14)$$

ここで、 k は波数、 G_{2D} はモノポール音源のグリーン関数、 j は虚数単位、 d は多重極音源に含まれるモノポール音源間の距離、 $\mathbf{r} = (r, \phi)$ は二次元極座標で表記した観測点の座標を表す。

4.5.2 客観評価尺度

提案手法の性能を評価するため、以下の音場の再現精度を表す式 (2.77) を用いた。音場の誤差評価尺度 $Err_{SF}(x, y)$ は、 $-1.5 \text{ m} \leq x \leq 1.5 \text{ m}$ 、 $1.0 \text{ m} \leq y \leq 3.0 \text{ m}$ の領域において、0.01 m 間隔の格子状に配置した観測点で計算した。

4.5.3 モノポール音源同士の間隔が再現精度に与える影響

多重極音源を構成するモノポール音源同士の間隔は、理論上は無限小とするのが妥当である。しかし、実際には空間微分を差分として計算するために有限の間隔を与えて実装する必要がある。したがって、モノポール音源同士の間隔が音場の誤差評価尺度に与える影響を把握しておくことは重要である。本項では、モノポール音源同士の間隔を変化させて音場の誤差評価尺度がどのように変化するかを明らかにする。実験条件は 4.5.1 節に記載のとおりであるが、モノポール音源同士の間隔のみ、0.01 m、0.02 m、0.04 m、0.07 m の 4 種類に変化させて、音場の誤差評価尺度を評価した。4.5.2 項に記載のとおり、 $-1.5 \text{ m} \leq x \leq 1.5 \text{ m}$ 、 $1.0 \text{ m} \leq y \leq 3.0 \text{ m}$ の領域において、0.01 m 間隔の格子状に配置した観測点で音場の誤差評価尺度を計算し、100 Hz から 7000 Hz までの範囲でそれぞれ平均値を算出した。

音場の誤差評価尺度のグラフを図4.8に示す。間隔0.04 m のときには、4300 Hz 付近から誤差が大きくなっていることがわかる。この周波数は、間隔0.04 m に対する空間エイリアシング周波数に対応している。すなわち、多重極音源を構成するモノポール音源同士の間隔も、空間エイリアシングが生じないように、再現したい周波数に対応する波長の半分の間隔とする必要があることを示している。なお、間隔0.07 m では波面の再現が困難であることがわかる。なお、間隔0.02 m に対応するエイリアシング周波数は約8600 Hz であるので、間隔0.02 m と0.01 m については、測定したレンジでは空間エイリアシングによる誤差の増加は観測できなかった。

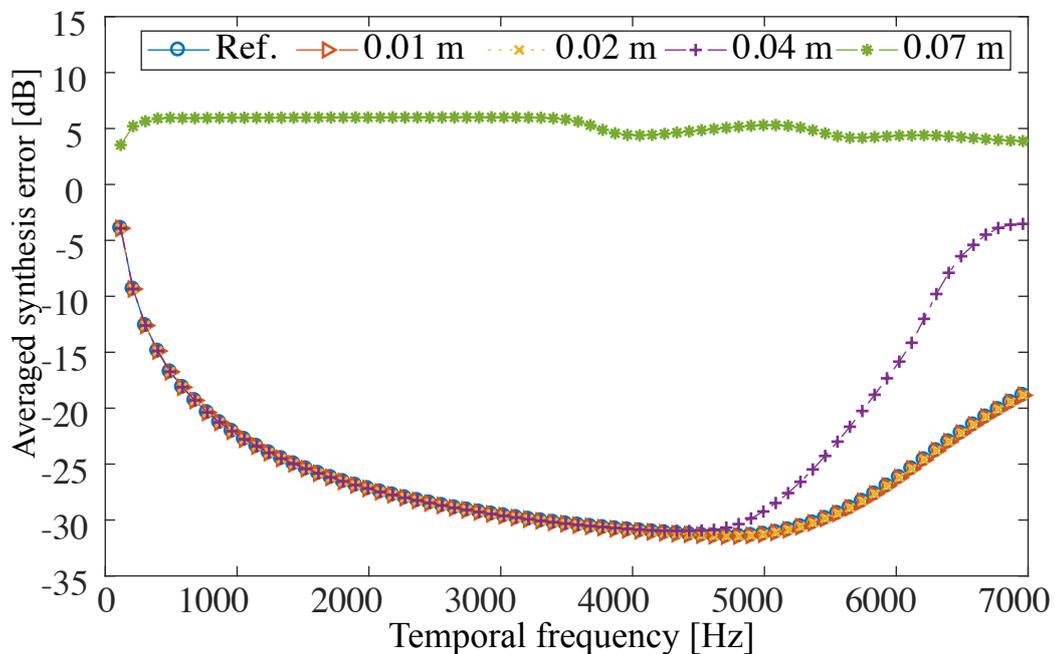


図4.8: 音場に関する誤差評価尺度の平均値。'Ref.': 従来手法（周波数領域での実装）。数値が記載されている凡例は、対応する値をモノポール音源同士の間隔としたときの提案手法（周波数領域での実装）。例）'0.01 m': モノポール音源間隔0.01 m のときの提案手法（周波数領域での実装）。

4.5.4 音場と再現精度

音場の再現精度を評価するため単一周波数に対する音場の再現実験を行った。シミュレーションには、1992.2 Hz の正弦波を用いた。なお、この周波数は1024点の離散フーリエ変換を行う際に、2000 Hz にもっとも近い周波数であったため採用した。指向性音源が回転した際の再現精度を比較するため、指向性音源が反時計方向に 10° 回転した際の音場についても生成して比較を行った。音源に回転を加えなかった際の音場の再現結果を図 4.9 に、音源を反時計方向に 10° 回転した際の再現結果を図 4.10 に示す。図 4.9 (a) は目標となる指向性音源が作る音場を表し、(b) (c) は周波数領域で実装した提案手法が作る音場と、目標音場と比較したときの音場の誤差評価尺度を示す。特に (c) から、焦点音源の受聴エリア内で高精度に音場を再現できていることがわかる。また、図 4.9 (d) (e) に、4.2 節で述べた効率の良い実装の多重極音源を用いて再現した音場と、音場の誤差評価尺度を示す。(b) (c) と比較すると、4.2 節で述べた効率的な実装を用いても音場の再現精度は維持できていることがわかる。(f) (g) は、時間領域で実装した提案手法が生成する音場と、目標音場からの音場の誤差評価尺度を表す。非整数サンプルの遅延処理ができない時間領域実装では、指向性音源の波面を合成できないことがわかる。(h) (i) は、時間領域で実装した提案手法を Thiran フィルタにより補間した際の音場と、目標音場からの誤差評価尺度を示す。精度は、周波数領域での実装に及ばないが、指向性音源の音場が再現できていることがわかる。図 4.10 より明らかなように、音源を反時計方向に 10° 回転した場合も同様の傾向が確認できる。

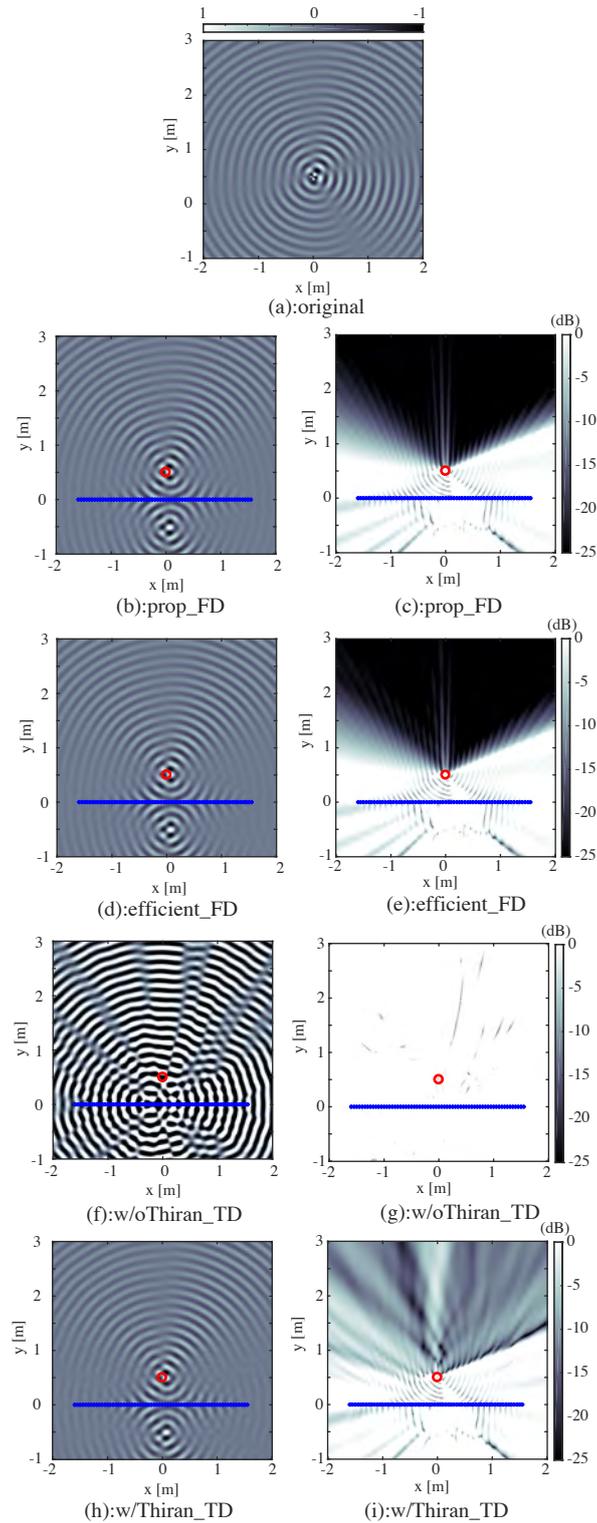


図 4.9: 再現された音場と、音場の誤差評価尺度. (a): 目標音場. (b): 周波数領域で実装した提案手法の再現音場. (c): (b) の音場の誤差評価尺度. (d): 周波数領域における効率的な実装を用いた提案手法の再現音場. (e): (d) の音場の誤差評価尺度. (f): 時間領域で実装した提案手法の再現音場. (g): (f) の音場の誤差評価尺度. (h): 時間領域で実装し、Thiran フィルタを適用した提案手法の再現音場. (i): (h) の音場の誤差評価尺度.

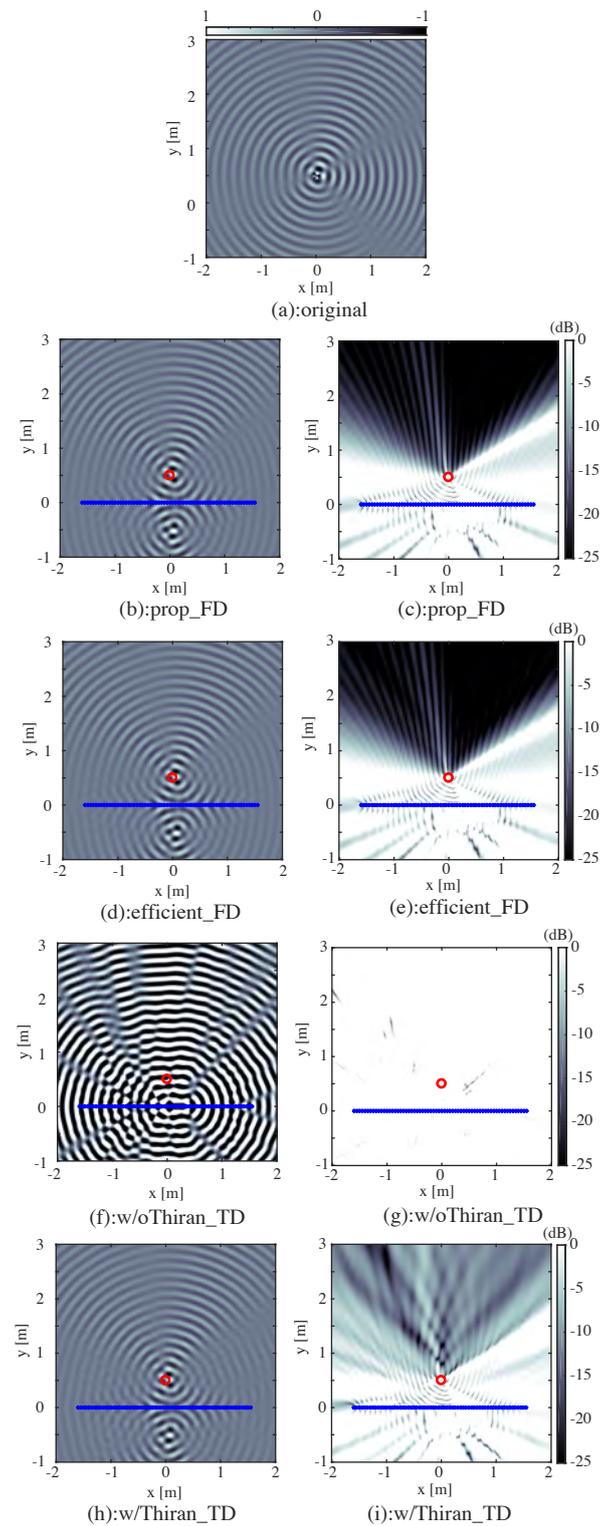


図 4.10: 指向性音源を反時計回りに回転 ($\phi = 10^\circ$) したときに再現された音場と、音場の誤差評価尺度. (a): 目標音場. (b): 周波数領域で実装した提案手法の再現音場. (c): (b) の音場の誤差評価尺度. (d): 周波数領域における効率的な実装を用いた提案手法の再現音場. (e): (d) の音場の誤差評価尺度. (f): 時間領域で実装した提案手法の再現音場. (g): (f) の音場の誤差評価尺度. (h): 時間領域で実装し、Thiran フィルタを適用した提案手法の再現音場. (i): (h) の音場の誤差評価尺度.

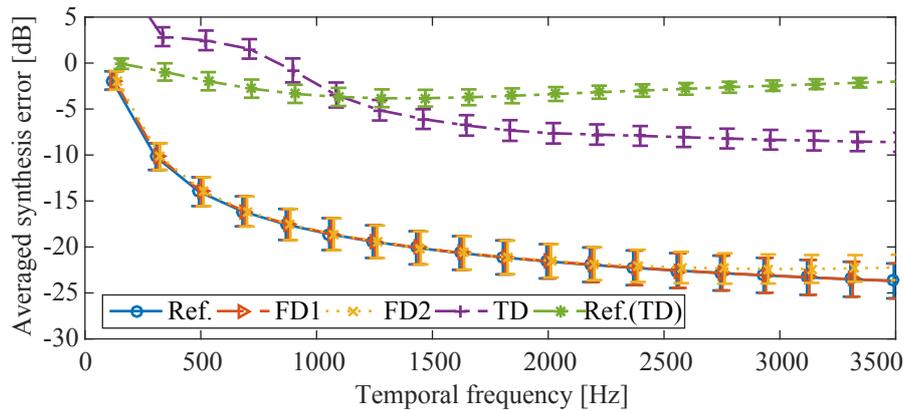


図 4.11: 音場の評価尺度について周波数ごとに算出した平均値. 'Ref.': 従来手法の周波数領域実装 [65, 68]. 'FD1': 提案手法の周波数領域実装. 'FD2': 効率的に実装した提案手法の周波数領域実装. 'TD': Thiran フィルタで補間した提案手法の時間領域実装. 'Ref.(TD)': 従来手法の時間領域実装 [65, 68].

音場の誤差評価尺度について算出した領域全体に関する平均値を, 空間エイリアシング周波数である 3500 Hz まで算出した結果を, 図 4.11 に示す. なお, 評価に用いる音源の指向特性に対する依存性を排するため, $2N + 1 = 9$ 個の乱数を 100 パターン生成し, 最大次数 $N = 4$ の円調和級数 $\check{S}_\nu^{(2)}(\omega)$ として用いた. 当該乱数は, 振幅 1.0 以下の複素数で生成した. また, 音源の平行移動が音場の再現精度におよぼす影響について確認するため, 仮想音源の位置を $-0.5 \text{ m} \leq x \leq 0.5 \text{ m}$, $0.25 \text{ m} \leq y \leq 0.75 \text{ m}$ の範囲でランダムに選択して, 上記 100 パターンの円調和展開級数のそれぞれに対応させた. さらに, 音源の回転が音場の再現精度におよぼす影響について確認するため, 仮想音源の向きを $-30^\circ \leq \phi \leq 30^\circ$ の範囲でランダムに選択して, 上記 100 パターンの円調和展開級数のそれぞれに対応させた. 100 パターンの計算結果の平均値を 95 % 信頼度区間とともに図 4.11 に示す. なお, 95 % 信頼度区間の算出においては student-t 分布を仮定した.

結果のグラフから, 周波数領域での実装では音場の再現精度に有意な差を確認できなかった. 一方, 時間領域の実装では, 1000 Hz を超える周波数帯域で提案手法の再現精度が従来手法の再現精度を有意に上回ることを確認した. また, 時間領域の実装と周波数領域の実装を比較すると, 時間領域の実装の再現精度は 15 dB 程度低下することが確認できる.

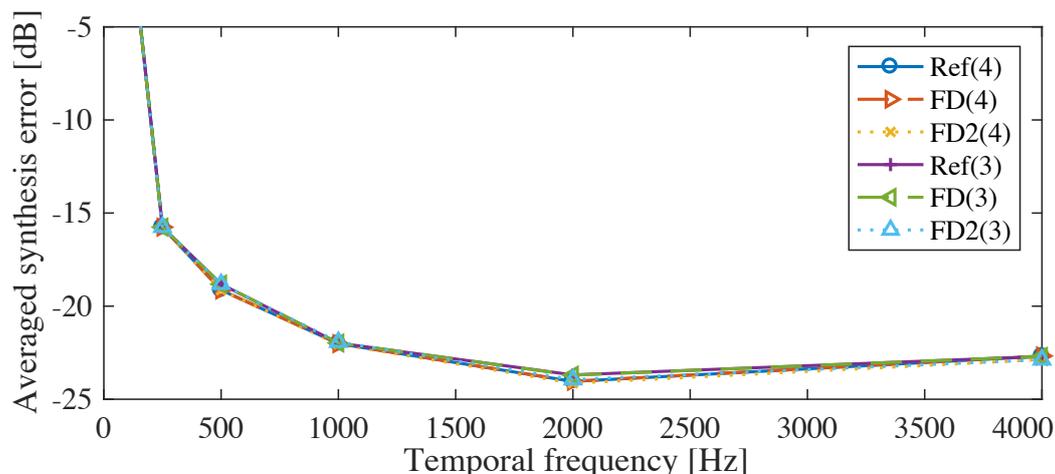


図 4.12: 周波数ごとの音場の誤差評価尺度. 凡例は以下のとおり. 'Ref(4)': 最大次数 $N = 4$ の従来手法 (周波数領域での実装). 'FD(4)': 最大次数 $N = 4$ の提案手法 (周波数領域での実装). 'FD2(4)': 最大次数 $N = 4$ の効率的に実装した提案手法の周波数領域実装. 'Ref(3)': 最大次数 $N = 3$ の従来手法 (周波数領域での実装). 'FD(3)': 最大次数 $N = 3$ の提案手法 (周波数領域での実装). 'FD2(3)': 最大次数 $N = 3$ の効率的に実装した提案手法の周波数領域実装.

次に, 2.5.2 項で用いた Monson らによる指向性のデータ [104] を用いて, 円調和展開級数の最大次数打ち切りがおよぼす誤差について評価する. 円調和展開級数の最大次数 $N = 3$ と $N = 4$ の場合について, 周波数領域実装の提案手法と従来手法の音場の誤差評価尺度を算出した. 結果を図 4.12 に示す. なお, 音場再現に用いる直線状スピーカアレイの間隔から, 3400 Hz 以上の音声帯域で空間エイリアジングが発生するため, 3400 Hz にもっとも近い周波数である 4000 Hz まで, 音場の再現精度を算出した.

音声発話の指向性を再現する場合, 最大次数 $N = 4$ の場合と $N = 3$ の場合を比較すると, 従来手法も提案手法も最大次数を $N = 3$ としても, 1 dB 程度しか再現精度が低下しないことが確認できる. また, 3 つの手法の再現精度の差は 1 dB 以下におさまることが確認できる.

表 4.1: 各手法について測定した演算量の推定結果 (weighted millions of operations per seconds. 以下, WMOPS). 'Reference': 時間領域で実装した従来手法 [65]. 'Proposed': 時間領域で実装した提案手法. 括弧の中の数字は多重極の次数を表す. 'Mono': 焦点音源法. (FD) は周波数領域での実装を表し, (TD) は時間領域の実装を表す [74].

Method	(WMOPS)
Reference in TD[68]	3850.88
Proposed in TD ($N = 2$)	1943.78
Proposed in TD ($N = 3$)	2713.75
Proposed in TD ($N = 4$)	4253.68
Monopole in FD[48]	8130.42

4.5.5 演算量の評価

演算量を評価するため, 従来法と提案法を C により実装し, ITU-T G.192 にて提供される浮動小数点版の演算量カウンタ [111] を各実装に組み込んだ. 周波数領域での実装に比べて時間領域での実装に必要な演算量が非常に少ないことは明らかであるので, 時間領域での実装のみを C で実装した.

本シミュレーションでは, 指向特性そのものは変化せず, その向きと位置のみが 20ms 毎に時間的に変化すると仮定した. シミュレーションにおけるスピーカ数は 65 とし, 波面合成プレフィルタとスピーカ毎に決まる時空間フィルタの長さをそれぞれ 513 サンプルと, 129 サンプルとした. 提案法における時空間フィルタは, 事前に計算してメモリに格納した.

計測した演算量を表 4.1 に示す. 無指向性の焦点音源を周波数領域実装で生成する際に必要な演算量を参考として記載する. 提案手法の演算量は, 多重極の最大次数が増加するにつれて増加したが, $N = 4$ までであれば無指向性の焦点音源を周波数領域実装で生成するのに必要な演算量よりも大幅に少ないことがわかる. 一方, 従来手法の演算量は, 円調和級数の次数が上がっても変化しなかった. これは, 従来手法の時空間フィルタの長さが円調和級数の次数と独立であり, 長さが変化しないためである.

3 次までの低次の多重極展開に対しては, 提案手法の演算量を用いた演算量の削減効果が期待でき, 従来手法と比較して全体で 29.5 % 演算量の削減効果が確認できた.

4.6 本章のまとめ

3章に引き続き、本章でも直線状スピーカアレイを用いてスピーカよりも前面に飛び出る指向性音源を生成する手法を研究した。本手法は直線状スピーカアレイにより生成した複数の焦点音源で多重極音源を生成し、この多重極音源を用いて指向特性の制御を行うものである。多重極音源の実装には、隣り合う低次の多重極音源を重ね合わせて配置する効率的な実装手法を用いることで、多重極音源の生成に必要な焦点音源の数を削減できることを確認した。また、さらなる演算量の削減のための時間領域実装を検討した。多重極音源は複数の焦点音源を密に配置して生成する。そのため、各焦点音源を高精度に生成できない場合、ある焦点音源を作るための音響信号が近傍の焦点音源に干渉を与え、結果として所望の精度の指向特性が再現できなかった。そこで、非整数遅延処理を実現する Thiran フィルタを時間領域実装に組み合わせることで、近傍の焦点音源への干渉を低減し、指向特性の再現精度が改善することを確認した。最後に、従来法と提案法について演算量の評価を行い、3次までの多重極音源の重ね合わせについて、従来手法より低演算量で実現できることを確認した。

4.6.1 本章の貢献と関連研究

本章の内容は、研究業績リスト [J-1] の内容をまとめたものである。本章の研究の貢献は、焦点音源で生成した仮想の円状スピーカアレイを制御して指向性音源が作る音場を再現し、特に時間領域で従来法より低演算量で高精度な音場再現ができることを示した点にある。従来も多重極音源の一部を重ね合わせることで多重極音源を構成するのに必要なモノポール音源の数を削減する手法について検討が行われてきたが、1次元かつ4次までの場合に限定したものであった。この手法を拡張し、2次元かつ任意の次数の多重極音源について陽に定式化したことも本章の貢献の一つである。シミュレーションにより、モノポール音源の数を削減した提案手法が従来手法と同等の音場再現精度を達成することも確認できた。本章の研究により、多重極展開を用いた手法において多重極音源を構成するモノポール音源の数を削減することが可能となるため、演算量を削減することが可能となる。多重極音源の原理に基づき音場を再現する多重極スピーカアレイの実用化には演算量の削減が不可欠である。本章の研究は、演算量の削減を通じて多重極スピーカアレイの実用化に貢献するといえる。

第5章 音場再現における円調和級数展開 と多重極展開

直線状スピーカアレイを用いた焦点音源法を拡張し、スピーカよりも前面に飛び出す仮想音源に指向特性を与える手法について3章、及び4章で説明した。3章で提案した手法は、指向性音源が作る音場を円調和級数でモデル化する手法である。一方、4章で提案した手法は、指向性音源が作る音場を多重極音源の重ね合わせ（多重極展開）でモデル化する手法である。円調和級数展開と多重極展開は、いずれも平面の音場をモデル化するために用いられるが、これらの間の関係については幾つかの先行研究の中で議論されてきた [54] [58] [70]。その中では、低次の多重極音源であるダイポールやクアドラポールが円調和関数や球面調和関数などでモデル化されることや [54]、低次の多重極音源が生成する指向特性を球面調和級数を基底とする駆動関数と直線状スピーカアレイを用いて生成する方法 [58]、円調和関数や球面調和関数を直交基底として持つ空間を多重極展開で等価に表現できること [70]、が明らかにされてきた。しかし、目標となる音場をモデル化した円調和級数から解析的な変換により、多重極音源に与える重み係数を求める手法については検討がされてこなかった。本章では、円調和級数展開でモデル化した音場と、多重極展開でモデル化した音場を係数比較することで、円調和級数を多重極音源の重み係数に変換する手法について提案する。本研究のモチベーションは、4章において最小2乗法で求めていた多重極の展開係数を、解析的に求める手法を与えることである。当該手法を4章で提案した手法に組み合わせることで、直線アレイで生成した仮想音源に円調和級数でモデル化した指向特性が付与できることを示す。

5.1 直線アレイによる音場再現手法の比較

直線状スピーカアレイを用いた指向性音源の音場再現手法として、円調和展開を用いる手法を3章で、多重極展開を用いる手法を4章で議論した。本節では、これらの手法のメリットとデメリットについて述べる。

直線状スピーカアレイで生成した焦点音源で円状アレイを生成し、これを制御して指向性音源の音場を再現する3章の手法によれば、周波数領域の実装を用いて従来手法と同等の音場再現精度が達成できる上、時間領域の実装では従来手法よりも低演算

量で、より高精度な音場再現ができることを確認した。一方、3章の手法は開放型の円状スピーカアレイを用いた音場制御と等価であるため、駆動関数の式(3.8)に示すとおり分母に高次のベッセル関数が現れる。高次のベッセル関数は、図2.10に示すとおり零点を有するため、零点付近で駆動関数が過剰なゲインを有することがある。したがって、実装に当たっては零点由来の過剰なゲインが生じていないか注意を必要とする。

次に、直線状スピーカアレイで生成した焦点音源で多重極音源を生成し、これを制御して指向性音源の音場を再現する4章の手法によれば、周波数領域の実装により従来手法と同等の音場再現精度が達成できる。時間領域の実装では、多重極の最大次数3まで、かつ1500 Hz以上の高周波数帯域に限定ではあるが、従来手法よりも低演算量かつ高精度な音場再現ができることを確認した。4章の多重極音源に基づく手法は、3章の円調和級数展開に基づく手法のように零点由来の過剰なゲインに注意する必要はないが、次数ごとの多重極音源に対して適切な重み係数が算出できるよう、実装において正則化パラメータに注意する必要がある。

以上を踏まえると、ライブイベントなどのようにリアルタイム動作が要求されるアプリケーションにおいては、円調和級数展開に基づく3章の手法を時間領域で実装して用いるのが良いと言える。また、十分な時間をかけてレンダリングすることが許される、例えばコンテンツ制作のようなアプリケーションでは、零点由来の過剰なゲインに注意する必要がない多重極展開に基づく4章の手法を周波数領域で実装して用いることで、指向特性を高精度に再現したコンテンツを制作できると考えられる。

5.2 単位円上での係数比較による円調和級数の解析的変換

5.1節において、4章の多重極音源に基づく手法では、次数ごとの多重極音源に対して適切な重み係数が算出できるよう、実装において最小2乗法における正則化パラメータに注意する必要があると述べた。本節では、これら重み係数を正則化パラメータに依存することなく、指向性音源が生成する音場から直接算出する手法について検討する。具体的には、円状マイクアレイなどで収録された複数チャンネルの音声データから円調和展開級数を取得し、この円調和展開級数を解析的に変換することで多重極音源の重み係数を算出する手法を検討する。円調和展開級数は円状マイクアレイと高速フーリエ変換を用いて高速に取得できることが知られているため、本節の手法と組み合わせることで、正則化パラメータなどへの依存性を排した形で多重極音源による音場再現を実現できるようになる。

5.2.1 関係式の導出

2次元の直交座標において、原点を中心とする半径 1 m の円上に設定した観測点 $\mathbf{x} = (\cos \alpha, \sin \alpha)$ での音場は、原点まわりでのテイラー展開を用いて次式で表現できる。

$$S(\mathbf{x}) = \sum_{m+n=0}^{\infty} \sum_{m=0}^{m+n} \frac{\partial^{m+n} S(\mathbf{x}')}{\partial x^m \partial y^n} \cdot \frac{\cos^m \alpha \cdot \sin^n \alpha}{m! \cdot n!} \Bigg|_{\mathbf{x}'=0} \quad (5.1)$$

一方、観測点 $\mathbf{x} = (\cos \alpha, \sin \alpha)$ の音場は、円調和級数展開を用いて次式で表現できる。

$$S(\mathbf{x}) = \sum_{\nu=-\infty}^{\infty} \check{S}_{\nu}^{(2)}(\omega) H_{\nu}^{(2)}(kr) e^{j\nu\alpha} \quad (5.2)$$

ここで、 $\check{S}_{\nu}^{(2)}$ は ν 次の円調和級数を表し、 $H_{\nu}^{(2)}$ は ν 次の第2種ハンケル関数を表す。ここで、複素指数関数の項 $e^{j\nu\alpha}$ に着目する。一般に、複素数には次式で定義されるオイラーの関係式が成り立つ。

$$e^{j\alpha} = \cos \alpha + j \sin \alpha \quad (5.3)$$

この関係式を複素指数関数の項 $e^{j\nu\alpha}$ に適用して二項展開すると、次式のように $\cos \alpha$ と $\sin \alpha$ のべき乗で書き換えることができる。

$$\begin{aligned} e^{j\nu\alpha} &= (\cos \alpha + j \sin \alpha)^{\nu} \\ &= \sum_{\zeta=0}^{\nu} \binom{\nu}{\nu - \zeta} j^{\zeta} \cos^{\nu - \zeta} \alpha \sin^{\zeta} \alpha \end{aligned} \quad (5.4)$$

この関係式を式 (5.2) に適用すると次式を得る。

$$\begin{aligned} S(\mathbf{x}) &= \sum_{\nu=-\infty}^{\infty} \check{S}_{\nu}^{(2)}(\omega) H_{\nu}^{(2)}(kr) e^{j\nu\alpha} \\ &= \check{S}_0^{(2)} H_0^{(2)}(kr) + \sum_{\nu=1}^{\infty} \check{S}_{\nu}^{(2)}(\omega) H_{\nu}^{(2)}(kr) e^{j\nu\alpha} + \sum_{\nu=1}^{\infty} \check{S}_{-\nu}^{(2)}(\omega) H_{-\nu}^{(2)}(kr) e^{-j\nu\alpha} \\ &= \check{S}_0^{(2)} H_0^{(2)}(kr) + \sum_{\nu=1}^{\infty} \check{S}_{\nu}^{(2)}(\omega) H_{\nu}^{(2)}(kr) \sum_{\zeta=0}^{\nu} \binom{\nu}{\nu - \zeta} j^{\zeta} \cos^{\nu - \zeta} \alpha \sin^{\zeta} \alpha + \\ &\quad \sum_{\nu=1}^{\infty} \check{S}_{-\nu}^{(2)}(\omega) H_{-\nu}^{(2)}(kr) \sum_{\zeta=0}^{\nu} \binom{\nu}{\nu - \zeta} (-j)^{\zeta} \cos^{\nu - \zeta} \alpha \sin^{\zeta} \alpha \\ &= \check{S}_0^{(2)} H_0^{(2)} + \sum_{\nu=1}^{\infty} \sum_{\zeta=0}^{\nu} \binom{\nu}{\nu - \zeta} j^{\zeta} \left\{ \check{S}_{\nu}^{(2)} H_{\nu}^{(2)} + (-1)^{\zeta} \check{S}_{-\nu}^{(2)} H_{-\nu}^{(2)} \right\} \cos^{\nu - \zeta} \alpha \sin^{\zeta} \alpha \end{aligned} \quad (5.5)$$

ここで、式 (5.1) との係数比較のために、 $m = \nu - \zeta$, $n = \zeta$ と置き換えると、

$$S(\mathbf{x}) = \check{S}_0^{(2)} H_0^{(2)} + \sum_{m+n=1}^{\infty} \sum_{n=0}^{m+n} \binom{m+n}{m} j^n \left\{ \check{S}_{m+n}^{(2)} H_{m+n}^{(2)} + (-1)^n \check{S}_{-m-n}^{(2)} H_{-m-n}^{(2)} \right\} \cos^m \alpha \sin^n \alpha \quad (5.6)$$

今、式 (5.1) および式 (5.6) に現れる $\cos^m \alpha \cdot \sin^n \alpha$ の係数を比較すると、以下の関係式を得る。

$$\left. \frac{\partial^{m+n} S(\mathbf{x})}{\partial x^m \partial y^n} \right|_{\mathbf{x}=\mathbf{0}} = j^n \cdot (m+n)! \cdot \left\{ \check{S}_{m+n}^{(2)}(\omega) H_{m+n}^{(2)}(kr) + (-1)^n \check{S}_{-m-n}^{(2)}(\omega) H_{-m-n}^{(2)}(kr) \right\} \quad (5.7)$$

さて、式 (2.64) のとおり、指向性音源がつくる音場はグリーン関数の高次微分の重みつき和で表すことができる。一方、同じ音場を式 (5.1) に示すように原点まわりのテイラー展開でも表せることから、係数比較により式 (5.7) とあわせて次の関係式を得る。

$$w_{m,n} = \left. \frac{\partial^{m+n} S(\mathbf{x})}{\partial x^m \partial y^n} \right|_{\mathbf{x}=\mathbf{0}} \quad (5.8)$$

式 (5.8) を式 (2.64) に代入することにより、円調和展開級数でモデル化した指向性音源の音場を多重極音源の重ねあわせで表現することができる。

5.2.2 予備実験

焦点音源で構成した仮想の多重極音源を用いる前に、理想の多重極音源を利用した場合に 5.2.1 項の手法で求めた重み係数を用いた際の音場の再現精度を評価する予備実験を行う。予備実験では、最大次数 $N = 4$ 、多重極音源の中心 $(x_c, y_c) = (0, 0)$ とした。なお、9 個 ($2N + 1 = 9$) の最大振幅 1 の複素数の乱数を生成して指向性音源をモデル化し、円調和展開級数とした。この級数を用いて、目標音場を式 (2.71) で生成し、提案手法で生成される音場は式 (5.8) で計算される多重極音源の展開係数を式 (2.64) に代入して算出した。

比較のため、pressure-matching method [51] で重み係数を求めて、式 (2.70) で音場を計算した。なお、ここでは半径 1 m の円周上に 2° 刻みで 180 個の制御点を設定した。

図 5.1 に計算結果を示す。(a) は目標音場、(b) はその放射パワー、(c) は提案手法の再現音場、(d) は提案手法の放射パワー、(e) は式 (2.77) で計算できる提案手法の音場と目標音場に関する音場の誤差分布を表す。また、(f) は pressure-matching method の音場を表し、(g) は pressure-matching method の放射パワーを表し、(h) は式 (2.77) で計算される pressure-matching method の目標音場との音場に関する誤差分布を表す。目標音場に関する放射パワーと、提案手法および pressure-matching method の放射パワーの間には大きな差はないように見える。しかし、目標音場と各手法の音場の

間の音場の誤差分布（右列）を比較すると, pressure-matching method では高い精度で音場を再現できているのに対し, 提案手法では音場の再現ができていない. この結果から, 放射されるパワーのみ再現精度が高いものの, 提案手法では正確な波面を再現できていないことがわかる. その理由は, 式 (5.1) に示すとおり半径 1 m の円上に設定した制御点での音圧について, 原点まわりのテイラー展開を行うことで関係式を導出しているためと考えられる. つまり, この円上を除く任意の観測点では仮定が成り立たず, 誤差が拡大すると考えられる.

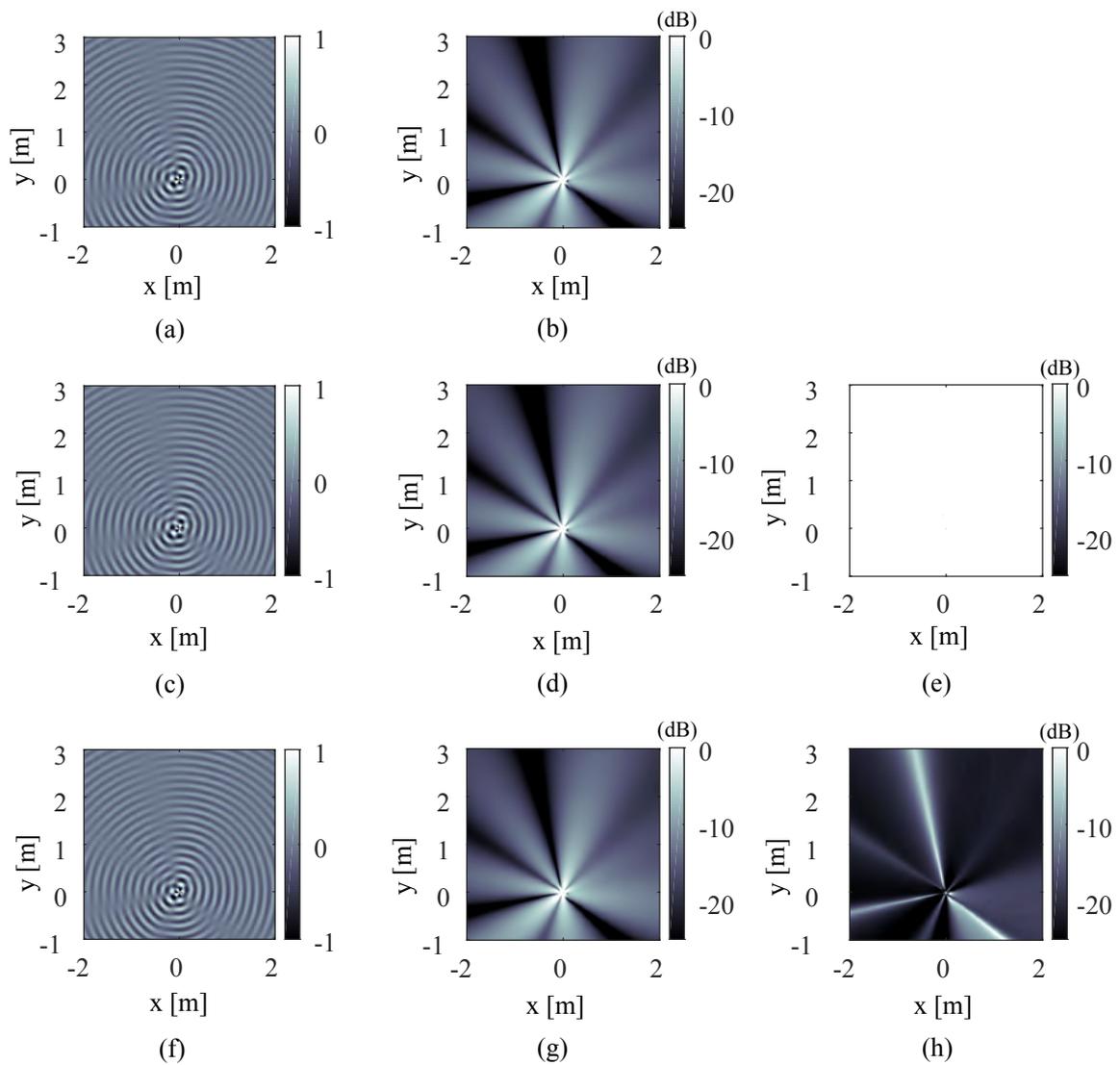


図 5.1: 左列：目標音場 (a) と提案手法および pressure-matching method の音場. 中央列：各手法の放射パワー. 右列：式 (2.77) で計算できる目標音場と各手法の音場の間の誤差分布. 上段：目標音源. 中段：提案手法. 下段：pressure-matching method.

5.3 多重極展開級数と円調和展開級数の係数比較による解析的変換

5.2.1項で導出した手法によれば, 指向性音源が生成する音場をモデル化した円調和展開係数から, 正規化パラメータなどに依存することなく元の音源の指向特性を再現することができる. しかし, 方向ごとに放射されるパワーの再現性は十分なものの, 再現される音場の再現精度は pressure-matching method よりも劣化することが確認された. 単位円上で音圧を一致させることを前提とした 5.2.1項の手法に対し, 本項では音源から遠方の領域全体で音圧を一致させることを前提として, 円調和展開係数を多重極音源の重み係数に変換する手法を導出する. 本節の手法により, 再現対象の音源の指向特性だけでなく音場そのものを高精度に再現可能となることが期待できる.

5.3.1 関係式の導出

5.2.1項では, 音場の原点周りでのテイラー展開と同音場の円調和展開を単位円上で比較することにより, 式(2.64)の多重極音源の重ね合わせに用いる重み係数 $w_{m,n}$ を円調和展開級数からの解析的変換により算出する手法を導出した. 単位円上において, 原点周りでテイラー展開した展開級数は多重極音源の重み係数に対応する. 一方, 円調和級数展開に現れる複素指数関数を二項展開して \cos 関数と \sin 関数のべき乗に整理すると, これらに係る係数も多重極音源の重み係数に対応した. この関係を用いた係数比較から, 円調和展開係数を解析的に変換して多重極音源の重み係数を得る手法を導出した. しかし, 音場の表現が原点周りのテイラー展開であることや, 制御点が単位円上に限定されることから, 原点から離れた観測点において音場のモデル化誤差が拡大すると考えられる. そこで本項ではこの手法を拡張し, 制御点が原点から遠方に離れても音場を正確に再現できる手法を導出する.

式(2.61)からもわかるとおり, (m, n) 次の多重極音源は $\cos^m \phi \sin^n \phi$ に比例する音場を生成する. そこで, 式(2.71)の音場の円調和級数展開を $\cos^m \phi \sin^n \phi$ の線形和の形に変形し, その係数を比較することで円調和級数展開の係数から多重極音源の重み係数を解析的に導出する. 式(2.71)を以下のとおり変形する. なお簡単のため, 以下 $H_\mu^{(2)}(kr)$ を $H_\mu^{(2)}$ と表記する.

$$\begin{aligned} S(\mathbf{r}, \omega) &= \sum_{\mu=-\infty}^{\infty} C_\mu H_\mu^{(2)} e^{j\mu\phi} \\ &= C_0 H_0^{(2)} + \sum_{\mu=1}^{\infty} H_\mu^{(2)} \{C_\mu e^{j\mu\phi} + (-1)^\mu C_{-\mu} e^{-j\mu\phi}\} \end{aligned}$$

これに, $H_\mu^{(2)} \approx j^\mu H_0^{(2)}$ の関係式を適用して以下のように変形する.

$$S(\mathbf{r}, \omega) = H_0^{(2)} \left[C_0 + \sum_{\mu=1}^{\infty} j^\mu \{ C_\mu e^{j\mu\phi} + (-1)^\mu C_{-\mu} e^{-j\mu\phi} \} \right]$$

ここに現れる複素指数関数 $e^{j\mu\phi}$ に, 式 (5.3) のオイラーの関係式と式 (5.4) の二項展開を適用することで, さらに以下のように変形する.

$$S(\mathbf{r}, \omega) = H_0^{(2)} \left[C_0 + \sum_{\mu=1}^{\infty} j^\mu \sum_{n=0}^{\mu} j^n \binom{\mu}{n} \{ C_\mu + (-1)^{\mu-n} C_{-\mu} \} \cdot \cos^{\mu-n} \phi \sin^n \phi \right] \quad (5.9)$$

一方, 式 (5.9) で表される多重極音源の重ね合わせを以下のとおり変形する.

$$\begin{aligned} S(\mathbf{r}, \omega) &= \sum_{\mu=0}^{\infty} \sum_{n=0}^{\mu} w_{\mu-n,n} \frac{\partial^\mu G(\mathbf{r}, \omega)}{\partial x^{\mu-n} \partial y^n} \\ &= \sum_{\mu=0}^{\infty} \sum_{n=0}^{\mu} w_{\mu-n,n} G_{2D}(\mathbf{r}, \omega) (j2dk)^\mu \cos^{\mu-n} \phi \sin^n \phi \\ &= \sum_{\mu=0}^{\infty} \sum_{n=0}^{\mu} w_{\mu-n,n} \frac{j}{4} H_0^{(2)} (j2dk)^\mu \cos^{\mu-n} \phi \sin^n \phi \\ &= \sum_{\mu=0}^{\infty} j^\mu \sum_{n=0}^{\mu} w_{\mu-n,n} \frac{j}{4} H_0^{(2)} (2dk)^\mu \cos^{\mu-n} \phi \sin^n \phi \end{aligned} \quad (5.10)$$

$\cos^{\mu-n} \phi \sin^n \phi$ の係数を比較して, 多重極音源に与える重み係数は円調和展開の級数を用いて次式のとおり書ける.

$$w_{\mu-n,n} \frac{j}{4} (2dk)^\mu = j^n \binom{\mu}{n} \{ C_\mu + (-1)^{\mu-n} C_{-\mu} \}$$

ここで μ を $m+n$ で置き換えると, 重み係数が次式のとおり得られる.

$$w_{m,n} = \frac{4j^{n-1}}{(2dk)^{m+n}} \binom{m+n}{n} \{ C_{m+n} + (-1)^m C_{-m-n} \} \quad (5.11)$$

この重み係数を式 (5.10) で定義される多重極音源の重ね合わせに代入することで, 円調和展開級数でモデル化された指向性音源の音場を再現できる.

5.3.2 単位円上での係数比較による解析的変換との関係性

5.2.1 項では, 単位円上の音圧分布を原点周りのテイラー展開と円調和級数展開の 2 とおりの手法で表現して係数比較することで, 多重極音源の重み係数を導出した [91].

この重み係数は次式で与えられる.

$$\begin{aligned} D_{m,n}^{(old)} &= \frac{j^n}{(j2dk)^{m+n}} \binom{m+n}{n} \cdot \left\{ C_{m+n} H_{m+n}^{(2)}(k) + (-1)^n C_{-m-n} H_{-m-n}^{(2)}(k) \right\} \\ &= \frac{j^n H_0^{(2)}(k)}{(2dk)^{m+n}} \binom{m+n}{n} \{ C_{m+n} + (-1)^m C_{-m-n} \} \end{aligned} \quad (5.12)$$

本項で提案する式 (5.11) に基づく解析的変換手法と比較して, 式 (5.7) は次式のとおり変形できる.

$$w_{m,n}^{(old)} = \frac{j}{4} H_0^{(2)}(k) D_{m,n} \quad (5.13)$$

この式から, 5.2.1 項で提案した式 (5.7) の解析的変換は, 本項で提案する式 (5.11) の解析的変換に単位円上での自由空間グリーン関数 $G_{2D}(0)$ の値がかかったものになっている. この項が音場全体にかかっているために, 各点で位相が回転して音場の正確な再現ができなかったと考えられる. しかし, 定数がかかっているだけであるので, 波面の正確な再現よりも指向特性が重要となるアプリケーションに対しては, 5.2.1 項で導出した解析的変換が有効に利用できる.

5.3.3 予備実験

焦点音源で構成した仮想の多重極音源を用いる前に, 理想の多重極音源を利用した場合に 5.3.1 項の手法で求めた重み係数が達成する指向特性や音場の再現精度を確認する予備実験を行う. 実験条件は 5.2.2 項に述べたものと同様である.

予備実験では, 最大次数 $N = 4$, 多重極音源の中心 $(x_c, y_c) = (0, 0)$ とした場合に式 (4.3) および式 (4.6) に示す位置にモノポール音源を配置して多重極音源を構成した. なお, 9 個 ($2N + 1 = 9$) の最大振幅 1 の複素数の乱数を生成して指向性音源をモデル化した円調和展開係数とした. 多重極音源の位置は座標系の原点とし, 多重極音源に含まれるモノポール音源の間隔を 0.01 m とした. これを用いて, 目標音場を式 (2.71) を用いて算出し, 提案手法で生成される音場は式 (2.64) により算出した. また, 5.2.1 項の予備実験と同様に, 式 (2.70) の pressure-matching method で求めた重み係数を式 (2.64) に代入して, 比較のための音場を計算した.

図 5.2 に計算結果を示す. (a) は目標音場, (b) はその放射パワー, (c) は提案手法の音場, (d) は提案手法の放射パワー, (e) は式 (2.77) で計算した提案手法の音場と目標音場に関する音場の誤差分布を表す. また, (f) は pressure-matching method の音場, (g) は pressure-matching method の放射パワー, (h) は式 (2.77) で計算した pressure-matching method の目標音場との音場に関する誤差分布を表す. 目標音場に関する放

射パワーと、提案手法および pressure-matching method の放射パワーの間には大きな差はない。目標音場と各手法の音場との音場の誤差分布（右列）を比較すると、提案手法も pressure-matching method も高い精度で音場を再現できていることがわかる。

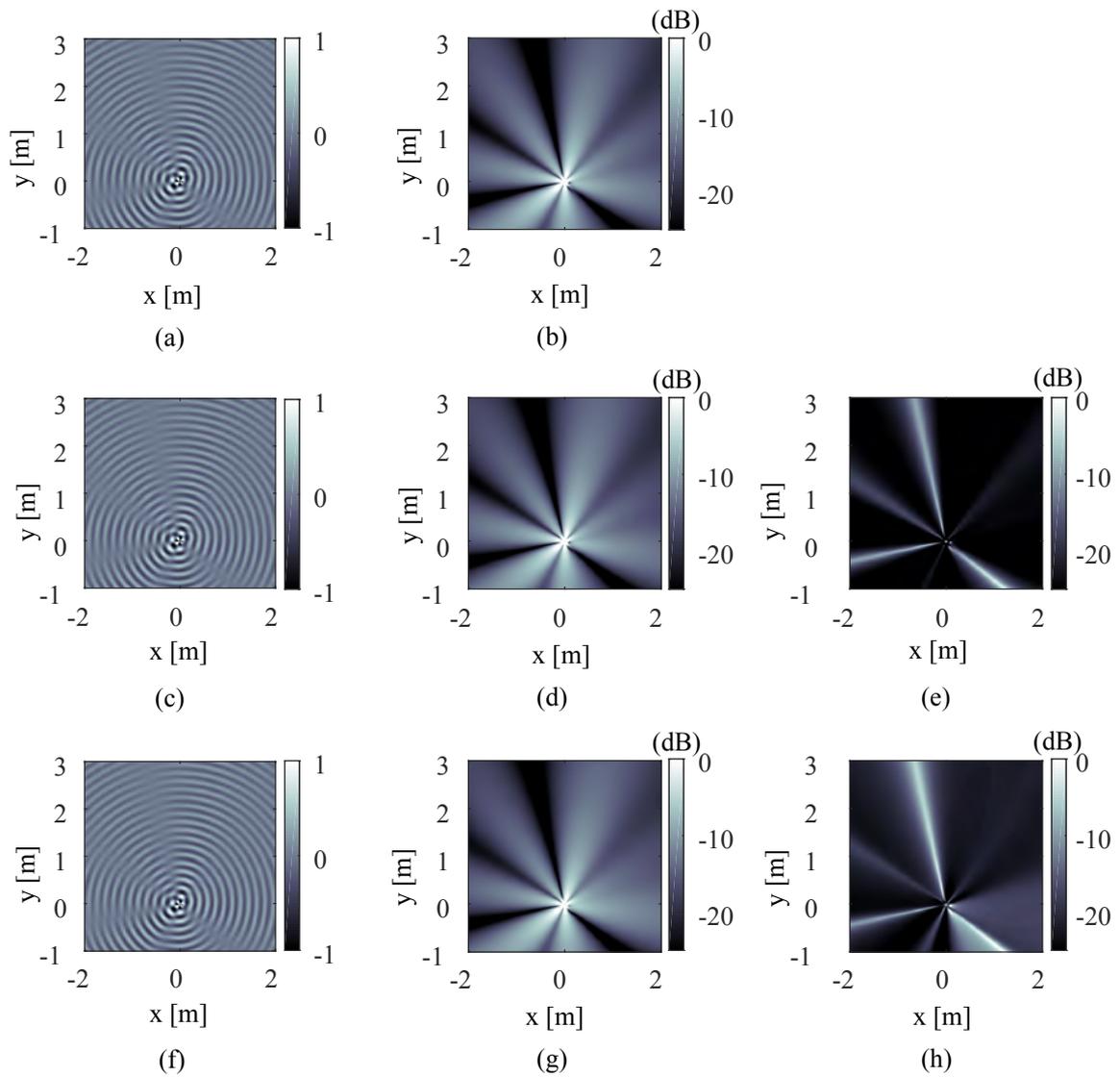


図 5.2: 左列：目標音場 (a) と提案手法および pressure-matching method の音場. 中央列：各手法の放射パワー. 右列：式 (2.77) による目標音場と各手法の音場との誤差分布. 上段：目標音源. 中段：提案手法. 下段：pressure-matching method.

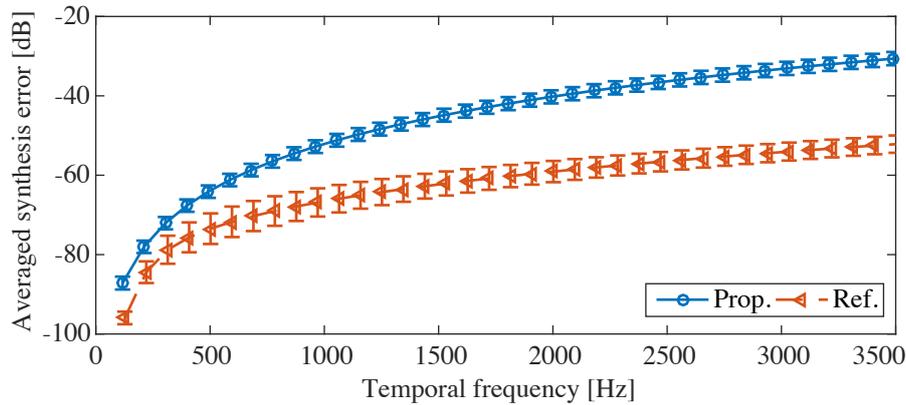


図 5.3: 音場の評価尺度について周波数ごとに算出した平均値. 'Ref.': Pressure-matching method で計算した重み係数を用いた場合. 'Prop.': 提案手法で計算した重み係数を用いた場合.

音場の誤差評価尺度について算出した領域全体に関する平均値を, 空間エイリアシング周波数である 3500 Hz まで算出した結果を, 図 5.3 に示す. なお, 評価に用いる音源の指向特性に対する依存性を排するため, $2N + 1 = 9$ 個の乱数を 100 パターン生成し, 最大次数 $N = 4$ の円調和級数 $\check{S}_v^{(2)}(\omega)$ として用いた. 乱数は, 振幅 1.0 以下の複素数で生成した. 100 パターンの計算結果の平均値を 95 % 信頼度区間とともに図 5.3 に示す. なお, 95 % 信頼度区間の算出においては student-t 分布を仮定した. 図からわかるとおり, 提案手法よりも pressure-matching method の方が高い再現精度を実現することがわかる. 提案手法は, 円調和展開級数の解析的な変換の導出にあたり, ハンケル関数の漸近近似などを用いているが, この近似誤差が音場の再現精度に影響していると考えられる. しかし, pressure-matching method では, 重み係数の算出精度が正則化パラメータ λ に依存するため, 精度のよい重み係数を求めるためにしばしば試行錯誤が必要になることがある. 一方, 提案手法である円調和展開級数の解析的な変換を用いれば, 指向性音源の音場をモデル化した円調和展開級数から高速に重み係数を決定することができる.

5.4 性能評価実験

焦点音源で生成した仮想の多重極音源に、円調和展開級数から解析的変換で得られた重み係数をかけて重ね合わせることで、所望の指向性音源の音場を再現できることを示す。また、提案手法を用いた場合の音場の再現精度を、pressure-matching methodで重み係数を求めた場合の音場の再現精度、および2.3節の従来手法を用いた場合の音場の再現精度と比較する。なお、ここでは各手法で算出される重み係数を用いた場合の音場精度に焦点を当てるため、時間領域での駆動関数については評価しない。

5.4.1 実験条件

本章で実施するシミュレーションは、本項にまとめた実験条件で行う。各実験の項においては本項で述べる条件からの差分のみを陽に記載する。なお、直線アレイの配置などの条件は、4.5.1項と同じであるが、念のため再度記載する。

シミュレーションはMATLABにより実装した。音速は、343.36 m/sとした。また、 x 軸上の $-1.6 \leq x \leq 1.6$ の範囲に0.05 m間隔で65個のスピーカを直線状に並べてスピーカアレイを構成した。焦点音源で構成する仮想の多重極音源に含まれるモノポール音源同士の間隔は0.01 mとした。仮想の多重極音源の中心は、 $(x_c, y_c) = (0.0, 0.5)$ とした。焦点音源の座標は多重極音源の中心座標と4.5.1項に記載の相対座標から容易に計算できる。また、目標となる音場及び指向特性は最大次数 $N = 4$ の円調和級数 $\check{S}_\nu^{(2)}(\omega)$ でモデル化することとし、 $2N + 1 = 9$ 個の乱数で生成した。当該乱数は、振幅1.0以下の複素数で生成した。なお、目標音場は式(2.71)で算出した。

5.4.2 客観評価尺度

提案手法の性能を評価するため、 $-1.5 \text{ m} \leq x \leq 1.5 \text{ m}$, $1.0 \text{ m} \leq y \leq 3.0 \text{ m}$ の領域において、0.01 m間隔の格子状に配置した観測点で式(2.77)の音場の再現精度を計算した。

5.4.3 音場と指向特性の再現精度

指向特性、音場の両方について再現精度を評価するため単一周波数に対する音場の再現実験を行った。シミュレーションには、1992.2 Hzの正弦波を用いた。なお、この周波数は1024点の離散フーリエ変換を行う際に、2000 Hzにもっとも近い周波数であったため採用した。指向性音源が回転した際の再現精度を比較するため、指向性音源が時計方向に 20° 回転した際の音場についても生成して比較を行った。音源に回転を加えなかった際の音場の再現結果を図5.4に、音源を時計方向に 20° 回転した際の

再現結果を図 5.5 に示す. 図 5.4 (a) は目標となる指向性音源が作る音場を表し, (b), (c) は pressure-matching method で求めた重み係数を多重極音源の重ね合わせに用いた音場と音場の誤差評価尺度を示す. また, 図 5.4 (d), (e) は, 本節の提案手法で多重極音源の重み係数を算出し, 多重極音源の重ね合わせに用いた際の再現音場と音場に関する誤差分布である. (f), (g) は, 従来法による再現音場と音場の誤差分布を示す. いずれの手法も同程度に高い再現精度を達成している. 次に音源を時計方向に 20° 回転した場合についての結果を図 5.5 に示す. 図より明らかなように, 音源が回転した場合も 3 手法ともに高い精度を達成していることがわかる.

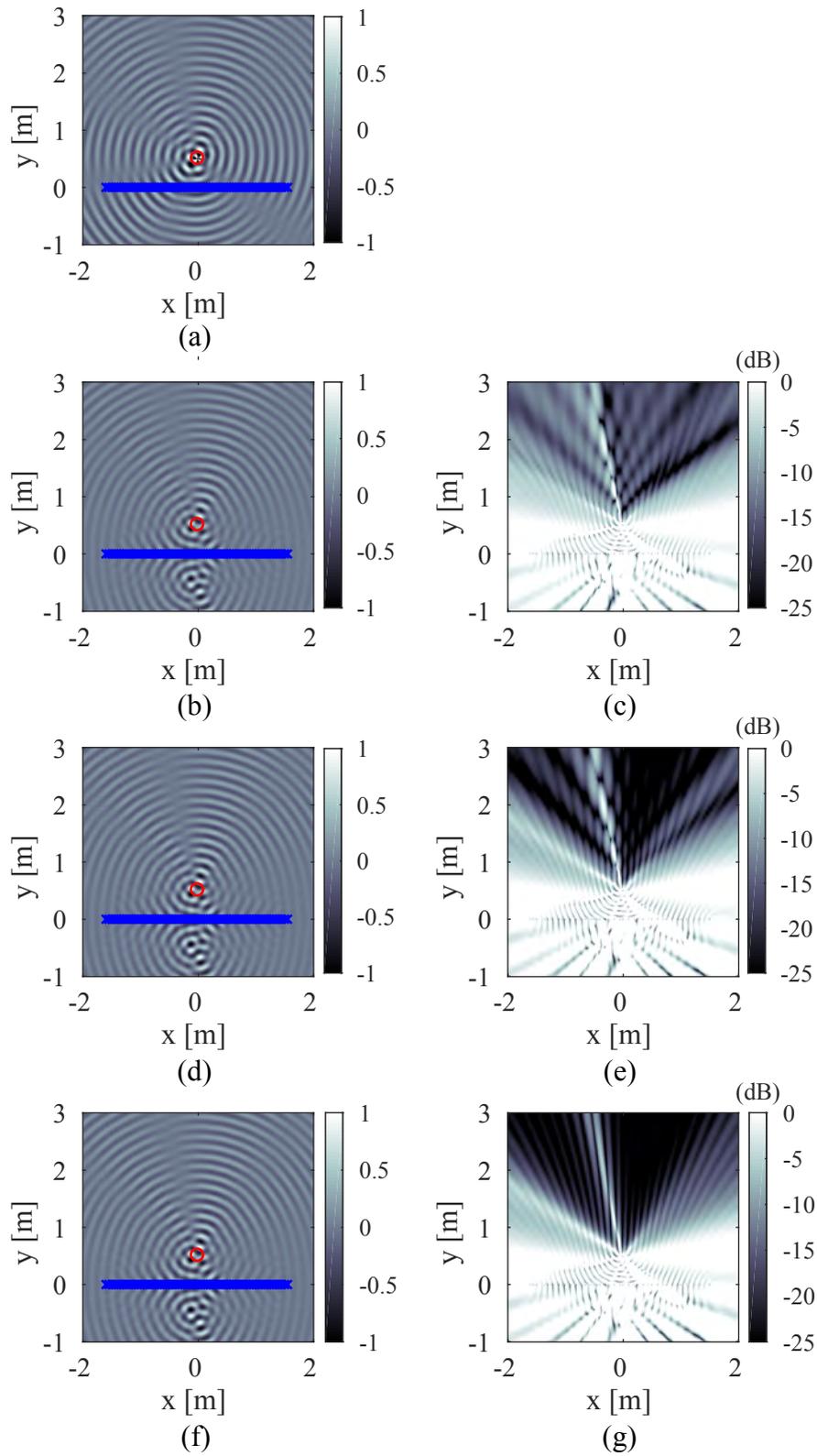


図 5.4: 左列:再現された音場. 右列:音場の誤差評価尺度. (上から) 1 段目:目標音場. 2 段目: pressure-matching method で重み係数を求めた多重極音源の重ね合わせ. 3 段目:提案手法で重み係数を求めた多重極音源の重ね合わせ. 4 段目:従来手法による指向性音源の音場再現.

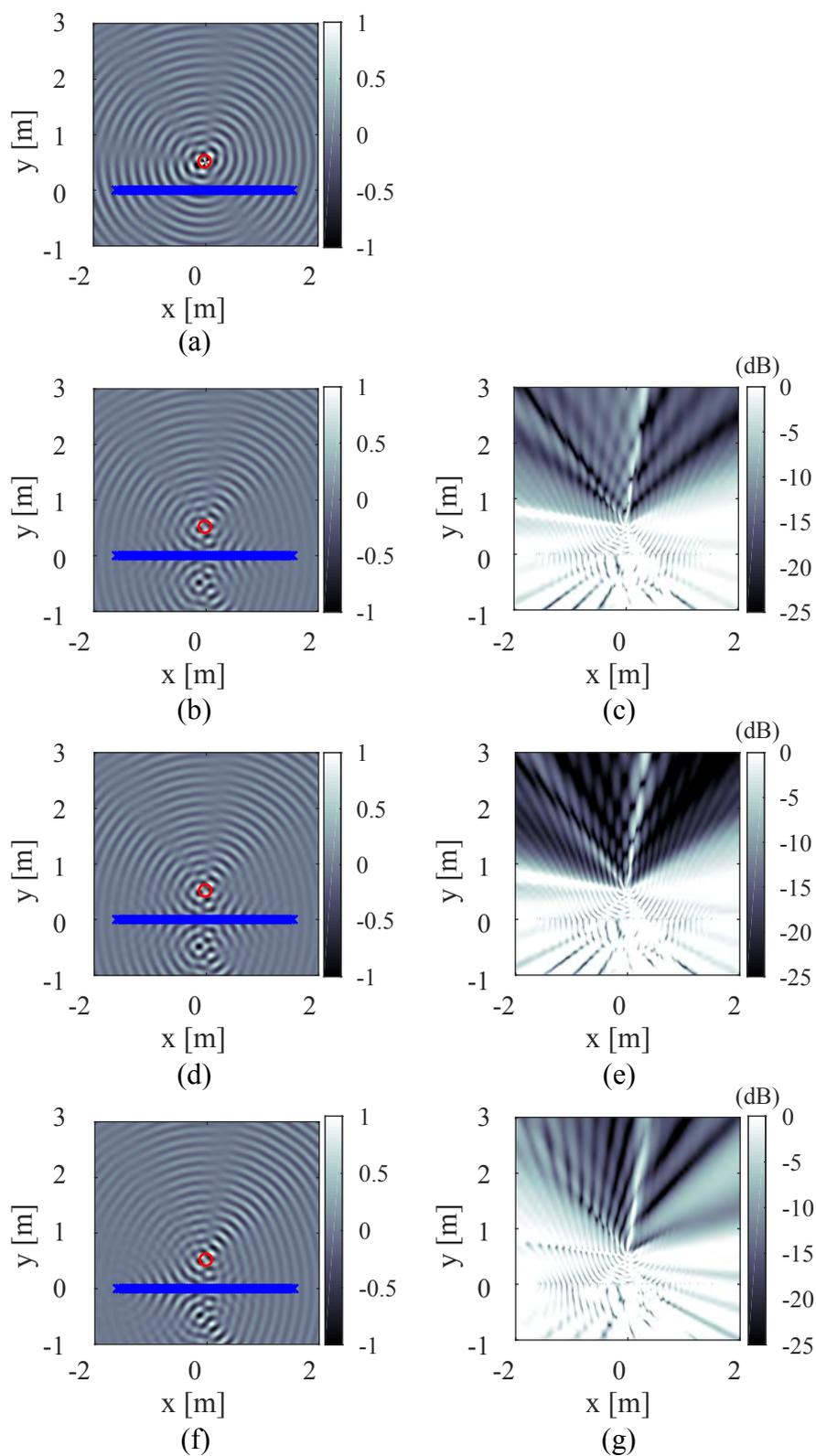


図 5.5: 指向性音源を時計回りに回転 ($\phi = 20^\circ$) したときに再現された音場と、音場の誤差評価尺度. 左列: 再現された音場. 右列: 音場の誤差評価尺度. (上から) 1 段目: 目標音場. 2 段目: pressure-matching method で重み係数を求めた多重極音源の重ね合わせ. 3 段: 提案手法で重み係数を求めた多重極音源の重ね合わせ. 4 段: 従来手法による指向性音源の音場再現.

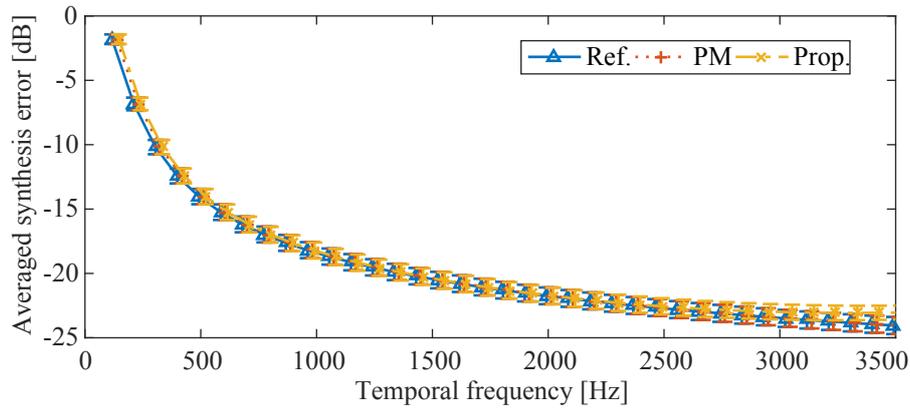


図 5.6: 音場の評価尺度について周波数ごとに算出した平均値. 'Ref.': 従来手法の周波数領域実装 [65, 68]. 'PM': Pressure-matching method で重み係数を算出した場合. 'FD2': 提案手法の周波数領域実装.

音場の誤差評価尺度について算出した領域全体に関する平均値を, 空間エイリアシング周波数である 3500 Hz まで算出した結果を, 図 5.6 に示す. なお, 評価に用いる音源の指向特性に対する依存性を排するため, $2N + 1 = 9$ 個の乱数を 100 パターン生成し, 最大次数 $N = 4$ の円調和級数 $\check{S}_v^{(2)}(\omega)$ として用いた. 当該乱数は, 振幅 1.0 以下の複素数で生成した. また, 音源の平行移動が音場の再現精度におよぼす影響について確認するため, 仮想音源の位置を $-0.5 \text{ m} \leq x \leq 0.5 \text{ m}$, $0.25 \text{ m} \leq y \leq 0.75 \text{ m}$ の範囲でランダムに選択して, 上記 100 パターンの円調和展開級数のそれぞれに対応させた. さらに, 音源の回転が音場の再現精度におよぼす影響について確認するため, 仮想音源の向きを $-30^\circ \leq \phi \leq 30^\circ$ の範囲でランダムに選択して, 上記 100 パターンの円調和展開級数のそれぞれに対応させた. 100 パターンの計算結果の平均値を 95 % 信頼度区間とともに図 5.6 に示す. なお, 95 % 信頼度区間の算出においては student-t 分布を仮定した.

結果のグラフから, 高周波数帯域になるにしたがい提案手法と従来手法の誤差が広がっていくが, 最大でも誤差の平均値の差は 1 dB であった. また, 重み係数を pressure-matching method で算出して多重極音源を生成した場合と従来手法の音場の再現精度を比較すると, 誤差の平均値の差が 0.01 dB 未満であり, ほぼ差がないことが確認できた.

5.5 本章のまとめ

本章では、まず直線状スピーカアレイを用いた指向性音場再現について、3章で提案した円調和級数展開に基づく手法と4章で提案した多重極音源に基づく手法の比較を行った。この比較により、リアルタイムアプリケーションなど演算量の削減が求められるユースケースでは、円調和級数展開に基づく手法を時間領域で実装して用いるのが良いことを確認した。しかし、円調和級数展開に基づく手法は高次ベッセル関数の零点由来の過剰なゲインが現れる可能性があるため実装に注意が必要である。そのため、リアルタイム性が求められないケースでは多重極展開に基づく手法あるいは従来手法を周波数領域で実装するのが適切であることを確認した。

次に、これまで関係性が指摘されつつも陽に定式化されることがなかった円調和展開係数と多重極展開の重み係数の間に解析的な関係が定義できることを2とおりの係数比較を用いて示し、これら解析的変換と多重極音源の重ね合わせに基づき直線状スピーカアレイで再現された仮想音源の指向性が所望のパターンを再現することを確認した。この解析的変換を用いて再現された音場は、pressure-matching method で算出した重み係数を多重極音源の重ね合わせに用いる手法や従来手法と同等の再現精度を達成することを確認した。

5.5.1 本章の貢献と関連研究

本章の内容は、研究業績リスト [C-2], [C-3] の内容をまとめたものである。本章の貢献は、従来研究で関連が指摘されながらも陽に定式化されてこなかった円調和展開級数と多重極展開の係数の関係を、解析的に示した点にある。指向性音源が生成する音場は、高速フーリエ変換により、円調和展開級数を用いて高速にモデル化できる。一方、同じ音場を多重極展開でモデル化するためには、pressure-matching method のような最適化に基づいて多重極展開の係数を算出する必要があった。pressure-matching method のような手法を用いて多重極展開の係数を算出するためには、最小2乗法を用いるしかなく、またこの際、正則化パラメータを適切に設定する必要があるが、実装に試行錯誤を要する上に、正則化パラメータの設定次第では元の音源の指向性を再現するために適切な多重極展開の係数を求めることができない場合があった。このような課題に対し、本章で研究した円調和展開級数の解析的変換を用いることで、正則化パラメータに依存することなく適切な多重極展開の係数を算出することが可能となった。解析的変換の過程でいくつかの近似を導入したため、pressure-matching method を用いてチューニングを行った結果と比較すると精度は劣るものの、音場の再現誤差

は -30 dB 以下に抑えられており, 本手法は十分な精度を達成していると言える.

第6章 結論

本研究は、歌舞伎などの演舞場においてマイクロフォンなどで収録した音響信号を、音源の位置や音源が持つ方向別の強弱のパターンである指向特性を含めて遠隔地のサテライト会場で再現する空間音響技術に関する研究である。本研究では空間音響技術の中でも音像定位技術に注目し、複数人で視聴するスタイルを前提とした場合の音像定位技術に焦点を当てた。個人で視聴する場合と大勢で視聴する場合の大きな違いは、別の観客との会話や別会場で視聴する観客との間の歓声等の共有であるとの視点に立ち、受聴者の近くまで飛び出る音像である焦点音源を生成可能な波面合成技術を基礎として検討を進めた。

また、楽器や人間の発声器官といった音源が持つ指向特性の再現と受聴者が感じる原音らしさの間には関連があることも指摘されていることから、再現性の高い音像定位技術の実現を目指し、前後左右といった音源の位置情報のみならず、音源が持つ指向特性を再現する波面合成技術を検討した。従来、複数のスピーカからなるスピーカアレイを用いた指向特性の再現には、円調和級数展開などの空間フーリエ変換に基づく手法や、直交格子状に密に並んだ音源が生成する多重極音源の重み付き和である多重極展開に基づく手法などが検討されてきた。これらの手法では、スピーカアレイを中心として、その周囲に指向性を再現することが前提であった。本研究では、実スピーカアレイで生成した複数の焦点音源を新たな仮想スピーカとみなし、この仮想スピーカ群に対して、円調和展開や多重極展開に基づく指向性制御技術を適用することで、指向特性を有する仮想の音源を、スピーカよりも前面に生成する検討を行った。

1章では、研究の背景、従来技術、および研究の位置づけについて述べた。

2章では、波面合成技術と指向性制御技術について、本論文と関わりの深い焦点音源法、空間フーリエ変換、多重極音源について述べた。特にこれらを直線状スピーカアレイや円状スピーカアレイを用いて実現するために各スピーカに適用するフィルタの導出を確認した。また、周波数領域で定義されたこれらのフィルタに近似を加えることで、低演算量で動作する時間領域のフィルタを導出できることを確認した。

3章では、直線状スピーカアレイで生成した複数の焦点音源を制御することで、これらの組を仮想の円状スピーカアレイとして扱う手法を研究した。従来の波面合成技術を用いて指向性音源の音場を合成した場合、音源の回転や平行移動に対してもフィ

ルタの再計算が必要となるため演算面で冗長な部分があった。また、演算量を削減するための時間領域のフィルタも導出されていたが、あらかじめ算出されたフィルタの補間処理により目的のフィルタを算出するため、目的の音場を再現することも困難であった。そこで3章では、複数の焦点音源からなる仮想の円状スピーカアレイを生成するための駆動関数と、仮想スピーカアレイの制御により指向特性の形状を生成するための駆動関数を独立に算出し、これらを結合して得られる駆動関数により指向性音源が作る音場を再現する手法を提案した。このような駆動関数を用いることで、指向性音源の平行移動や回転を仮想スピーカアレイの生成に関わる駆動関数の更新のみで実現できるようになり、また仮想音源の平行移動や回転の際にも高い再現精度を維持できることを確認した。仮想の円状アレイを用いることで、再現したい円調和展開級数の次数に比例する個数の焦点音源のみで従来手法と同等の再現精度が達成できることを確認した。本章の貢献は、直線状アレイで生成した焦点音源を用いてスピーカ前面に指向性音源が作る音場を再現できることを示したところであり、円調和展開級数の最大次数が6次以下のとき、提案手法が従来手法よりも低演算量かつ高精度に音場を表現できることを示したところである。

4章では、直線状スピーカアレイで生成した複数の焦点音源を制御することで、これらの組を多重極音源として扱う手法を研究した。3章でも述べたとおり、従来の波面合成技術を用いた指向性音源の音場合成では、音源の回転や平行移動について演算面で冗長な部分があった。そこで4章においても4章と同様に、複数の焦点音源からなる仮想のスピーカアレイを生成するための駆動関数と、仮想スピーカアレイの制御により指向特性の形状を生成するための駆動関数を独立に算出し、これらを結合して得られる駆動関数により指向性音源が作る音場を再現する手法を提案した。ただし、仮想のスピーカアレイとして円形アレイではなく、多重極音源を検討した。多重極音源を実現するためには、一般的に多数のモノポール音源が必要となるため、多重極音源の次数が増加した際の演算量の増加が本章で提案した手法の課題となる。そこで、隣り合う低次の多重極音源の一部を重ねた効率的構造を持つ多重極音源を検討し、2次元かつ一般の次元を持つ多重極音源の中のモノポール音源の位置と、それらに与えるべき重みを定式化した。さらに時間領域の駆動関数を導出し、3次の多重極音源の重ね合わせまでは従来手法よりも高精度かつ低演算量に動作可能なことを確認した。本章の貢献は、直線状アレイで生成した焦点音源を用いてスピーカ前面に指向性音源が作る音場を再現できることを示したところであり、提案手法は仮想の指向性音源の平行移動や回転を効率的かつ高精度に表現できることを示したところである。また本章

では、これまで陽に定式化されてこなかった、一般の次数の多重極音源の構成法について定式化したことで、2次元平面上の任意の次数の多重極展開を実現する手法を示した。

5章では、3章及び4章で提案した2つの手法の比較をまずは行った。その結果、ライブイベントなどのようにリアルタイム動作が要求されるアプリケーションにおいては、円調和級数展開に基づく3章の手法を時間領域で実装して用いることで仮想音源の指向性制御を低演算量で実現できるといえる。一方、リアルタイム性が要求されないアプリケーションにおいては、多重極展開に基づく4章の手法を周波数領域で実装して用いることで、指向特性を高精度に再現することができる。次に、これまで密な関係があることが指摘されながらも陽に定式化されることがなかった円調和展開級数と多重極展開級数の間の解析的な関係を研究した。この解析的な関係を2とおりの表現手法を用いて研究した。1つめは、多重極音源を原点に配置した際に単位円上で観測される音圧を原点周りのテイラー展開により表現したものと、多重極音源が単位円上で作る音場を比較して得られる変換式である。この式は、放射される音の方向別の強弱は再現可能なものの、単位円上のみでの係数比較となっているため、音場再現対象となる領域全体をみると再現音場の位相が精度よく再現されない課題があった。これに対し2つめの表現手法では、2次元グリーン関数の高次微分に現れる多重極の係数と、円調和展開級数に含まれる複素正弦波を二項展開することで現れる多重極の係数を比較すること解析的な変換式を導出した。この変換式は、1つめの変換式と異なり、音源から遠方の領域全体で成り立つため音場の位相も再現できることを確認した。このようにして求めた重み係数を4章に示した仮想の多重極音源に乗じて重ね合わせることにより、円調和展開級数でモデル化された指向性音源の音場を、仮想の多重極音源により再現できることを示した。本章の貢献は、これまで関連性が指摘されてきた円調和展開級数と多重極展開級数との関係性を考察し、円調和展開級数を直接、多重極展開係数に変換する式を導出したところにある。これまで多重極展開級数は、pressure-matching methodのように最小2乗法など正則化パラメータなどへの依存性が高い手法を用いており、正則化パラメータの選択次第では音場の再現精度が大きく低下してしまう課題があった。これに対し、本章で提案する変換手法を用いれば安定した再現精度を実現可能な重み係数を得ることができる。また、円調和展開級数は円状マイクアレイなどを用いて高速に算出できることも知られている。本章で提案した手法により、円状マイクアレイで得た円調和展開級数から多重極音源により音場を再現することもできるようになる。一般に、円状スピーカアレイよりも多重極

音源の方が小型化が可能であることが知られているため、多重極音源を用いた小型のパーソナルオーディオなどへの応用が期待できる。

本論文では、あたかも演舞場で観劇しているかのような再現性の高い音響再生を目指し、主に直線状スピーカアレイを用いて指向特性を有する仮想音源をスピーカアレイよりも前面に生成する技術について研究した。比較的大きな会場において、複数人で同じコンテンツを楽しむためのシステムを想定した技術に焦点を当てたが、今後はこういった音響コンテンツを自宅などで楽しんだり、大きな会場であっても客席ごとに呈示する音響コンテンツを変更したいといった要望が生まれることが予想される。本論文で研究した音像定位技術の本質は、スピーカアレイを用いた指向性制御技術であり、音を聞かせたいエリアと音を聞かせないエリアを生成するといった応用にも適用可能な技術である。

付録A 波面合成技術を活用した パブリックビューイング

NTT サービスエボリューション研究所では、競技空間をまるごとリアルタイムに日本国内はもとより世界へ配信するイマーシブテレプレゼンス技術「Kirari!」の研究開発を行っている。Kirari! は、時間的に同期伝送された競技映像・音声をホログラフィックディスプレイに投影・再生することで、あたかも選手が上映会場で競技を行っているかのような高い臨場感のパブリックビューイングを実現する。サッカーボールを蹴る音や柔道選手が受け身を取るときの音といった競技音を再生する際、スクリーンに投影された選手の映像と時間的・空間的に一致する点から音を鳴らすことで、単なるステレオ再生などよりもはるかに高い臨場感が期待できる。このように再生する音に空間を移動する特殊効果を与える音像定位技術は、高い臨場感や迫力が要求される映画やゲームなどで広く用いられており、音響で高い臨場感を引き出すために非常に重要である。そこで、NTT は、奥行き方向の音像制御も実現可能な波面合成音響技術を用いて、音が客席近くまで飛び出る音像定位技術を実現した。波面合成音響技術は Kirari! に組み込まれ、NTT 西日本グループコレクション 2016, R & D フォーラム 2017, 歌舞伎シアターバーチャル座, ニコニコ超歌舞伎 2017, 科学技術館常設展示（サイバーリンク）, といった実フィールドで多数採用された。本章では、フィールドトリアルの中でもニコニコ超歌舞伎 2017 に出展した際に用いたシステムについて述べる。

A.1 Kirari! と波面合成技術

これまで NTT で検討していた波面合成技術は、マイクアレイで収録した音場をスピーカアレイで高精度に再現する波面再構成技術であった。音源の数や位置を検出することなく、収録した音进行处理して再生することで収録会場の音場をそのまま上映会場で再現できる優れた技術であるものの、マイクアレイを被写体の近くに配置して収録する必要があるため、同時に収録している映像への写り込みを考慮すると実際の中継に用いることは困難であった。また、個々の音源を分離するのが難しく、特定の音に限定して編集を加えるといった柔軟性にかかる課題もあった。そこでコンテンツに演出や編集が加わる可能性を考慮し、モノラルの音源から仮想的な音源を作り出す波面

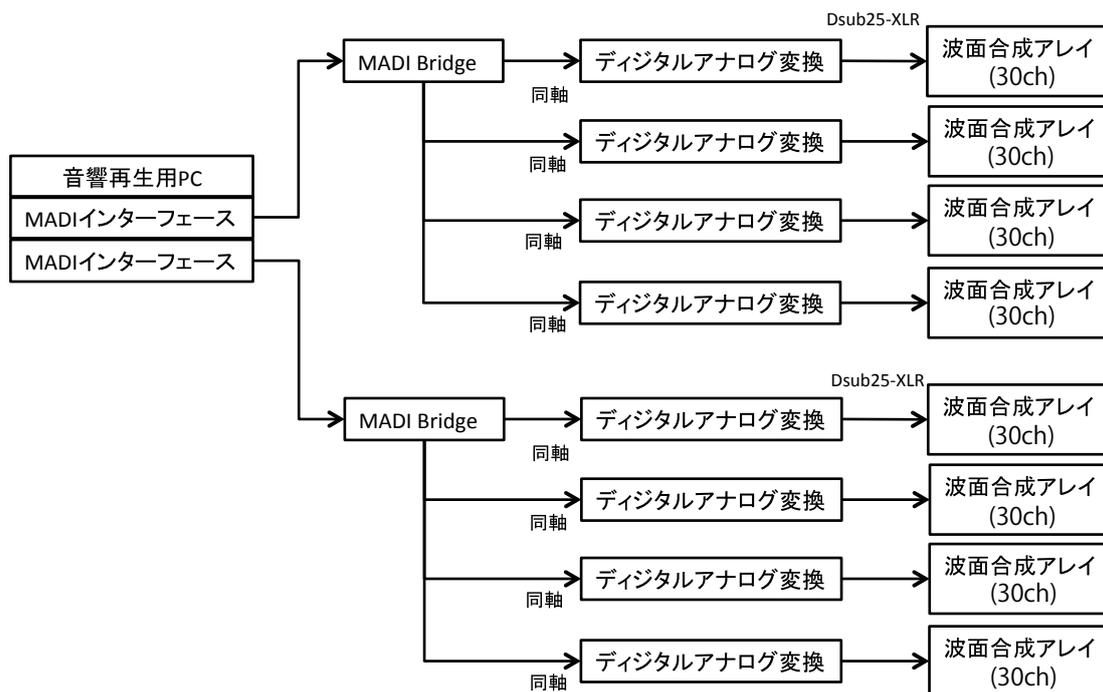


図 A.1: システム構成図

合成技術を検討した。通常の収録でも、ガンマイクなどを用いて個別の音源を高品質に収録した上で、編集を加えてコンテンツを制作しており、モノラル音源に対して音像定位効果を付与する手法は現実のワークフローに合わせた手法となっている。

A.2 システム構成

ニコニコ超歌舞伎 2017 では、事前に波面合成の音響効果を付与した効果音を音響再生用ソフトウェアに準備しておき、役者の演技に合わせてオペレータが操作することで効果音を再生する方式とした。システム全体の構成を図 A.1 に示す。ワークステーション (Mac Pro) に 2 台並列接続したオーディオインターフェース (RME 社 MADIFace XT) から MADI 形式にフォーマットされた音声のデジタルデータを送り、デジタルアナログ変換 (DirectOut 社 Andiamo 2 DA) でアナログ信号に変換された出力音声信号を、アンプ内蔵のスピーカ (Genelec 社 Genelec 8020A) を並べて構成した波面合成アレイから再生することにより、波面合成の効果が付与された効果音がされる。波面合成技術は、2.1 節でも述べた通り、理論上無限に長い連続な 2 次音源を前提とするが、実際には市販のスピーカを用いて有限長かつ離散的な 2 次音源を用いざるをえない。さらに、消防法などの制約条件の下でスピーカの設置場所を決定することから、直線の途中で通路分の隙間を開ける必要があることも少なくない。

その結果、観客が音像を知覚する位置と事前に設定した焦点音源の位置が一致するスイートスポットが狭くなり、飛び出す音響の効果を同時に体験できる客席数が少なくなる課題がある。このスイートスポットの課題に対し、直線状アレイの一部を曲げて折れ線アレイ [113] とし、線分ごとに独立なアレイとして焦点音源を生成する手法を導入することで、スイートスポットを拡大する設計とした [114]。直線アレイを用いて生成した焦点音源の音場を図 A.2 に、提案手法の折れ線アレイで生成した焦点音源の音場を図 A.3 に示す。点線で囲んだエリアを比較すると、直線アレイではスイートスポットから外れてモノポール音源の波面が形成されていないが、提案手法の折れ線アレイでは所望の波面が形成されていることがわかる。

このような折れ線アレイによる波面合成システムを幕張メッセのアリーナにおいて構築し歌舞伎の演目の公演を行った。図 A.4 に示すとおり、システムに用いたスピーカ数は 240 チャンネル（フロント 80 チャンネル、リア 80 チャンネル、レフト・ライトそれぞれ 40 チャンネル）、全長 36m のスピーカアレイとなった。幕張メッセのアリーナには本来、5000 名程度の観客を収容することができるが、スピーカ数の制約上、1 階席 350 名分の客席のみを対象に波面合成を用いた特殊効果を呈示した。石灯籠が爆発して龍が飛び出すシーンでは爆発で飛び散った石灯籠のかけらが客席に降り注ぐ様子を、大詰の火事のシーンでは焼け落ちた家屋が客席近くに落下する様子を再現するなど、様々な演出に波面合成音響を利用した。落下する瓦礫や石を観客の耳元でならずことにより、これまでにない高い臨場感を実現することができたと考える。

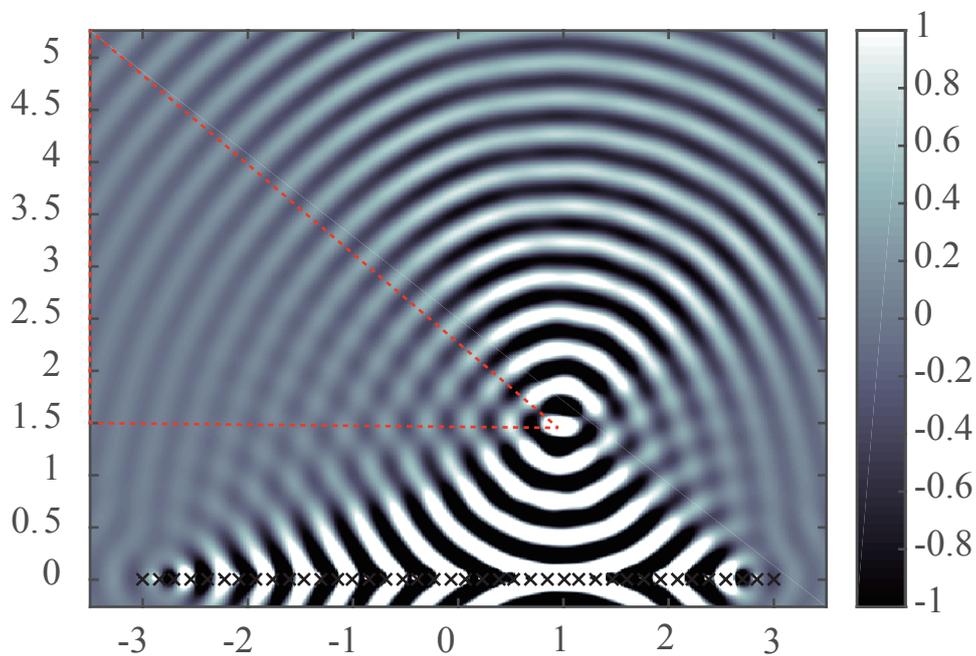


図 A.2: 焦点音源法で生成した音場

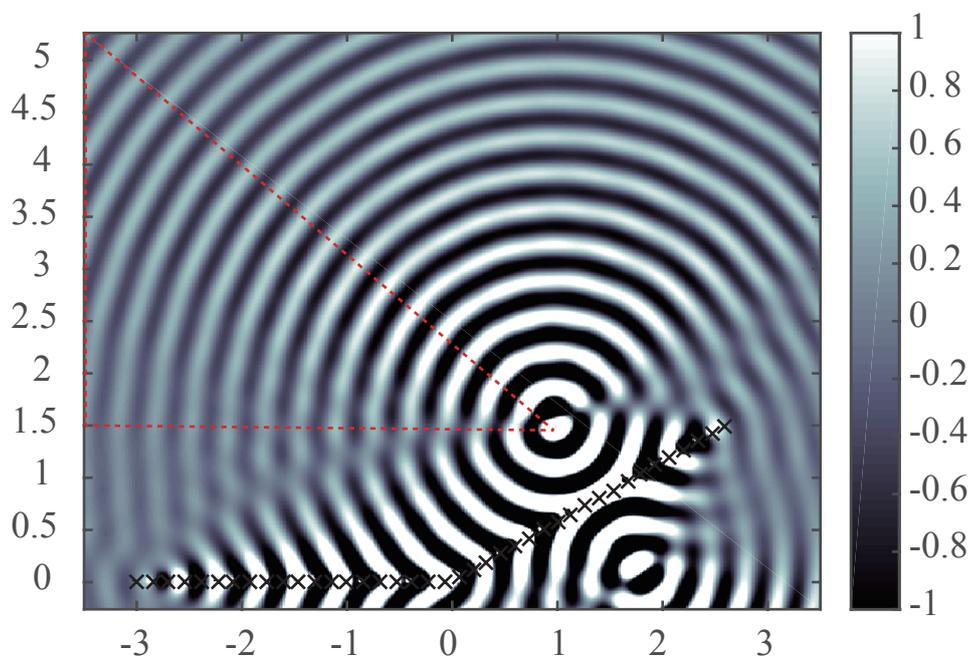


図 A.3: 折れ線アレイを用いて生成した焦点音源の音場

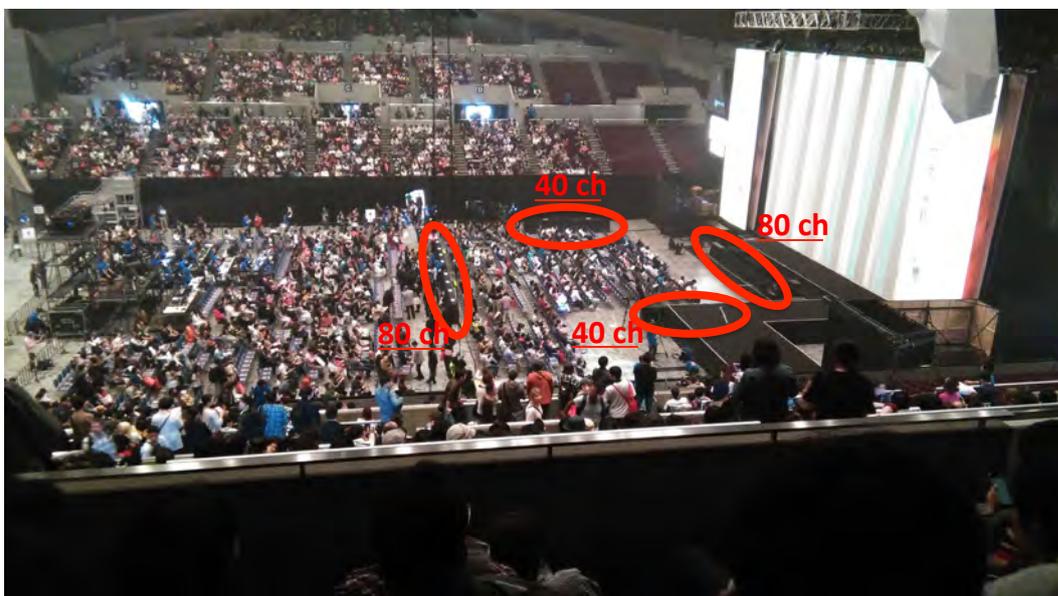


図 A.4: 会場でのシステムレイアウト

参考文献

- [1] F. Rumsey. *Spatial Audio*. Focal Press, 2001.
- [2] H. A. M. Clark, G. F. Dutton, and P. B. Vanderlyn. The 'stereosonic' recording and reproducing system: A two-channel systems for domestic tape records. *Journal of the Audio Engineering Society*, Vol. 6, No. 2, pp. 102–117, 1958.
- [3] Robert Charles Alexander. *The Inventor of Stereo The life and works of Alan Dower Blumlein*. Focal Press, 1999.
- [4] 平原達也, 大谷真, 戸嶋巖樹. 頭部伝達関数の計測とバイノーラル再生にかかわる諸問題. 電子情報通信学会 基礎・境界ソサイエティ Fundamentals Review, Vol. 2, No. 4, pp. 68–85, 2009.
- [5] A. J. Berkhout, D. de Vries, and P. Vogel. Acoustic control by wave field synthesis. *The Journal of the Acoustical Society of America*, Vol. 93, No. 5, pp. 2764–2778, 1993.
- [6] Jens Ahrens, R. Rabenstein, and Sacha Spors. The theory of wave field synthesis revisited. *Audio Engineering Society Convention 124*, 2008.
- [7] 勝本道哲, 山肩洋子, 木村敏幸. 異なる放射指向性を持つスピーカによる臨場感に関する聴覚実験. 聴覚研究会資料, Vol. 40, No. 1, pp. 13–16, Jan 2010.
- [8] 牧勝弘, 木村敏幸, 勝本道哲. 26チャンネル球形スピーカによる楽器の実演奏音の再生: 楽器の放射パターンの再現性に関する音響比較. 日本音響学会誌, Vol. 67, No. 10, pp. 447–458, 2011.
- [9] T. du Moncel. The telephone at the paris opera. *Scientific American*, Vol. 45, No. 27, pp. 422–423, 12 1881.

- [10] E Torick. Highlights in the history of multichannel sound. *Journal of the Audio Engineering Society*, Vol. 46, pp. 27–31, 01 1998.
- [11] S. Spors, H. Wierstorf, A. Raake, F. Melchior, M. Frank, and F. Zotter. Spatial sound with loudspeakers and its perception: A review of the current state. *Proceedings of the IEEE*, Vol. 101, No. 9, pp. 1920–1938, Sep. 2013.
- [12] J. Ahrens. *Analytic Methods of Sound Field Synthesis*. T-Labs Series in Telecommunication Services. Springer Berlin Heidelberg, 2012.
- [13] 安藤彰男. 高臨場感音響技術とその理論. 電子情報通信学会 基礎・境界ソサイエティ Fundamentals Review, Vol. 3, No. 4, pp. 33–46, 2010.
- [14] 安藤彰男. 音響の高臨場感技術. 映像情報メディア学会誌, Vol. 66, No. 8, pp. 671–677, aug 2012.
- [15] Otis C. Trimble. Localization of sound in the anterior-posterior and vertical dimensions of auditory space. *British Journal of Psychology*, Vol. 24, No. 3, Jan. 1934.
- [16] Benjamin Bernfeld. Attempts for better understanding of the directional stereophonic listening mechanism. In *Audio Engineering Society Convention 44*, Mar 1973.
- [17] J. Blauert. *Spatial hearing – The psychophysics of human sound localization*. The MIT Press, 1999.
- [18] D. M. Leakey. Some measurements on the effects of interchannel intensity and time differences in two channel sound systems. *The Journal of the Acoustical Society of America*, Vol. 31, No. 7, pp. 977–986, 1959.
- [19] International Telecommunication Union (ITU-R). *Recommendation BS.775: Multichannel stereophonic sound system with and without accompanying picture*. Geneva, 2012.
- [20] Kimio Hamasaki, Koichiro Hiyama, and Reiko Okumura. The 22.2 multichannel sound system and its application. In *Audio Engineering Society Convention 118*, May 2005.

- [21] Ville Pulkki. Virtual sound source positioning using vector base amplitude panning. *Journal of the Audio Engineering Society*, Vol. 45, No. 6, pp. 456–466, 1997.
- [22] UC-Davis. *The CIPIC HRTF Database*. http://interface.cipic.ucdavis.edu/CIL_html/CIL_HRTF_database.htm.
- [23] MIT Media Lab. *HRTF of a KEMAR Dummy-Head*. <http://sound.media.mit.edu/resources/KEMAR.html>.
- [24] IRCAM. *LISTEN HRTF DATABASE*. <http://recherche.ircam.fr/equipes/salles/listen/>.
- [25] 東北大学. HRTF データベース. <http://www.ais.riec.tohoku.ac.jp/lab/db-hrtf/index-j.html>.
- [26] 名古屋大学. HRTF データベース. <http://www.sp.m.is.nagoya-u.ac.jp/HRTF/index-j.html>.
- [27] 岩谷幸雄. バイノーラル収録・再生における課題：ヘッドホンの自由空間等価特性の影響. *日本音響学会誌*, Vol. 61, No. 7, pp. 386–391, 2005.
- [28] Henrik Mller. Fundamentals of binaural technology. *Applied Acoustics*, Vol. 36, No. 3, pp. 171 – 218, 1992.
- [29] Henrik Mller, Dorte Hammershi, Clemen Boje Jensen, and Michael Friis Srensen. Transfer characteristics of headphones measured on human ears. *Journal of the Audio Engineering Society*, Vol. 43, No. 4, pp. 203–217, 1995.
- [30] 植松尚, 柏野牧夫, 平原達也. 頭外音像定位における自発的な頭部回転の影響. *日本音響学会全国研究発表会講演論文集*, pp. 501–502, Oct 2001.
- [31] M. Kato, H. Uematsu, M. Kashino, and T. Hirahara. The effect of head motion on the accuracy of sound localization. *Acoustical Science & Technology*, Vol. 24, No. 5, pp. 315–317, 2003.
- [32] A. Kudo, H. Higuchi, H. Hokari, and S. Shimada. Improved method for accurate sound localization. *Acoustical Science & Technology*, Vol. 27, No. 3, pp. 134–146, 2006.

- [33] 戸嶋巖樹, 青木茂明, 平原達也. 頭部運動を再現するダミーヘッド: テレヘッド. 日本音響学会論文誌, Vol. 61, No. 4, pp. 197–207, 2005.
- [34] Makoto Otani and Shiro Ise. Fast calculation system specialized for head-related transfer function based on boundary element method. *The Journal of the Acoustical Society of America*, Vol. 119, No. 5, pp. 2589–2598, 2006.
- [35] H. Takemoto, P. Mokhtari, H. Kato, R. Nishimura, and K. Iida. A simple pinna model for generating head-related transfer functions in the median plane. In *Proc. 20th International Congress on Acoustics (ICA)*, pp. 1–5, 2010.
- [36] Yōiti Suzuki, Douglas Brungart, Yukio Iwaya, Kazuhiro Iida, Densil Cabrera, and Hiroaki Kato. *Principles and Applications of Spatial Hearing*. WORLD SCIENTIFIC, 2011.
- [37] M. A. Gerzon. Periphony: With-height sound reproduction. *Journal of the Audio Engineering Society*, Vol. 21, No. 1, pp. 2–10, 1973.
- [38] M. A. Poletti. Three-dimensional surround sound systems based on spherical harmonics. *Journal of the Audio Engineering Society*, Vol. 53, No. 11, pp. 1004–1025, 2005.
- [39] Jerome Daniel. Spatial sound encoding including near field effect: Introducing distance coding filters and a viable, new ambisonic format. In *Audio Engineering Society Conference: 23rd International Conference: Signal Processing in Audio Recording and Reproduction*, May 2003.
- [40] Jens Ahrens and Sascha Spors. An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions. *Acta Acustica United With Acustica*, Vol. 94, No. 6, pp. 988–999, 2008.
- [41] Y. J. Wu and T. D. Abhayapala. Theory and design of soundfield reproduction using continuous loudspeaker concept. *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 17, No. 1, pp. 107–116, Jan 2009.
- [42] Filippo Fazi, Philip Nelson, Jens E. Christensen, and Jeongil Seo. Surround system based on three-dimensional sound field reconstruction. In *Audio Engineering Society Convention 125*, Oct 2008.

- [43] Jorge Trevino Lopez, Takuma Okamoto, Yukio Iwaya, and Yôiti Suzuki. High order ambisonic decoding method for irregular loudspeaker arrays. In *Proceedings of 20th International Congress on Acoustics, ICA 2010*, August 2010.
- [44] 岩谷幸雄, 岡本拓磨, トレビーニョホルヘ, 鈴木陽一. 球面調和解析による音場表現 : 高次アンビソニックス技術の可能性. *日本音響学会誌*, Vol. 67, No. 11, pp. 544–549, 2011.
- [45] A.J. Berkhout. A holographic approach to acoustic control. *Journal of the Audio Engineering Society*, Vol. 36, pp. 977–995, 1988.
- [46] 古家賢一, 小山翔一. 6. 波面合成技術の原理と応用. *映像情報メディア学会誌*, Vol. 68, No. 8, pp. 621–624, 2014.
- [47] Sascha Spors and Jens Ahrens. Spatial sampling artifacts of wave field synthesis for the reproduction of virtual point sources, 2009.
- [48] S. Spors and H. Wierstorf and M. Geier and J. Ahrens. Physical and perceptual properties of focused virtual sources in wave field synthesis. In *Audio Engineering Society Convention 127*, Oct 2009.
- [49] H. Wierstorf, A. Raake, M. Geier, and S. Spors. Perception of focused sources in wave field synthesis. *Journal of the Audio Engineering Society*, Vol. 61, No. 1/2, pp. 5–16, 2013.
- [50] Mincheol Shin, Filippo Maria Fazi, Philip Arthur Nelson, and Fabio Casagrande Hirono. Controlled sound field with a dual layer loudspeaker array. *Journal of Sound and Vibration*, Vol. 333, No. 16, pp. 3794–3817, August 2014.
- [51] Filippo Maria Fazi, Mincheol Shin, Ferdinando Olivieri, Simone Fontana, and Yue Lang. Comparison of pressure-matching and mode-matching beamforming for methods for circular loudspeaker arrays. In *Audio Engineering Society Convention 137*, Oct 2014.
- [52] F. M. Fazi and M. Shin and F. Olivieri and S. Fontana. Low frequency performance of circular loudspeaker arrays. In *Audio Engineering Society Convention 138*, May 2015.

- [53] K. Sato and Y. Haneda. Filter design of a circular loudspeaker array considering the three-dimensional directivity patterns reproduced by circular harmonic modes. In *Audio Engineering Society Convention 142*, May 2017.
- [54] E. G. Williams. *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. Elsevier Science, 1999.
- [55] Gary H. Koopmann, Limin Song, and John B. Fahline. A method for computing acoustic fields based on the principle of wave superposition. *The Journal of the Acoustical Society of America*, Vol. 86, No. 6, pp. 2433–2438, 1989.
- [56] J. Stuart Bolton and Thierry A. Beauvilain. Multipole sources for cancellation of radiated sound fields. *The Journal of the Acoustical Society of America*, Vol. 91, No. 4, pp. 2349–2349, 1992.
- [57] A.J. Kempton. The ambiguity of acoustic sources—a possibility for active control? *Journal of Sound and Vibration*, Vol. 48, pp. 475–483, 10 1976.
- [58] Y. Haneda, K. Furuya, and H. Itou. Design of multipole loudspeaker array based on spherical harmonic expansion. In *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 141–144, May 2011.
- [59] 羽田陽一, 古家賢一, 島内末廣. 球調和関数展開に基づく多重極音源を用いた指向性合成. *日本音響学会誌*, Vol. 69, No. 11, pp. 577–588, 2013.
- [60] M. Eichler and A. Lacroix. Broadband superdirective beamforming using multipole superposition. In *2008 16th European Signal Processing Conference*, pp. 1–5, Aug 2008.
- [61] E. Corteel. Synthesis of directional sources using wave field synthesis, possibilities, and limitations. *EURASIP Journal on Advances in Signal Processing*, Vol. 2007, No. 1, p. 090509, Feb 2007.
- [62] S. Koyama, K. Furuya, Y. Hiwasaki, and Y. Haneda. Reproducing virtual sound sources in front of a loudspeaker array using inverse wave propagator. *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 20, No. 6, pp. 1746–1758, Aug 2012.

- [63] S. Koyama, K. Furuya, Y. Hiwasaki, and Y. Haneda. Analytical approach to wave field reconstruction filtering in spatio-temporal frequency domain. *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 21, No. 4, pp. 685–696, April 2013.
- [64] S. Koyama. *Boundary Integral Approach to Sound Field Transform and Reproduction*. PhD thesis, The University of Tokyo, Tokyo, 2013.
- [65] Jens Ahrens and Sacha Spors. Implementation of directional sources in wave field synthesis. *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, pp. 447–458, 2008.
- [66] J. Ahrens and S. Spors. Wave field synthesis of moving virtual sound sources with complex radiation properties. *Journal of the Acoustic Society of America*, vol. 130, no. 5, pp. 2807–2816, 2011.
- [67] J. Ahrens and S. Spors. Rendering of virtual sound sources with arbitrary directivity in higher order ambisonics. *Audio Engineering Society Convention 123*, 2007.
- [68] J. Ahrens and S. Spors. Notes on rendering of focused directional sound sources in wave field synthesis. *Proc. 34rd German Annual Conference on Acoustics (DAGA)*, pp. 2807–2816, 2011.
- [69] T. Okamoto. Angular spectrum decomposition-based 2.5d higher-order spherical harmonic sound field synthesis with a linear loudspeaker array. *Proc. IEEE ICASSP, New Paltz, USA*, 2017.
- [70] Yangfan Liu and J. Stuart Bolton. On the completeness and the linear dependence of the cartesian multipole series in representing the solution to the helmholtz equation. *The Journal of the Acoustical Society of America*, Vol. 140, No. 2, pp. EL149–EL153, 2016.
- [71] William B. Snow. Basic principles of stereophonic sound. *Audio, IRE Transactions on*, Vol. AU-3, pp. 42 – 53, 04 1955.
- [72] K. Tsutsumi and H. Takada. Powerful sound effects at audience seats by wave field synthesis. *NTT Technical Review*, vol. 15, no. 12, 2017.

- [73] Yuki Mitsufuji, Asako Tomura, and Kazunobu Ohkuri. Creating a highly-realistic "acoustic vessel odyssey" using sound field synthesis with 576 loudspeakers. Jul 2018.
- [74] Sacha Spors, H. Wierstorf, M. Geier, and Jens Ahrens. Physical and perceptual properties of focused virtual sources in wave field synthesis. *Audio Engineering Society Convention 127*, 2009.
- [75] E.W. Start. *Direct sound enhancement by wave field synthesis*. PhD thesis, Delft University of Technology, Newark, 1997.
- [76] Sascha Spors and Jens Ahrens. Spatial sampling artifacts of wave field synthesis for the reproduction of virtual point sources. In *Audio Engineering Society Convention 126*, May 2009.
- [77] Diemer de Vries, E. W. Start, and V. G. Valstar. The wave-field synthesis concept applied to sound reinforcement restriction and solutions. In *Audio Engineering Society Convention 96*, Feb 1994.
- [78] Sylvain Yon, Mickael Tanter, and Mathias Fink. Sound focusing in rooms: The time-reversal approach. Vol. 113, No. 3, pp. 1533–43, 2003.
- [79] G.B. Arfken, H.J. Weber, and F.E. Harris. *Mathematical Methods for Physicists*. Elsevier, 2005.
- [80] Alan V. Oppenheim, R. Schaffer, and W. Ronald. *Digital signal processing*. Prentice-Hall Englewood Cliffs, N.J, 1975.
- [81] T. I. Laakso, V. Valimäki, M. Karjalainen, and U. K. Laine. Splitting the unit delay [fir/all pass filters design]. *Signal Processing Magazine, IEEE*, vol.13, no. 1, pp. 30–60, 1996.
- [82] F. Winter and S. Spors. On fractional delay interpolation for local wave field synthesis. *Proc. 24th European Signal Processing Conference*, 2016.
- [83] J. P. Thiran. Recursive digital filters with maximally flat group delay. *IEEE Transactions on Circuit Theory*, vol. 18, no. 6, pp. 659–664, 1971.

- [84] M. Kolundzija and C. Faller and M. Vetterli. Design of a compact cylindrical loudspeaker array for spatial sound reproduction. In *Audio Engineering Society Convention 130*, May 2011.
- [85] Ferdinando Olivieri, Filippo Maria Fazi, Mincheol Shin, and Philip Nelson. Pressure-matching beamforming method for loudspeaker arrays with frequency dependent selection of control points. In *Audio Engineering Society Convention 138*, May 2015.
- [86] H. M. Jones, R. A. Kennedy, and T. D. Abhayapala. On dimensionality of multipath fields: Spatial extent and richness. In *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 3, pp. III-2837–III-2840, May 2002.
- [87] T. D. Abhayapala, T. S. Pollock, and R. A. Kennedy. Characterization of 3d spatial wireless channels. In *2003 IEEE 58th Vehicular Technology Conference. VTC 2003-Fall (IEEE Cat. No.03CH37484)*, Vol. 1, pp. 123–127 Vol.1, Oct 2003.
- [88] R. A. Kennedy, P. Sadeghi, T. D. Abhayapala, and H. M. Jones. Intrinsic limits of dimensionality and richness in random multipath fields. *IEEE Transactions on Signal Processing*, Vol. 55, No. 6, pp. 2542–2556, June 2007.
- [89] T. D. Abhayapala and A. Gupta. Spherical harmonic analysis of wavefields using multiple circular sensor arrays. *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 18, No. 6, pp. 1655–1666, Aug 2010.
- [90] P.A. Nelson and S.J. Elliott. *Active Control of Sound*. Academic Press, 1992.
- [91] K. Tsutsumi, Y. Haneda, K. Noguchi, and H. Takada. Directivity synthesis with multipoles comprising a cluster of focused sources using a linear loudspeaker array. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 496–500, April 2018.
- [92] R. Coifman, V. Rokhlin, and S. Wandzura. The fast multipole method for the wave equation: a pedestrian prescription. *Antennas and Propagation Magazine, IEEE*, Vol. 35, No. 3, pp. 7–12, June 1993.

- [93] L. Greengard, Jingfang Huang, V. Rokhlin, and S. Wandzura. Accelerating fast multipole methods for the helmholtz equation at low frequencies. *IEEE Computational Science and Engineering*, Vol. 5, No. 3, pp. 32–38, July 1998.
- [94] F. Lacava. *Classical Electrodynamics: From Image Charges to the Photon Mass and Magnetic Monopoles*. Undergraduate Lecture Notes in Physics. Springer International Publishing, 2016.
- [95] David J Griffiths. *Introduction to electrodynamics; 4th ed.* Pearson, Boston, MA, 2013. Re-published by Cambridge University Press in 2017.
- [96] I. S. Gradshteyn and I. M. Ryzhik. *Table of integrals, series, and products*. Elsevier/Academic Press, Amsterdam, seventh edition, 2007.
- [97] James L. Flanagan. Analog measurements of sound radiation from the mouth. *The Journal of the Acoustical Society of America*, Vol. 32, No. 12, pp. 1613–1620, 1960.
- [98] W. Chu and A. Warnock. Detailed directivity of sound fields around human talkers. *Research Report (National Research Council of Canada. Institute for Research in Construction)*, pp. 1–47, 01 2002.
- [99] Fabio Bozzoli, Angelo Farina, and Michel Viktorovitch. Balloons of directivity of real and artificial mouth used in determining speech transmission index. In *Audio Engineering Society Convention 118*, May 2005.
- [100] Densil Cabrera, Pamela J. Davis, and Anna Connolly. Long-term horizontal vocal directivity of opera singers: Effects of singing projection and acoustic environment. *Journal of Voice*, Vol. 25, No. 6, pp. 291–303, 2011.
- [101] Felipe Otondo and Jens Rindel. A new method for the radiation representation of musical instruments in auralizations. *Acta Acustica united with Acustica*, Vol. 91, pp. 902–906, 09 2005.
- [102] Pittman A. L. Short-term word-learning rate in children with normal hearing and children with hearing loss in limited and extended high-frequency bandwidths. *Journal of speech, language, and hearing research*, Vol. 51, No. 3, pp. 785–797, 2008.

- [103] B. B. Monson, A. J. Lotto, and S. Ternström. Detection of high-frequency energy changes in sustained vowels produced by singers. *The Journal of the Acoustical Society of America*, Vol. 129, No. 4, pp. 2263–2268, 2011.
- [104] Brian B. Monson, Eric J. Hunter, and Brad H. Story. Horizontal directivity of low- and high-frequency energy in speech and singing. *The Journal of the Acoustical Society of America*, Vol. 132, No. 1, pp. 433–441, 2012.
- [105] C. P. Mathews and M. D. Zoltowski. Eigenstructure techniques for 2-d angle estimation with uniform circular arrays. *IEEE Transactions on Signal Processing*, Vol. 42, No. 9, pp. 2395–2407, Sep. 1994.
- [106] M. Mller and M. Olsen and F. T. Agerkvist and J. Dyreby and G. K. Munch. Circular loudspeaker arrays with controllable directivity. In *Audio Engineering Society Convention 128*, May 2010.
- [107] O. Kirkeby and P. A. Nelson. Reproduction of plane wave sound fields. *The Journal of the Acoustical Society of America*, Vol. 94, No. 5, pp. 2992–3000, 1993.
- [108] J. Ahrens and S. Spors. Local sound field synthesis by virtual secondary sources. In *Audio Engineering Society Conference: 40th International Conference: Spatial Audio: Sense the Sound of Space*, Oct 2010.
- [109] S. Spors and J. Ahrens and K. Helwani. Local sound field synthesis by virtual acoustic scattering and time-reversal. In *Audio Engineering Society Convention 131*, Oct 2011.
- [110] F. Winter and S. Spors. Physical properties of local wave field synthesis using linear loudspeaker arrays. In *Audio Engineering Society Convention 138*, May 2015.
- [111] International Telecommunication Union (ITU-T). *Recommendation G.192: A common digital parallel interface for speech standardization activities*. Geneva, 1996.
- [112] 羽田陽一, 古家賢一, 島内末廣. 球調和関数展開に基づく多重極音源を用いた指向性合成. *日本音響学会誌*, Vol. 69, No. 11, pp. 577–588, 2013.

- [113] E. W. Stuart. Application of curved arrays in wave field synthesis. In *Audio Engineering Society Convention 100*, May 1996.
- [114] 堤公孝, 高田英明. 補助アレイを用いた波面合成法による焦点音源の受聴エリア拡大. 日本音響学会全国研究発表会講演論文集, 2017.

謝辞

本論文は、筆者が日本電信電話株式会社（NTT）サービスエボリューション研究所および電気通信大学で行った3年間の研究成果をまとめたものです。本研究の遂行にあたり、多くの方々から御指導と御支援を頂きました。特に御世話になった皆様をここに記し、深い感謝の意を表します。

指導教官である電気通信大学羽田陽一教授には、本論文のテーマ選定から始まり、日頃の研究遂行の指導、本論文をまとめるにあたっては構成や内容についての丁寧な御指導を賜りました。羽田教授は、筆者がNTTに入社した際の上長であり、駆け出しの新人時代に業務の進め方を手取り足取りご指導いただきました。本論文の研究をスタートさせるきっかけとなったKirari!音響の研究テーマを立ち上げる際には、アカデミックな観点から高臨場感音響技術について様々なご意見と技術面でのご協力をいただきました。数多くのプロジェクトを成功させることができたのは羽田教授のおかげと言っても過言ではないと確信しております。お忙しい中、審査委員を務めていただいた電気通信大学の梶本裕之教授、広田光一教授、高橋裕樹准教授、野村英之准教授からは、本論文の内容や構成に関して有益なご意見を頂きました。

NTTにおいて、研究活動のみならず業務の進め方など、多くをご指導いただいたのが高田英明博士です。入社以来、アカデミックな活動に縁がなかった筆者を研究チームに誘ってくださり、研究者として活躍する機会を与えていただきました。高田博士は、研究活動をとおして会社の課題を解決すること、課題解決の成果を独占せず学術の発展に貢献すること、マネジメントとの適切なコミュニケーションをつうじてリソースを確保することなど、企業研究者としてやるべきことをすべて遂行しておられました。短い期間ではありましたが共に過ごし、その取り組みを手本とすることができたのは、筆者にとってかけがえのない資産となりました。

NTT サービスエボリューション研究所の中平篤博士には、技術の実用化や研究方針について議論をして頂きました。中平博士の流れを読む力と戦略を間近で見ることができ、大変勉強になりました。NTT サービスエボリューション研究所の今泉健太氏には、コア技術の研究から実証実験にいたるまで共に検討いただき、特に技術面でプ

プロジェクトの成功に貢献いただきました。

NTT ドコモ先進技術研究所の菊入圭氏には、エンジニアとしてのあり方をご指導頂きました。技術者としての基礎は菊入氏からのご指導がなければ習得できなかったと確信しております。

NTT サービスエボリューション研究所阿久津明人所長には、本研究にご理解とご支援を頂くと共に、温かい激励を頂きました。NTT サービスエボリューション研究所南憲一企画部長、ナチュラルコミュニケーションプロジェクト嵯峨田淳プロジェクトマネージャ、グループリーダーである佐藤隆博士、日高浩太博士には、上長として本研究にご理解頂き、本研究遂行の機会を与えて頂きました。NTT テクノクロス株式会社の長谷川啓介氏には、実証実験における開発やプロジェクトマネジメントを支援頂きました。

本研究の成果の実環境運用にあたり、多くの関係者の皆様にお世話になりました。個々の御名前は省略させていただきますが、NTT サービスエボリューション研究所の皆様とは数多くの御討論をさせて頂きました。本研究の遂行は、御支援頂いた皆様のご協力により成し得たものです。関係者の皆様方に深く感謝いたします。最後に、研究を何不自由なく行えるよう支援してくれた妻 智美、長女 友那、長男 太吾に心からの感謝の意を込めて本論文を捧げます。

学位論文を構成する論文

- J-1** K. Tsutsumi, K. Imaizumi, Y. Haneda, and H. Takada, “Sound field synthesis based on superposition of multipoles comprising focused monopole sources,” in *Acoustic Science and Technology*, Vol. 41, No. 2, pp. 489 - 500, 2020.
- C-1** K. Tsutsumi, Y. Haneda, K. Noguchi, and H. Takada, “A method to reproduce a directional sound source by using a circular array of focused sources in front of a linear loudspeaker array,” in *Audio Engineering Society Convention 144*, 2018. (Italy) .
- C-2** K. Tsutsumi, H. Takada, K. Noguchi, and Y. Haneda, “Directivity Synthesis with Multipoles Comprising a Cluster of Focused Sources Using a Linear Loudspeaker Array,” in *Proc. of International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018)*, 2018. (Canada) .
- C-3** K. Tsutsumi, K. Imaizumi, A. Nakadaira, and Y. Haneda, “Analytical conversion of circular harmonics expansion coefficients for sound field synthesis by using multipole loudspeaker array,” in *Proc. of the 27th European Signal Processing Conference (EUSIPCO 2019)*, 2019, (Spain).

研究業績リスト

受賞

A-1 日本音響学会 技術開発賞 (2015).

A-2 情報処理学会 FIT 論文賞 (2004).

学術論文誌

J-1 K. Tsutsumi, K. Imaizumi, Y. Haneda, and H. Takada, “Sound field synthesis based on superposition of multipoles comprising focused monopole sources,” in *Acoustic Science and Technology*, Vol. 41, No. 2, pp. 489 - 500, 2020.

J-2 南 憲一, 宮武 隆, 深津 真二, 外村 喜秀, 堤 公孝, “超高臨場感通信技術の取り組み,” *電子情報通信学会誌*, Vol. 103, No. 1, pp. 15 - 21, 2020.

査読付き国際会議

C-1 K. Tsutsumi, Y. Haneda, K. Noguchi, and H. Takada, “A method to reproduce a directional sound source by using a circular array of focused sources in front of a linear loudspeaker array,” in *Audio Engineering Society Convention 144*, 2018. (Italy) .

C-2 K. Tsutsumi, H. Takada, K. Noguchi, and Y. Haneda, “Directivity Synthesis with Multipoles Comprising a Cluster of Focused Sources Using a Linear Loudspeaker Array,” in *Proc. of International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018)*, 2018. (Canada) .

C-3 K. Tsutsumi, K. Imaizumi, A. Nakadaira, and Y. Haneda, “Analytical conversion of circular harmonics expansion coefficients for sound field synthesis by using multipole loudspeaker array,” in *Proc. of the 27th European Signal Processing Conference (EUSIPCO 2019)*, 2019, (Spain).

- C-4 K. Imaizumi, K. Tsutsumi, A. Nakadaira, and Y. Haneda, “Analytical Method of 2.5D Exterior Sound Field Synthesis by Using Multipole Loudspeaker Array,” in Proc. of IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA 2019), 2019, (America).

国内口頭発表

- D-1 今泉 健太, 堤 公孝, 中平 篤, 羽田 陽一, “直線スピーカアレイによる多重極音源を用いたビーム幅制御,” 日本音響学会 2019 年春季研究発表会, 2019.
- D-2 堤 公孝, 羽田 陽一, 野口 賢一, 高田 英明, “焦点音源法で生成した多重極音源の時間領域実装における音場再現精度改善,” 日本音響学会 2018 年春季研究発表会, 2018.
- D-3 堤 公孝, 高田 英明, 羽田 陽一, “焦点音源で生成した多重極音源による仮想音像の指向性制御,” 日本音響学会 2017 年秋季研究発表会, 2017.
- D-4 堤 公孝, 高田 英明, “補助アレイを用いた波面合成法による焦点音源の受聴エリア拡大,” 日本音響学会 2017 年春季研究発表会, 2017.

紀要

- K-1 堤 公孝, 高田 英明, “客席まで飛び出す音響を実現する波面合成音響技術,” NTT 技術ジャーナル, 2017. (Vol. 29, No. 10) .
- K-2 堤 公孝, 井坂 建, “東京五輪に向けて進化した「Kirari!」,” 月刊ニューメディア: 2018. (5月) .

その他の業績

- O-1 K. Tsutsumi, K. Kikuri, and J. Lecomte, “A packet loss recovery technique with line spectral frequency modification in 3GPP EVS codec,” in Proc. of 2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP), 2015. (America).
- O-2 J. Lecomte, A. Tomasek, G. Markovic, M. Schnabel, K. Tsutsumi, and K. Kikuri, “Enhanced time domain packet loss concealment in switched speech/audio codec,” in Proc. of International Conference on Acoustics, Speech and Signal Processing (ICASSP 2015), 2015. (Australia).

- O-3 堤 公孝, 菊入 圭, “VoLTE のさらなる高音質化と音楽の活用を実現する 3GPP 標準音声符号化方式 EVS,” NTT DOCOMO テクニカルジャーナル, Vol.22, No.4, pp.6-13, Jan. 2015.
- O-4 川嶋 宏彰, 堤 公孝, 松山 隆司, “動的イベントの分節化・学習・認識のための Hybrid Dynamical System,” 第 3 回情報科学技術フォーラム (FIT2004), 2004.

著者略歴

2004年3月 京都大学 工学部 電気電子工学科 卒業

2004年4月 京都大学大学院 情報科学研究科 知能情報学専攻 入学

2006年3月 京都大学大学院 情報科学研究科 知能情報学専攻 卒業

2006年4月 日本電信電話株式会社 NTT サイバースペース研究所 入所

2010年5月 株式会社NTTドコモ 先進技術研究所 入所

2015年4月 日本電信電話株式会社 NTT サービスエボリューション研究所 入所

2017年4月 電気通信大学大学院 情報理工学研究科 情報学専攻 博士後期課程入学