

ービギナーズセミナー 空間音響ことはじめー
スピーカ再生による空間音響
～音場収音・再現技術の基礎～

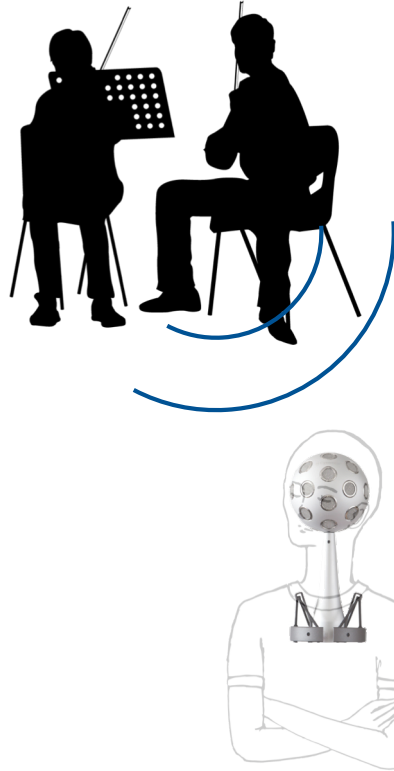
東京大学 大学院情報理工学系研究科／JSTさきがけ

小山 翔一

空間音響技術とは？

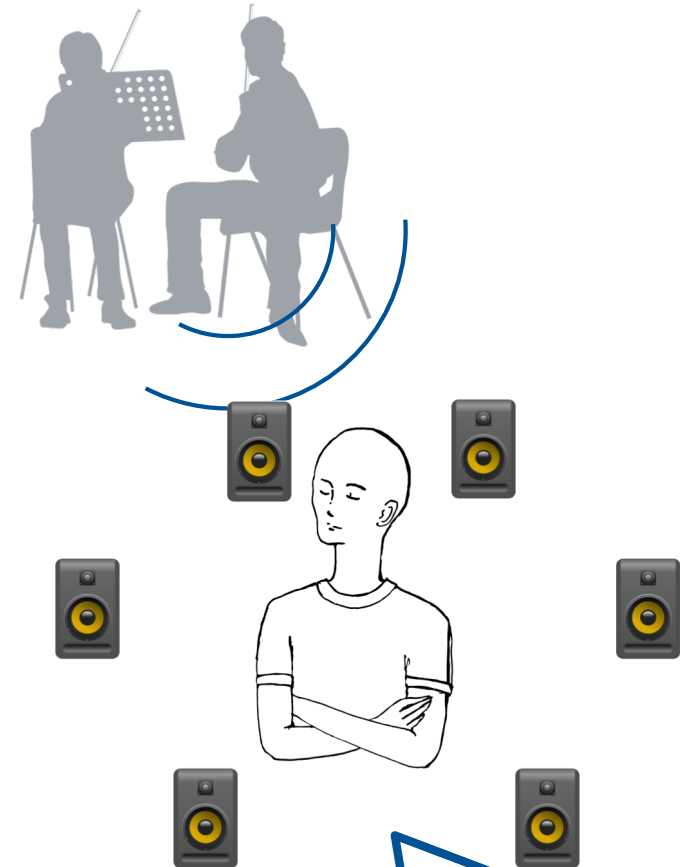
空間音響技術：ある特定の場所での音環境を別の場所で再現する技術

收音側



マイクによる音場の收音
／シミュレーションによる模擬

再現側



スピーカ／ヘッドフォンによる
音場の再現

空間音響技術とは？

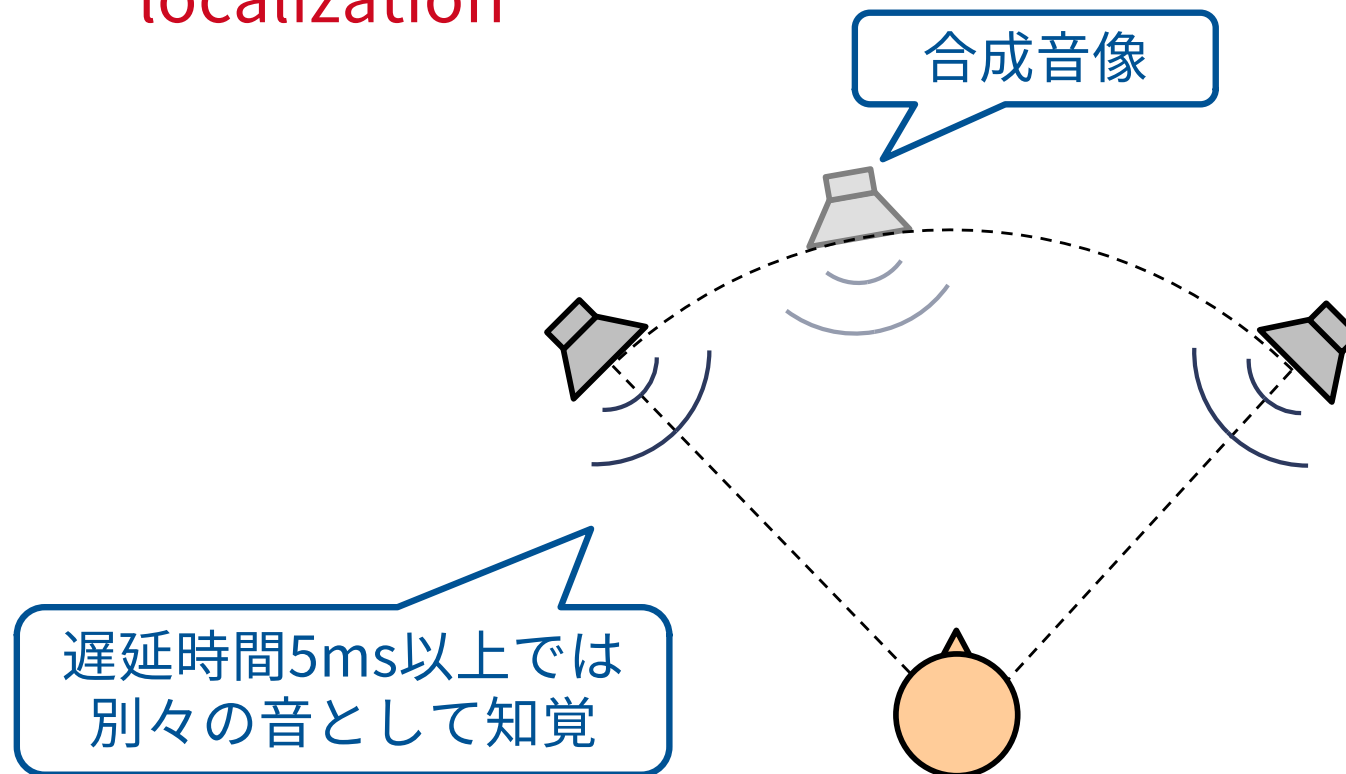
➤ 空間音響技術の分類

- バイノーラル合成 ➡ 前半の森川先生による解説
 - 頭部伝達関数(HRTF)やそれを近似したものを用いて両耳の信号（**バイノーラル信号**）を合成し，提示する
- ステレオフォニック・サラウンド方式
 - **Summing localization**と呼ばれる効果を用いて空間的な音像を提示
- 音場再現
 - 複数スピーカを用いて音空間そのものを物理的に合成

➡ ここでは，受聴者のHRTFを必要とせず，また基本的にヘッドフォンではなくスピーカを用いる空間音響技術である，音場再現技術を中心に解説する。

ステレオフォニック・サラウンド方式

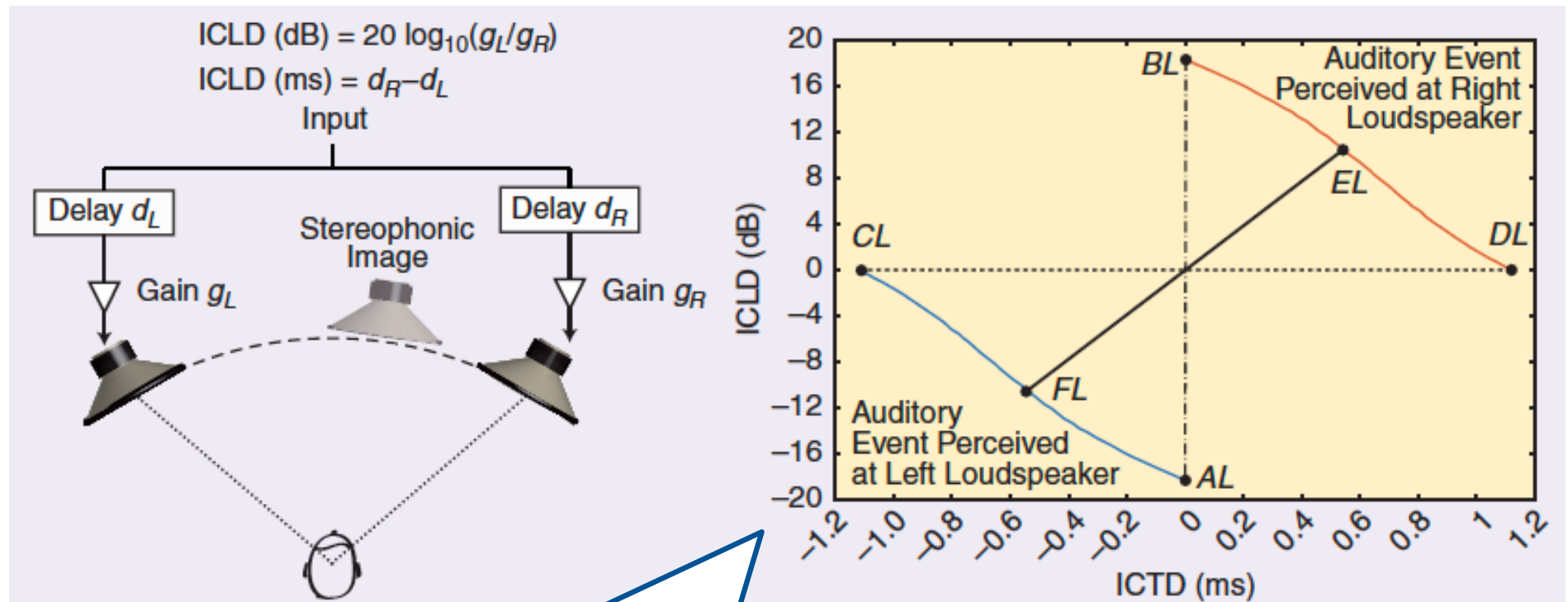
- 複数スピーカのチャンネル間での時間差やレベル差によって空間的な音像を提示：2chステレオ，5.1chサラウンド，22.2chマルチチャンネル音響など
- 遅延時間1ms以下で2つの音源から同一の信号が到達した場合，それらの音源の間に音像を知覚する：Summing localization



[Hacihabiboglu+ 2017]

2chステレオ

- Summing localizationを利用し，2つのスピーカのチャンネル間時間差（interchannel time difference: ICTD），チャンネル間レベル差（interchannel level difference: ICLD）によって音像を制御：パンニング（panning）



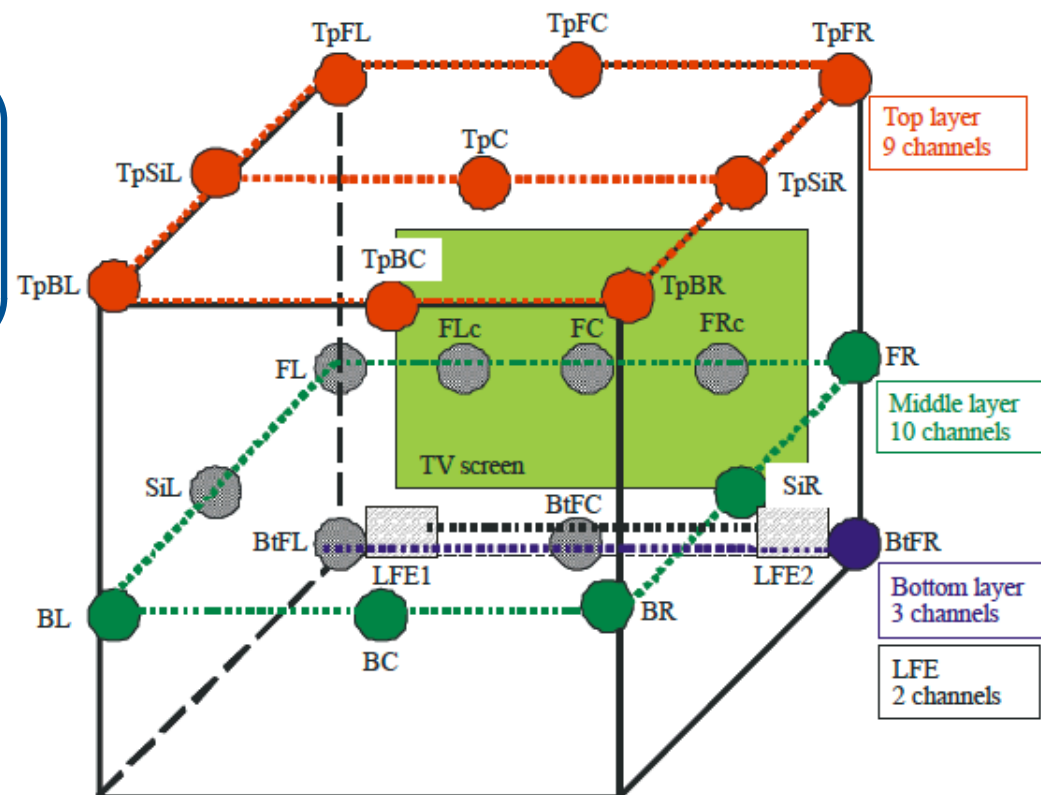
ICTD, ICLDの様々な操作方法が存在（panning law）

[Hacihabiboglu+ 2017]

サラウンド／マルチチャネルステレオ

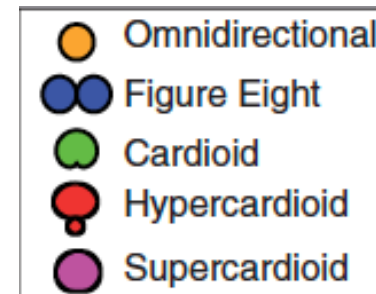
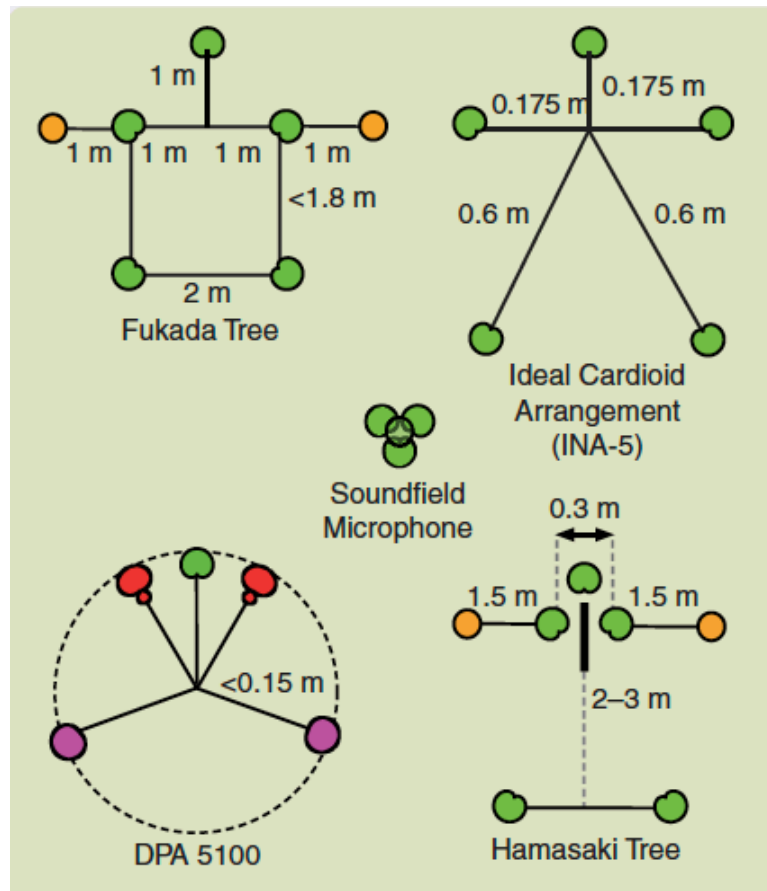
- サラウンド，マルチチャネルステレオ
 - 受聴者を取り囲むようにスピーカを配置
- 様々な方式が存在
 - 5.1, 7.1, 10.2, 22.2, Dolby ATMOS, DTS-X, Auro-3Dなど

2chステレオと同様に
パンニングによって
音像を制御



サラウンド方式における収録

- 音像を合成する場合は各音源のICTD・ICLDを任意に操作可能。では、マイクを使って収録する場合はどうする？
- ➡ 複数のマイクを組み合わせたマイクアレイを用いる。現状は経験的な設計方法によるものがほとんど。



5.1chサラウンド収録
のためのマイクアレイ

[Hacihabiboglu+ 2017]

ステレオフィニック・サラウンド方式

➤ Pros

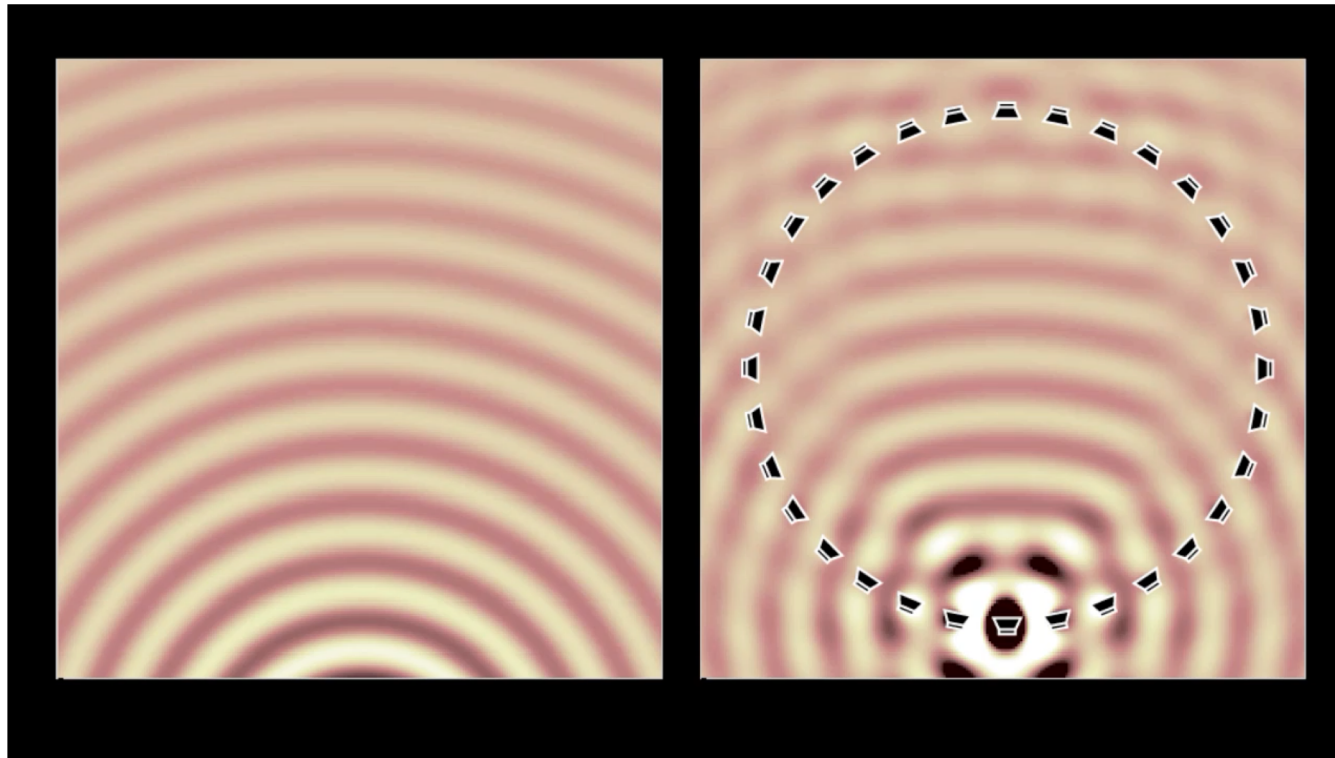
- 少数のスピーカとICTD・ICLDなどの簡単な処理だけで音像を操作できるため、放送・制作用途としては扱いやすい
- 実際よりも強調した表現を行うためのデフォルメもしやすい

➤ Cons

- Summing localizationに基づくため、物理的に音空間そのものが合成されるわけではなく、パンニングやエフェクトによって各音像をデザインすることが必要
- 音空間を合成するのではなく、マイクで収録して再生する場合には、現状の収録方法はあくまで経験的なもの
- スピーカの中心位置で受聴しなければ十分な効果が得られない（**スイートスポット**）

音場再現

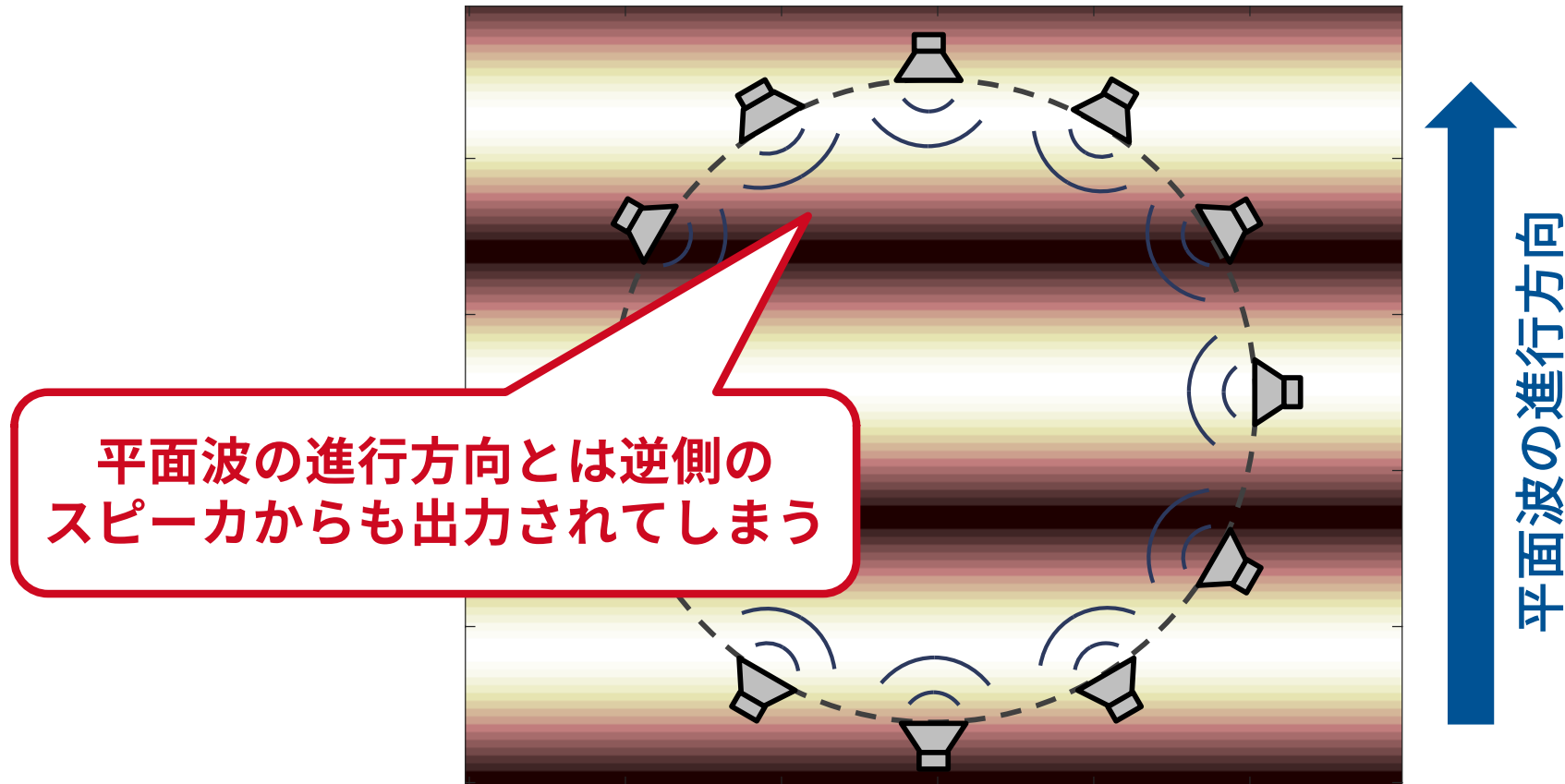
- 複数スピーカを用いて音空間そのものを物理的に合成
 - 広い受聴領域が実現可能であり，受聴者が複数いる場合や，受聴者が動く場合にも適応可能



概念としては1990年頃からあったが，
理論的な整理がなされたのはここ10年ほど

音場再現

- ホイヘンスの原理で考えると，スピーカを連続的に並べ，それぞれのスピーカ位置での音圧を出力すればよい？



各スピーカ位置での音圧を直接出力しても
音場を再現することはできない

ある領域内で所望の音場を合成するために，複数のスピーカ
が出力すべき駆動信号を求める一種の逆問題

➤ 駆動信号を求める方法は大きく分けて二つ

– 解析的な定式化に基づく方法

- 波面合成法 (Wave Field Synthesis: WFS)
- 高次アンビソニックス (Higher Order Ambisonics: HOA)

➡ 逆問題を順問題的に解けるので数値的に安定だが，
解析解が得られる条件が限られる。

– 数値的な最適化に基づく方法

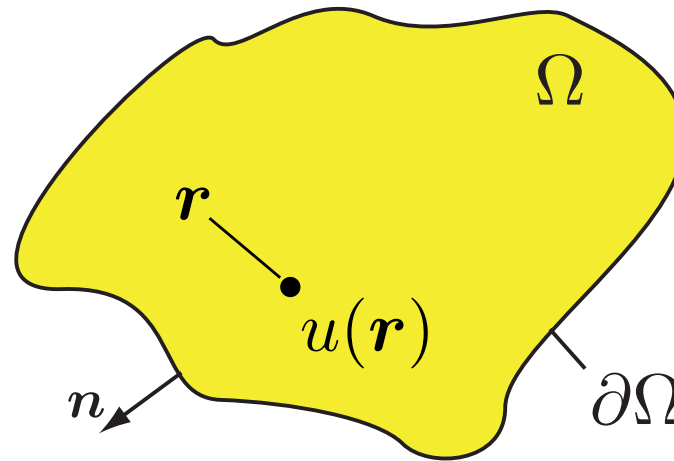
- 多点音圧制御法 (Pressure-Matching)
- モードマッチング法 (Mode-Matching)

➡ スピーカ配置など様々な条件に柔軟に適用可能だが，
数値的に不安定となる場合がある。

音場再現 — 解析的な方法

➤ Kirchhoff-Helmholtz積分方程式

- 波動方程式の周波数領域表現である，Helmholtz方程式に対してGreenの定理を適用することで導出可能。



$$u(\mathbf{r}) = \int_{\partial\Omega} \left(G(\mathbf{r}|\mathbf{r}_s) \frac{\partial u(\mathbf{r}_s)}{\partial \mathbf{n}} - u(\mathbf{r}_s) \frac{\partial G(\mathbf{r}|\mathbf{r}_s)}{\partial \mathbf{n}} \right) d\mathbf{r}_s$$

➡ 音源を含まない領域内の音場は，境界面上の音圧と音圧勾配が決めれば一意に決定できる。

音場再現 — 解析的な方法

➤ Kirchhoff-Helmholtz積分方程式

- 波動方程式の周波数領域表現である，Helmholtz方程式に対してGreenの定理を適用することで導出可能。

Diagram illustrating the Kirchhoff-Helmholtz integral equation. A volume Ω is shown with a boundary $\partial\Omega$. A point $\mathbf{r}$ is inside the volume, and $u(\mathbf{r})$ is the value of the field at that point. The boundary $\partial\Omega$ is labeled as the "二次音源面" (Secondary source surface). The integral equation is given by:

$$u(\mathbf{r}) = \int_{\partial\Omega} \left(G(\mathbf{r}|\mathbf{r}_s) \frac{\partial u(\mathbf{r}_s)}{\partial n} - u(\mathbf{r}_s) \frac{\partial G(\mathbf{r}|\mathbf{r}_s)}{\partial n} \right) d\mathbf{r}_s$$

The terms in the integrand are identified by callouts:

- $G(\mathbf{r}|\mathbf{r}_s)$ is the "モノポールの伝達関数" (Monopole transmission function).
- $\frac{\partial u(\mathbf{r}_s)}{\partial n}$ is the "境界面での音圧勾配" (Sound pressure gradient on the boundary surface).
- $u(\mathbf{r}_s)$ is the "境界面での音圧" (Sound pressure on the boundary surface).
- $\frac{\partial G(\mathbf{r}|\mathbf{r}_s)}{\partial n}$ is the "ダイポールの伝達関数" (Dipole transmission function).

➡ モノポール・ダイポール特性のスピーカを連続的に配置し，独立に制御することで音場を合成できる？

音場再現 — 解析的な方法

➤ Kirchhoff-Helmholtz積分方程式

- 波動方程式の周波数領域表現である，Helmholtz方程式に対してGreenの定理を適用することで導出可能。

Diagram illustrating the Kirchhoff-Helmholtz integral equation. A volume Ω is shown with a boundary $\partial\Omega$. A point $\mathbf{r}$ is inside Ω with value $u(\mathbf{r})$. The boundary $\partial\Omega$ is labeled as the "二次音源面" (Secondary source surface). The integral equation is:

$$u(\mathbf{r}) = \int_{\partial\Omega} \left(G(\mathbf{r}|\mathbf{r}_s) \frac{\partial u(\mathbf{r}_s)}{\partial n} - u(\mathbf{r}_s) \frac{\partial G(\mathbf{r}|\mathbf{r}_s)}{\partial n} \right) d\mathbf{r}_s$$

Callouts identify the terms in the integral:

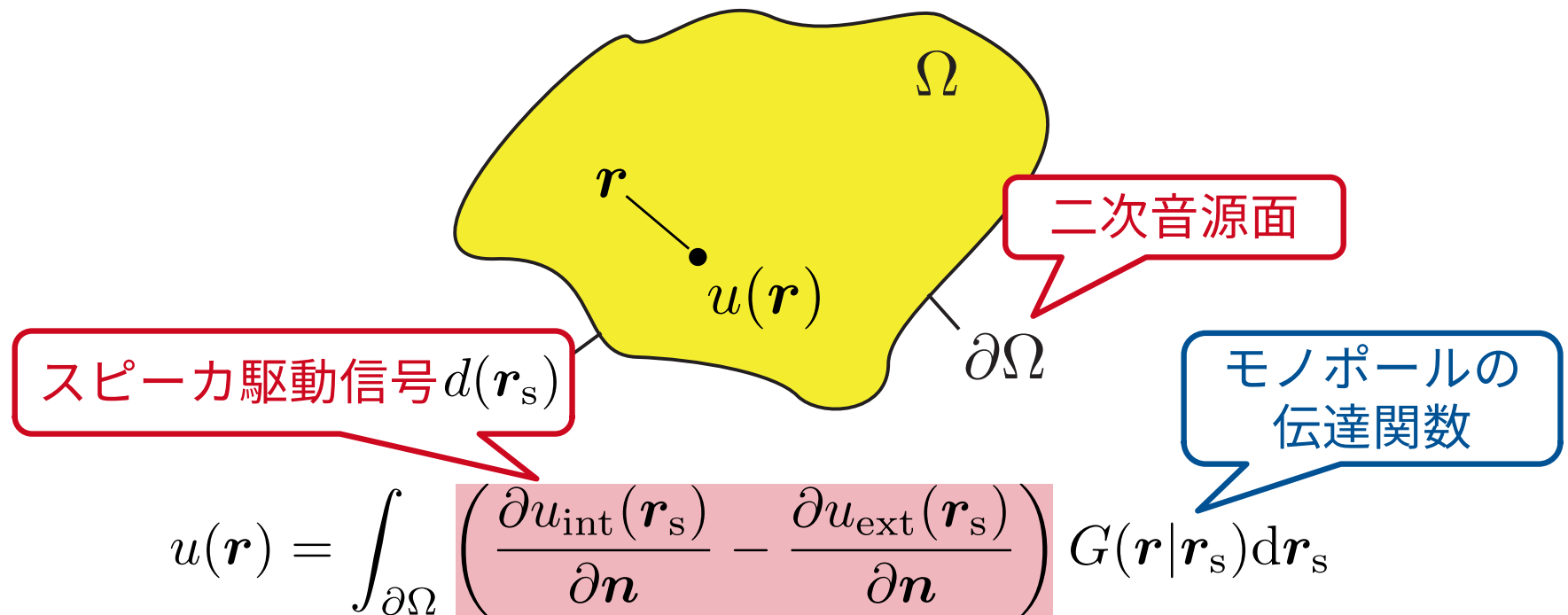
- モノポールの伝達関数 (Monopole transmission function) points to $G(\mathbf{r}|\mathbf{r}_s)$.
- 境界面での音圧勾配 (Boundary surface pressure gradient) points to $\frac{\partial u(\mathbf{r}_s)}{\partial n}$.
- 境界面での音圧 (Boundary surface pressure) points to $u(\mathbf{r}_s)$.
- ダイポールの伝達関数 (Dipole transmission function) points to $\frac{\partial G(\mathbf{r}|\mathbf{r}_s)}{\partial n}$.

二次音源の伝達関数（Green関数）は一般的なスピーカの指向特性として解くことが必要

音場再現 — 解析的な方法

➤ Single layer potential [Fazi+ 2009, 小山 2012]

- 音源を含まない領域内の任意の音場は、境界面上に連続的に分布するモノポール（点音源）によって表現可能



➡ モノポール特性を持つスピーカを境界面に連続的に配置することで任意の音場を合成可能。

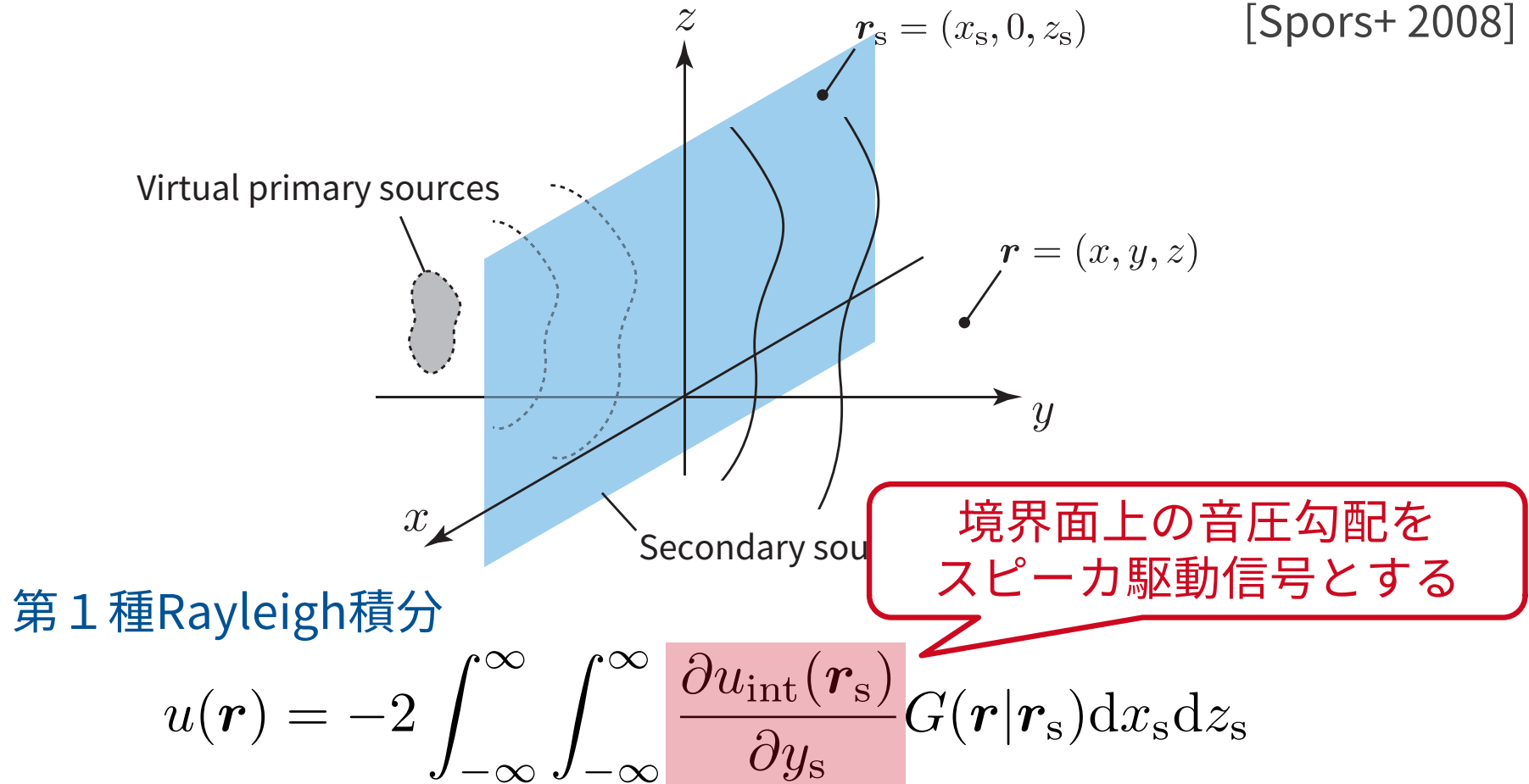
領域 Ω が単純な形状の場合は解析的に
スピーカ駆動信号を導出できる

波面合成法：WFS

➤ 波面合成法（Wave Field Synthesis: WFS）

- Single layer potentialにおいて境界面を無限大の平面とした場合

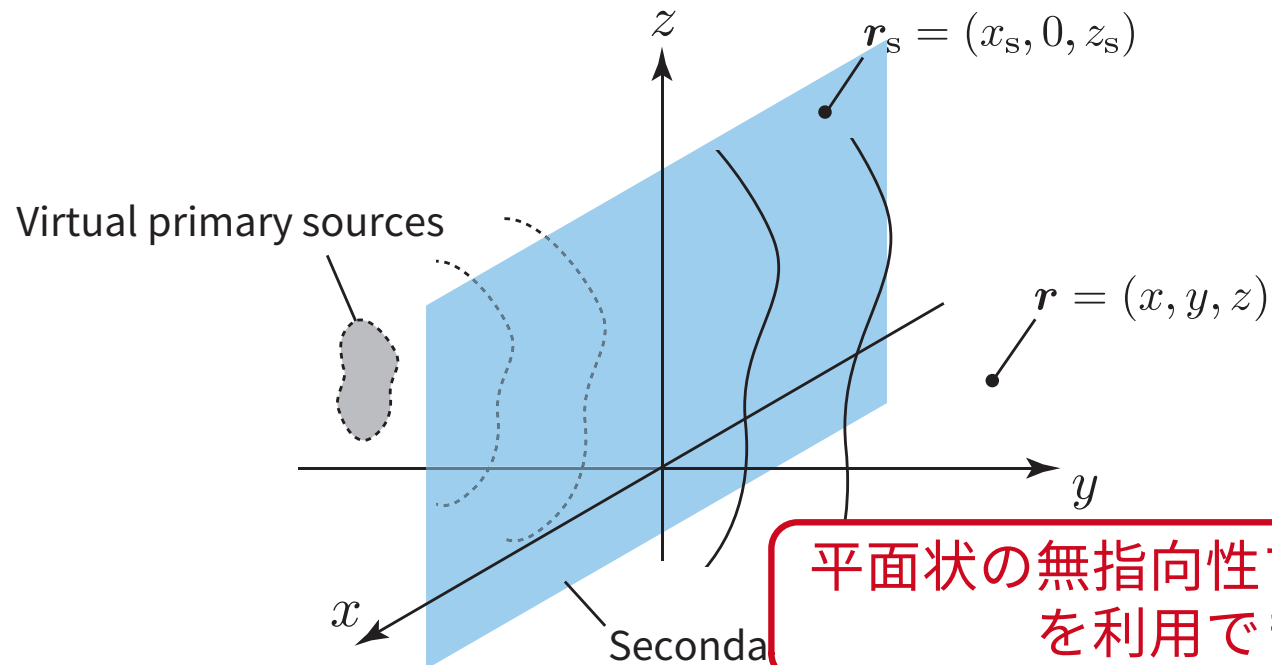
[Spors+ 2008]



- 音場を合成する場合は所望音場の音圧勾配を計算可能
- ただし音場を收音する場合，音圧勾配の計測は困難

波面再構成フィルタ法：WFR Filter

- 収音した音場を再現するにはどうする？
 - 波数領域で音圧分布から音圧勾配の分布を推定 [Koyama+ 2013]



平面状の無指向性マイクアレイ
を利用できる

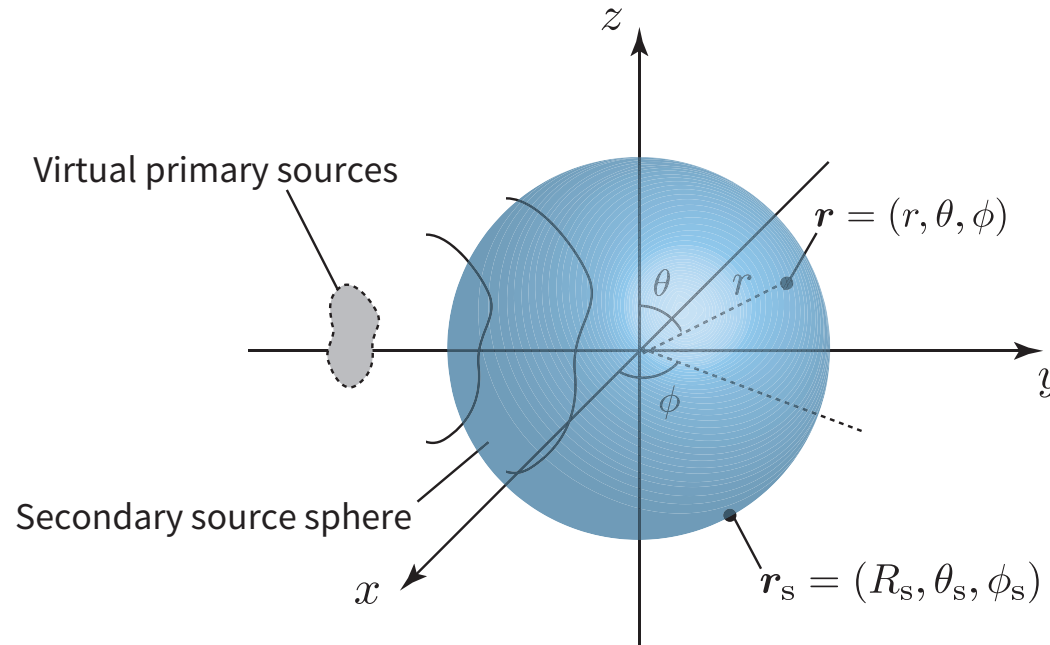
第1種Rayleigh積分の波数領域表現

$$\tilde{u}(k_x, y, k_z) = -2i\sqrt{k^2 - k_x^2 - k_z^2} \tilde{u}_{\text{int}}(k_x, 0, k_z) \tilde{G}(k_x, y, k_z)$$

- 平面状のマイク・スピーカアレイを用いた音場収音・再現
- 直線状アレイを用いる形に近似することも可能 (2.5D)

高次アンビソニックス：HOA

- 高次アンビソニックス (Higher Order Ambisonics: HOA)
 - Single layer potentialにおいて境界面を球状とした場合



音場の球波動関数展開表現

$$u_{\text{int}}(\mathbf{r}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_{nm} j_n(kr) Y_n^m(\theta, \phi)$$

球Bessel関数

$$u_{\text{ext}}(\mathbf{r}) = \sum_{n=0}^{\infty} \sum_{m=-n}^n B_{nm} h_n^{(1)}(kr) Y_n^m(\theta, \phi)$$

球面調和関数

球Hankel関数

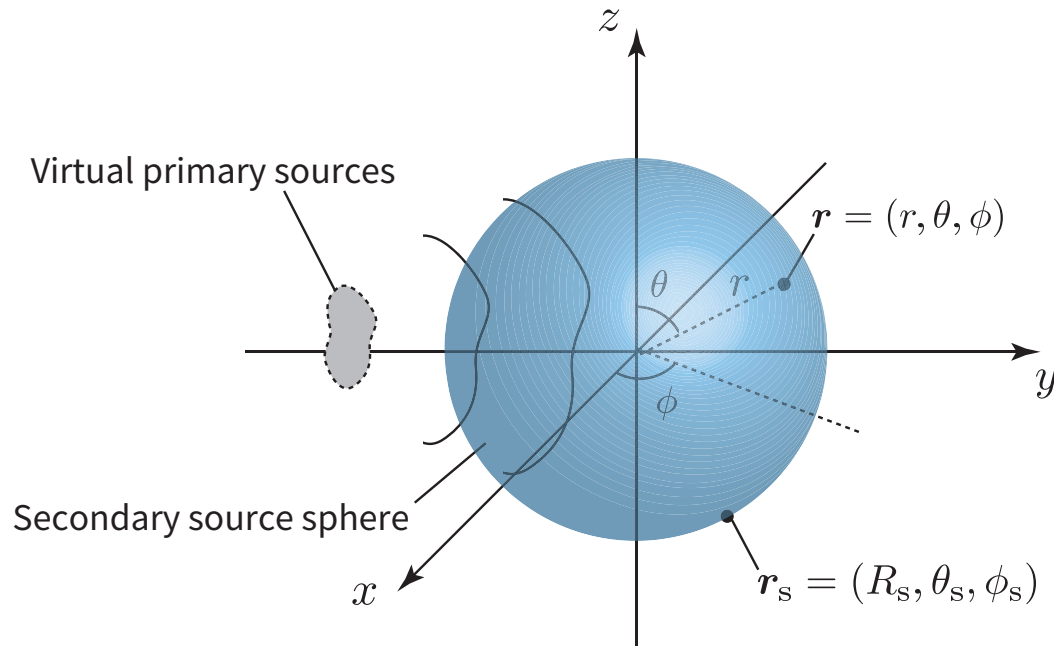
高次アンビソニックス：HOA

➤ 球波動関数展開表現をSingle layer potentialに代入

– 球の半径を R_s とすると，スピーカ駆動信号 $d(\mathbf{r}_s)$ は，

所望音場の
展開係数

$$d(\mathbf{r}_s) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \frac{A_{nm}}{ikR_s^2 h_n^{(1)}(kR_s)} Y_n^m(\theta, \phi)$$



[Poletti 2005]

- 球表面上に配置したスピーカによって内部の音場を再現
- 第1次までの展開係数を用い，各スピーカを平面波近似したものがいわゆるAmbisonicsに相当

高次アンビソニックス：HOA

➤ 収音した音場を再現するにはどうする？

- 収音場の球波動関数による展開係数 A_{nm} を推定することが必要
- 球状マイクアレイを用いることで推定可能だが、ある特定の周波数で推定ができない**禁止周波数問題**が知られている
- 音響的に剛体のバッフル上にマイクを設置する，指向性マイクによるアレイを用いる，多重の球状アレイ配置を用いる，などの方法で**禁止周波数問題を回避** [Poletti 2005, Koyama+ 2016]
- 分散配置したマイクアレイによる推定法も提案されている
[Samarasinghe+ 2012, Ueno+ 2018]



mh acoustics Eigenmike microphone
<https://mhacoustics.com/>

音場再現 — 数値的な方法

- 離散配置された L 個のスピーカにより，対象領域 Ω 内で所望の音場を合成

スピーカによる合成音場

所望音場

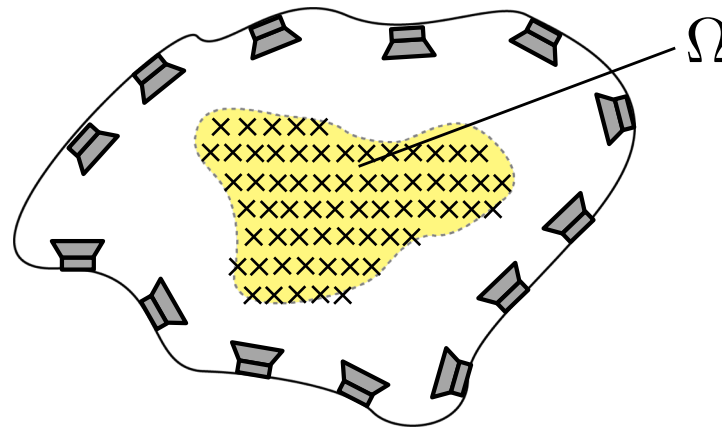
$$\text{minimize}_{d_l} \int_{\mathbf{r} \in \Omega} \left| \sum_{l=1}^L d_l g_l(\mathbf{r}) - u_{\text{des}}(\mathbf{r}) \right|^2 d\mathbf{r}$$

l 番目のスピーカ駆動信号 × 伝達関数

➡ 領域 Ω に関する積分を含むため，直接最適化することは困難

多点音圧制御法：Pressure-Matching

- 対象領域 Ω を離散化し，離散的な制御点上での音圧が所望音場と一致するように駆動信号を求める [Miyoshi+ 1988, Kirkeby+ 1993]



二次音源と制御点間の
伝達関数行列

スピーカ駆動信号
のベクトル

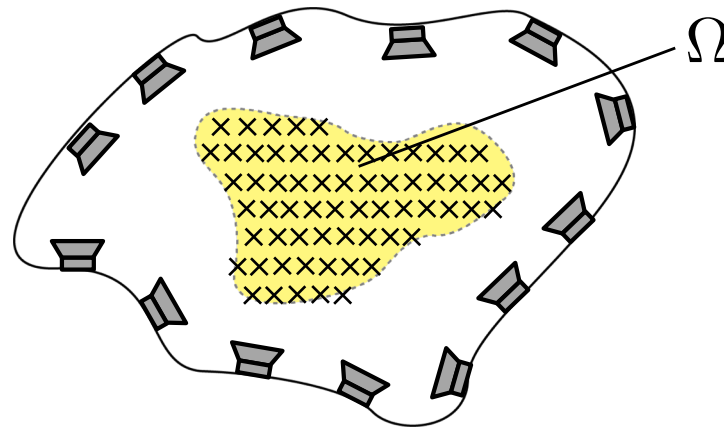
$$\underset{d}{\text{minimize}} \left\| Gd - u^{\text{des}} \right\|^2$$

所望音圧分布のベクトル

$$\Rightarrow d = (G^H G + \lambda I)^{-1} G^H u^{\text{des}}$$

多点音圧制御法：Pressure-Matching

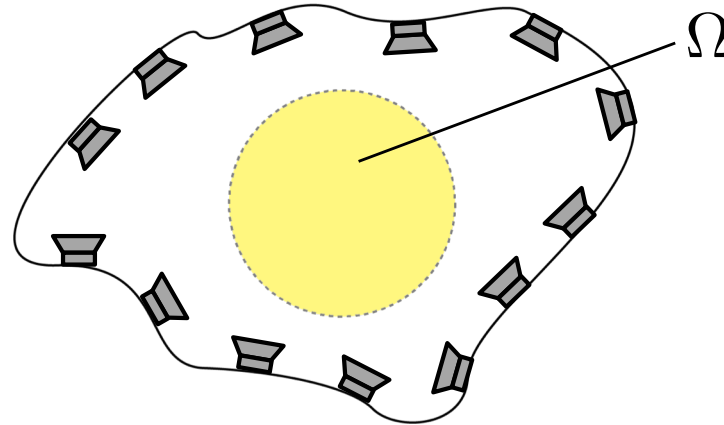
- 対象領域 Ω を離散化し，離散的な制御点上での音圧が所望音場と一致するように駆動信号を求める [Miyoshi+ 1988, Kirkeby+ 1993]



- 残響環境下で収音した音場を再現する場合に，対象領域内の音圧分布を密に計測することは難しい
- Kirchhoff-Helmholtz積分方程式によれば，境界面上の音圧と音圧勾配を一致させればよいが，音圧勾配の計測は困難
- できるだけ少数の素子でPressure-matchingを実現するためのスピーカ・制御点配置最適化法 [Koyama+ 2018]

モードマッチング法：Mode-Matching

- 球波動関数展開による表現を用いて，展開係数の二乗誤差を最小化 [Poletti 2005]



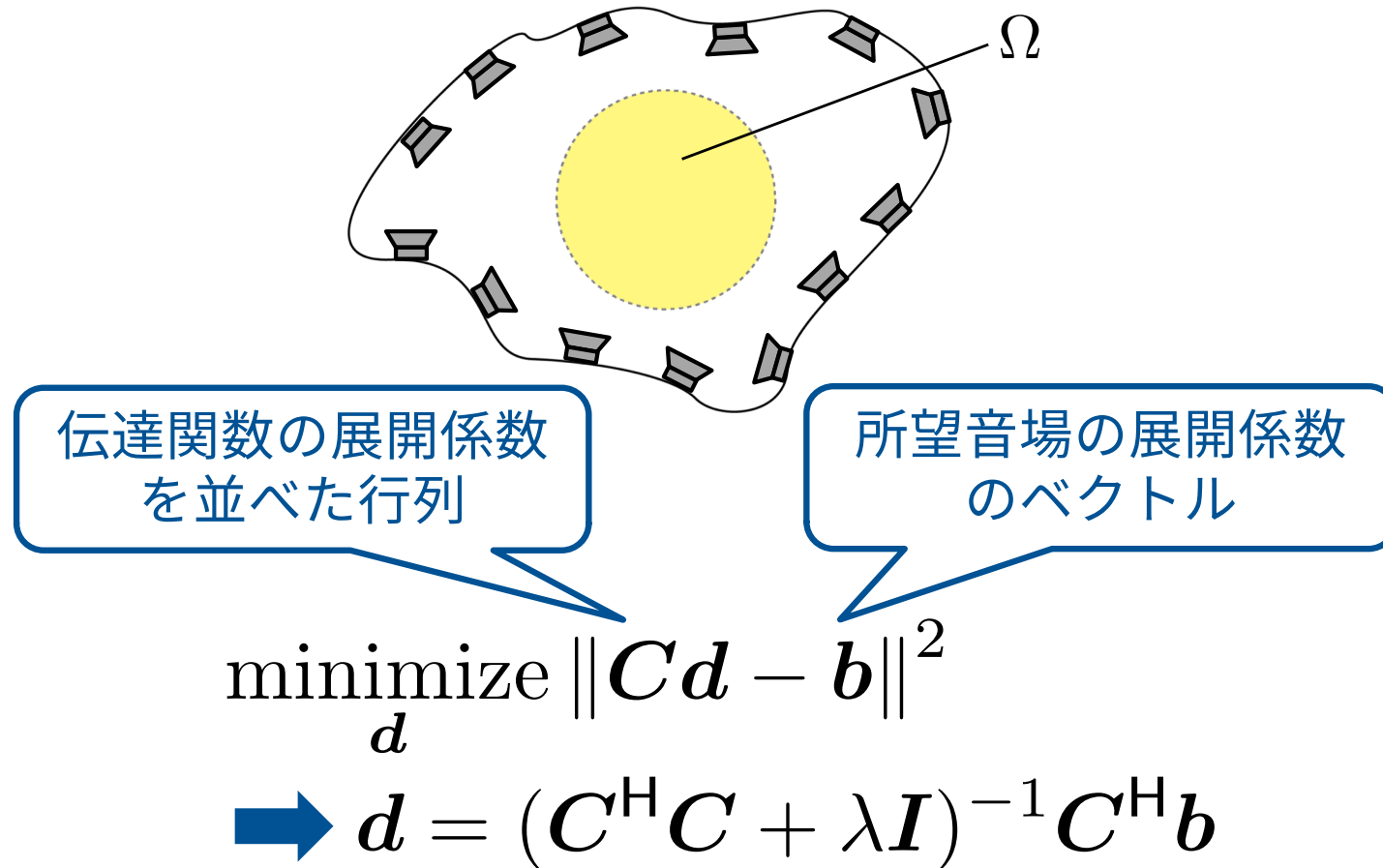
スピーカ伝達関数： $g_l(\mathbf{r}) = \sum_{\nu, \mu} c_{l, \nu, \mu} \varphi_{\nu, \mu}(\mathbf{r})$

所望音場： $u_{\text{des}}(\mathbf{r}) = \sum_{\nu, \mu} b_{\nu, \mu} \varphi_{\nu, \mu}(\mathbf{r})$

基底関数（球波動関数）： $\varphi_{\nu, \mu}(\mathbf{r}) = \sqrt{4\pi} j_{\nu}(kr) Y_{\nu, \mu}(\theta, \phi)$

モードマッチング法：Mode-Matching

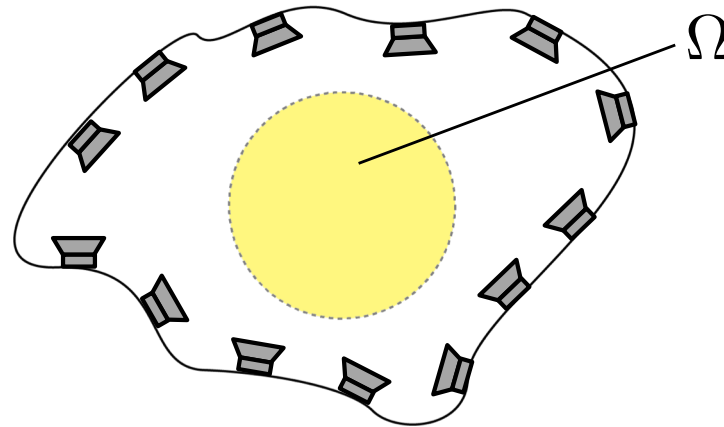
- 球波動関数展開による表現を用いて，展開係数の二乗誤差を最小化 [Poletti 2005]



球波動関数の展開係数を制御することで，
合成音場がHelmholtz方程式に従うことを暗に制約

モードマッチング法：Mode-Matching

- 球波動関数展開による表現を用いて，展開係数の二乗誤差を最小化 [Poletti 2005]

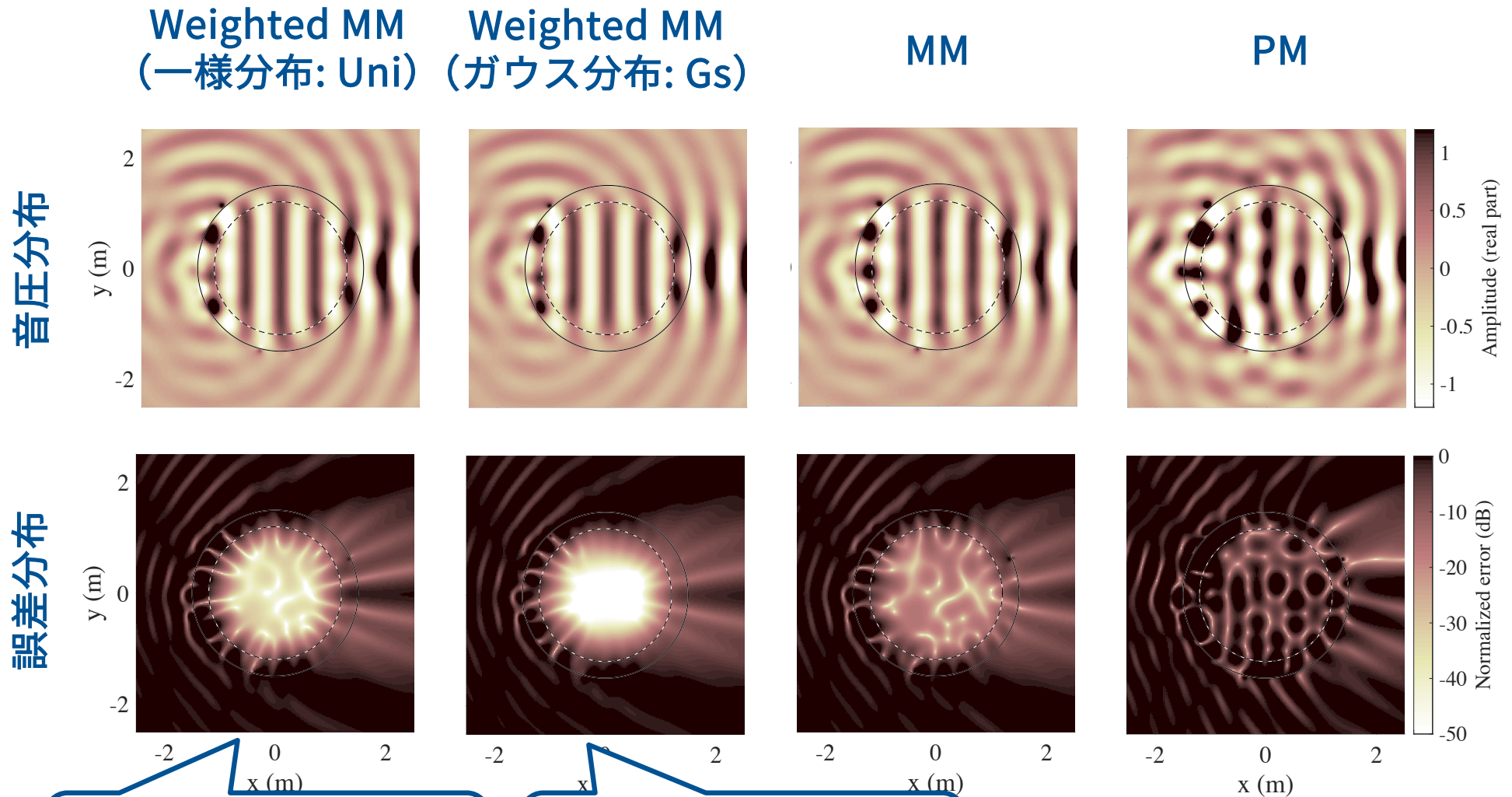


- 展開係数の打ち切りが必要となるが，制御領域の半径を R として， $\lceil kR \rceil$ や $\lceil ekR/2 \rceil$ など経験的な方法が用いられてきた
- 領域積分の定式化から，各展開係数の重みを厳密に導出することで，経験的な打ち切りを回避する**重み付きモード・マッチング法** (Weighted Mode-Matching) [Ueno+ 2019]

球状マイクアレイなどを用いて球波動関数の展開係数を推定する
音場収音手法との相性が良い

シミュレーション実験例

- 144個の単一指向性スピーカ による平面波の合成 [Ueno+ 2019]

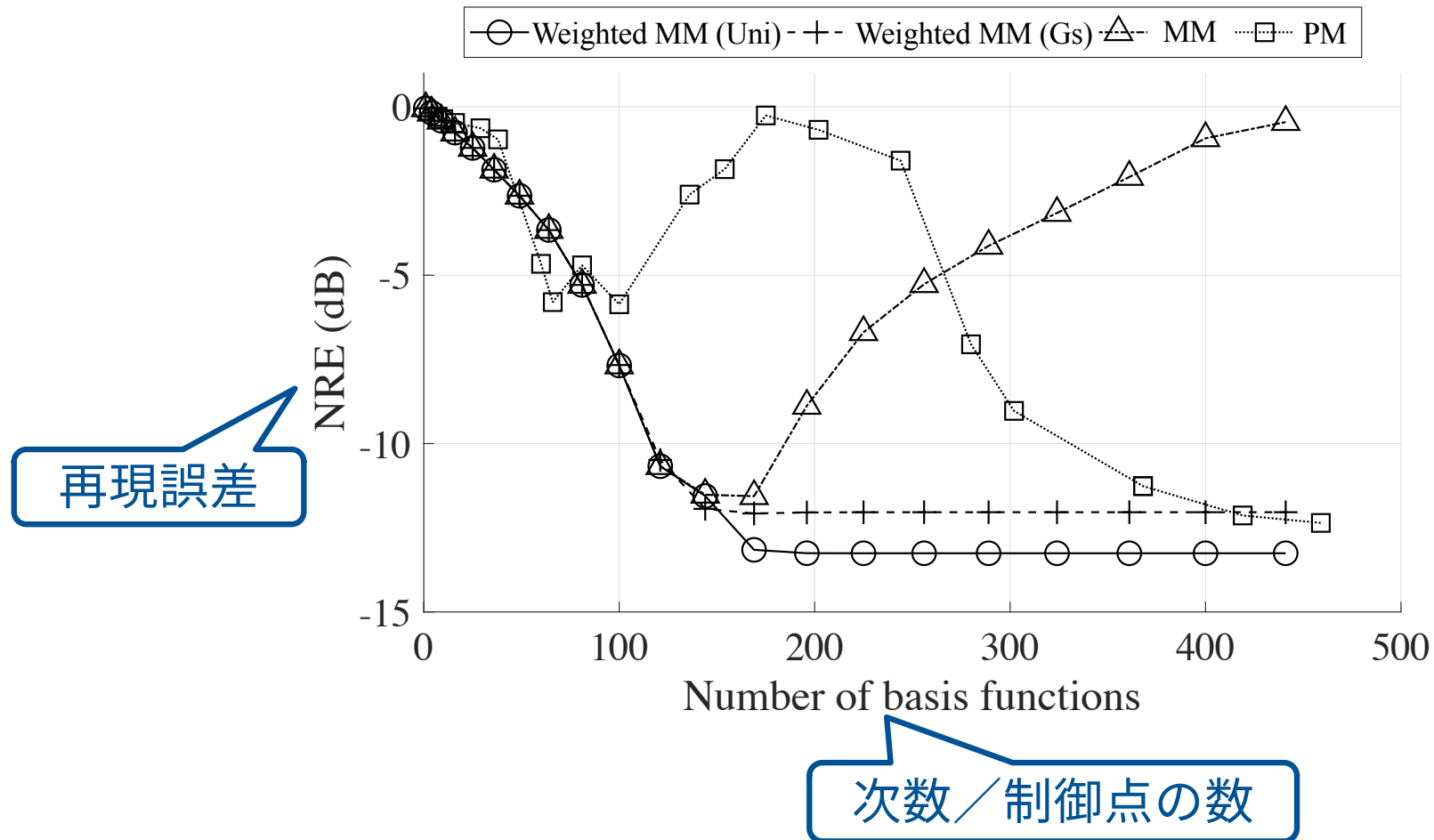


領域内で一様な精度
となる重みを利用

中心付近で精度を
高くする重みを利用

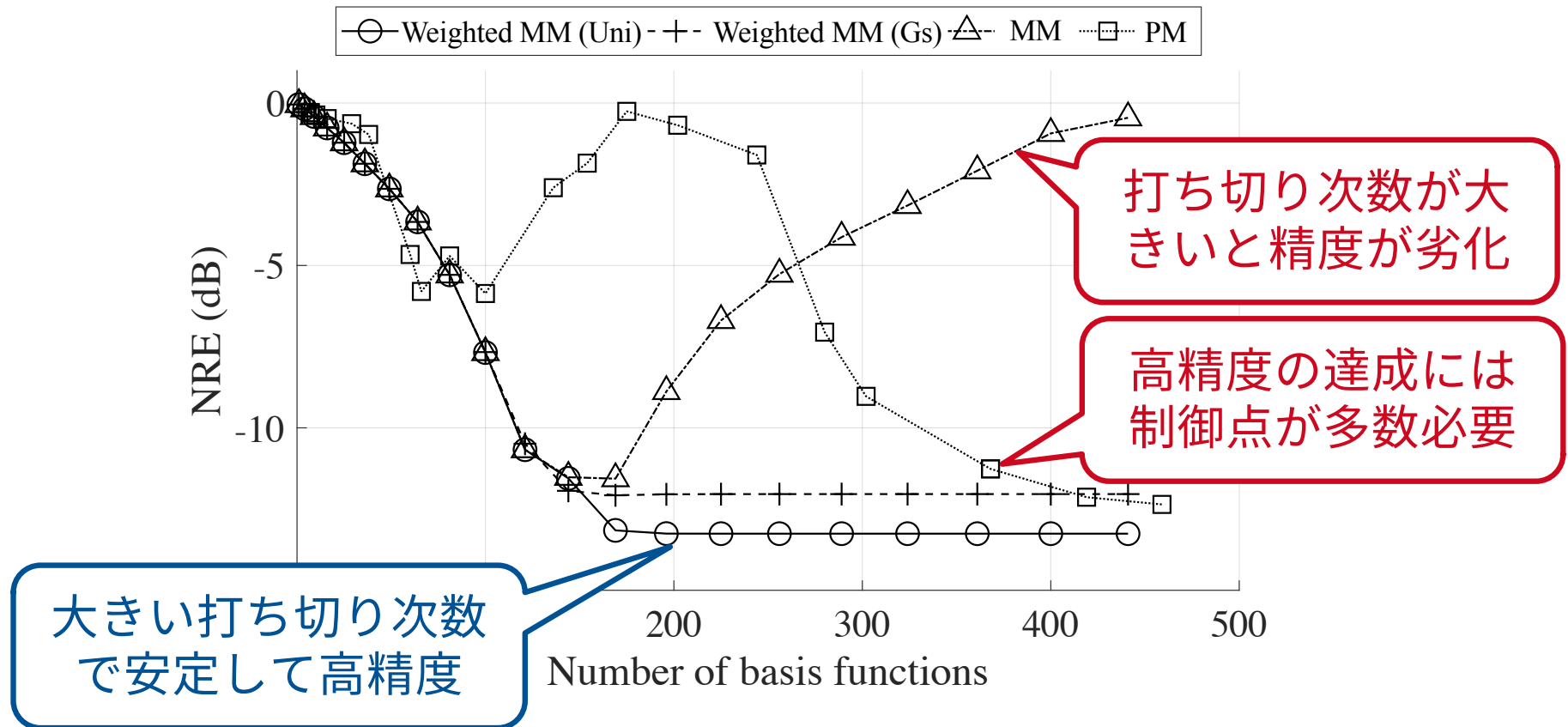
シミュレーション実験例

- 144個の単一指向性スピーカ による平面波の合成 [Ueno+ 2019]



シミュレーション実験例

- 144個の単一指向性スピーカ による平面波の合成 [Ueno+ 2019]



- 球波動関数展開により制御すべきパラメータ数を削減可能
- 重み付きモードマッチング法は次数の打ち切りに非依存

音場再現手法のまとめ

解析的な方法	数値的な方法
空間領域表現 による方法 - WFS [Berkhout+ 1993, Spors+ 2008] - Source-Location-Informed method [Koyama+ 2015]	- Pressure-Matching [Miyoshi+ 1988, Kirkeby+ 1993] - Boundary Surface Control [Ise 1999]
波数領域表現 に基づく方法 - HOA [Poletti 2005] - Spectral Division Method [Ahrens+ 2008, 2010] - WFR Filter [Koyama+ 2013, 2014, 2016]	- Mode-Matching [Poletti 2005, Betlehem+ 2005] - Weighted Mode-Matching [Ueno+ 2019]

解析的な方法／数値的な方法に加え，
空間領域／波数領域での方法に分類できる

音場再現手法のまとめ

解析的な方法

NFS
[Berkhout+ 1993, Spors+ 2008]

空間領域表
による方法

Pros:

数値的に安定，計算量を削減可能，
妥当な2次元近似が可能 (2.5D)，
解析がしやすい

Cons:

スピーカの配置や指向性に制約，
自由音場を仮定

[Koyama+ 2013, 2014, 2016]

数値的な方法

Pressure-Matching
[Koyoshi+ 1988, Kirkeby+ 1993]
Boundary Surface Control
[199]

Pros:

スピーカの配置・指向性に自由度，
残響環境下での適用が可能

Cons:

数値的に不安定になる場合がある，
2次元スピーカ配置での妥当な近
似が難しい

解析的な方法／数値的な方法に加え，
空間領域／波数領域での方法に分類できる

音場再現手法のまとめ

解析的な方法

数値的な方法

空間領域表現
による方法

- WFS
[Be

Pros:

直感的な方法であり実装が比較的簡単

Cons:

制御すべきパラメータ数が多くなる場合がある，収録した音場を再現する場合に適用が難しい

y+ 1993]
trol

波数領域表現
に基づく方法

- HO

- Sp
[Ah

Pros:

制御すべきパラメータ数を削減できる，収録との相性が良い

Cons:

展開係数の推定方法がやや複雑

m+ 2005]
hing

解析的な方法／数値的な方法に加え，
空間領域／波数領域での方法に分類できる

音場再現システムの実現例

音場の収音・伝送・再現をリアルタイムで実現 [Koyama+ IEICE Trans 2014]

Yokosuka



Network



Musashino



- Loudspeakers (for high freq.): 64, 6cm intervals
- Loudspeakers (for low freq.): 32, 12cm intervals
- Microphones: 64, 6cm intervals
- Array size: 3.84 m
- Sampling freq.: 48 kHz, Delay: 152 ms



音場再現

➤ Pros

- 物理的に音場そのものを再現するため，広い受聴領域が実現可能
- 複数の受聴者や受聴者の動きにもそのまま対応可能であり，空間知覚の観点から有利
- 収音場の再現も可能であり，人手での操作やデザインを介することなく残響のある音場を再現することも可能

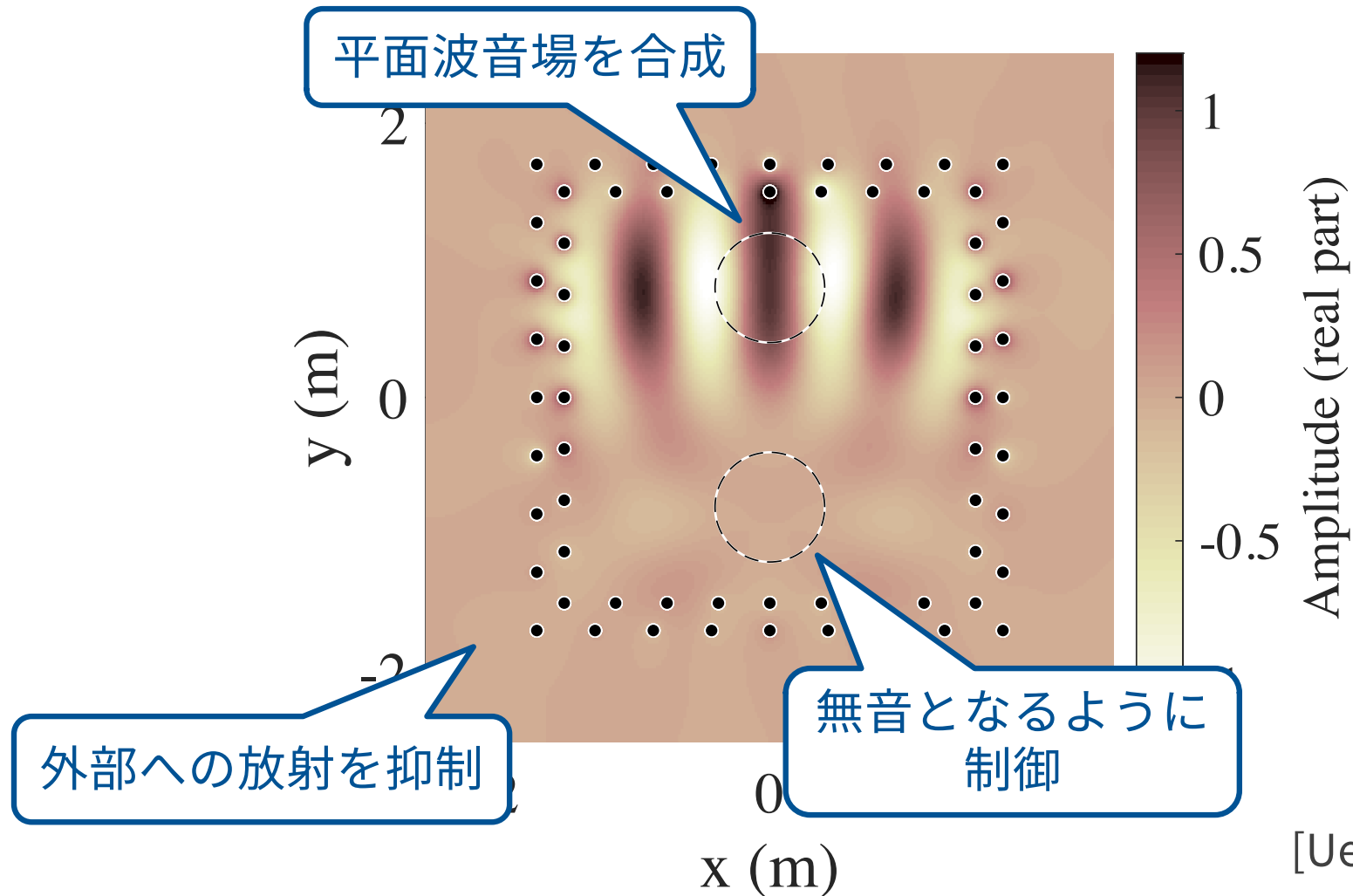
➤ Cons

- 再現可能な周波数の上限や範囲がスピーカ・マイクの数によって決まり，それ以上では空間エイリアシングと呼ばれる誤差が生じる
- 多数のスピーカ・マイクが必要となるため，システムの規模が大きくなる

➡ 素子の小型化・低コスト化，A/D・D/A変換器の多チャンネル化により実現可能性が高まっている

空間音響以外の音場再現

- 音場再現は単純に空間音響としてだけではなく、複数の領域に異なる音場を合成することや、空間的な騒音制御などにも応用できる



[Ueno+ 2019]

まとめ

- 音場再現技術は，ステレオ・サラウンド方式やバイノーラル方式に比べ，**広い受聴領域**を実現できることが特長
- 音場再現手法は，**解析的な方法**と**数値的な方法**の2つに大きく分けられ，それぞれ利点・欠点がある。
- 多数のスピーカ・マイク素子が必要となるが，これを削減するための手法も多数提案されている
 - 音場のスパーズ表現に基づく方法 [Koyama+ 2014, 2018, 2019]
 - 音源位置の事前情報を利用したマイク数削減法 [Koyama+ 2015]
 - 受聴エリアの事前情報を利用したスピーカ数削減法 [Ueno+ 2017, 2019]
 - スピーカ・マイク(制御点)の最適配置法 [Koyama+ 2018]
- 最近では，空間音響以外の応用として，**空間能動騒音制御**への適用なども行われている [Ito+ 2019]

Related Publications

- N. Ueno, S. Koyama, and H. Saruwatari, “Three-dimensional sound field reproduction based on weighted mode-matching method,” *IEEE/ACM Trans. ASLP*, 2019.
- S. Koyama and L. Daudet. “Sparse representation of a spatial sound field in a reverberant environment,” *IEEE J. STSP*, 2019.
- H. Ito, S. Koyama, *et al.* “Feedforward spatial active noise control based on kernel interpolation of sound field,” *Proc. IEEE ICASSP*, 2019.
- S. Koyama, *et al.* “Sparse sound field decomposition for super-resolution in recording and reproduction,” *JASA*, 2018.
- S. Koyama, *et al.* “Joint source and sensor placement for sound field control based on empirical interpolation method,” *Proc. IEEE ICASSP*, 2018.
- N. Ueno, S. Koyama, and H. Saruwatari, “Sound field recording using distributed microphones based on harmonic analysis of infinite order,” *IEEE Signal Process. Letters*, 2018.
- N. Ueno, S. Koyama, and H. Saruwatari, “Listening-area-informed sound field reproduction based on circular harmonic expansion,” *Proc. IEEE ICASSP*, 2017.
- S. Koyama, *et al.* “Analytical approach to transforming filter design for sound field recording and reproduction using circular arrays with a spherical baffle,” *JASA*, 2016.
- S. Koyama, *et al.* “Source-location-informed sound field recording and reproduction,” *IEEE J. STSP*, 2015.
- S. Koyama, *et al.* “Sparse sound field representation in recording and reproduction for reducing spatial aliasing artifacts,” *Proc. IEEE ICASSP*, 2014.
- S. Koyama, *et al.* “Real-time sound field transmission system by using wave field reconstruction filter and its evaluation,” *IEICE Trans. Fundamentals*, 2014.
- S. Koyama, *et al.* “Wave field reconstruction filtering in cylindrical harmonic domain for with-height recording and reproduction,” *IEEE/ACM Trans. ASLP*, 2014.
- S. Koyama, *et al.* “Analytical approach to wave field reconstruction filtering in spatio-temporal frequency domain,” *IEEE Trans. ASLP*, 2013.