

タイトル	音声生成系の放射過程のモデルについて
著者	元木, 邦俊
引用	工学研究 : 北海学園大学大学院工学研究科紀要, 9: 3-14
発行日	2009-09-30

音声生成系の放射過程のモデルについて

元 木 邦 俊*

Model of Radiation Process of Speech Production System

Kunitoshi Motoki*

1. はじめに

音声生成系の工学的モデルとして、音源部、声道部、放射部の各段階からなる縦続システムとして扱う方法が広く用いられている。音源部は声帯音源波形を生成する部分で、粘弾性体としての声帯と声門を通過する高速空気流の相互作用が生じる部分でもあり、その物理モデルの検討やシミュレーションが盛んに行われている。声道部は、音源部により生成された音源波形のスペクトル特性に音韻に従ったスペクトル特徴を与える共鳴器で、声門から口唇に至る空間部分（この部分を声道と呼ぶ）の形状によりその共鳴特性が決定される。声道を形が変化する管と考えると、管の横断長と音波の波長の関係で伝搬の様子は異なる。もし、管の横断長が波長よりも十分に小さい場合は、音波は管軸に沿って一様に伝搬し、管軸に垂直な断面上では音圧と粒子速度の変動は無視できるほど小さいとみなす。この近似の仕方を平面波伝搬の近似という。実際的には、横断長が半波長程度までを平面波伝搬と考える場合が多い。声道の横断長は口腔部で比較的長いですが、最大の横断長を5 cm程度と仮定すると3.5 kHz程度までを平面波伝搬の有効な近似として考えることができ、母音で重要な周波数帯域を含むことができる。平面波伝搬の仮定の下に、声道の特徴は断面積関数によって表すことができ、その共振特性は比較的容易に計算することができる。放射部は、声道終端の境界条件を定めると同時に、遠方点への音響伝達特性を定める部分でもあり、そのスペクトル特性は、+6 dB/octaveの一階微分特性として扱わ

れることが多い。

本稿では、このような音声生成系のモデル化手法において、詳しく解説されることが少ないと思われる音声の放射過程について、1次元モデルに適合させるための扱い方、筆者が行ってきた実験結果、さらには3次元的な音響放射モデルについて紹介する。

2. 放射過程のモデル

音声を空間に放射する部位は、口唇、鼻孔（以下ではこれらを開口端と呼ぶ）が考えられる。実際には、声帯振動や声道壁の振動が体表まで伝わり、頸部などの皮膚からも音波が放射されるものの、これらの影響については十分小さいと考えて音声生成系のモデルでは考慮しないのが普通である。放射過程は、共鳴器の一部から音波が洩れ出て自由空間に放散していく1つの音響現象であるが、工学的な取り扱いが容易になるように、次の2つの部分からなるものとして考える。第一は、終端境界条件を定め、声道の共鳴特性を定める働きをする部分である。これは、声道終端における音圧 P_i と体積速度 U_i の複素比として定義される放射インピーダンス Z_r （駆動点放射インピーダンスと呼ぶこともある）

$$Z_r = \frac{P_i}{U_i} \quad (1)$$

によって特徴付けられる。 Z_r は、変形しないピストン振動板がバツフル表面に設置されたモデルを仮定して計算することができる。放射インピーダンスが定まると、声道モデルの最終区間の特性イ

* 北海学園大学大学院工学研究科電子情報工学専攻

Graduate School of Engineering (Electronics and Information Eng.) Hokkai-Gakuen University

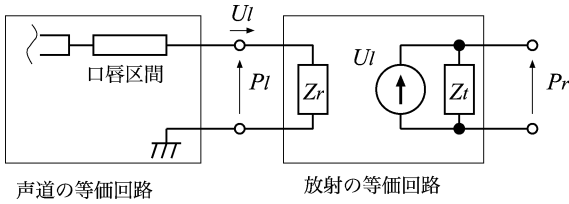


図1 放射過程の等価回路

ンピーダンスとの不整合により開口端での音波の反射係数が決まる。これにより、声道モデルの出力としての体積速度が求まり、この体積速度が外部（遠方点）からみた音声の“音源”とみなされる。第二の部分は、開口端で定まった体積速度を遠方の観測位置（音声を聴取する位置）での音圧 P_r に変換する働きである。これは、観測位置における音圧と開口端の体積速度の複素比として定義される伝達インピーダンス Z_t

$$Z_t = \frac{P_r}{U_i} \quad (2)$$

によって特徴付けられる。図1は、放射過程を電氣的な等価回路で表したもので、声道モデルの終端を表す端子に接続されている。伝達インピーダンスは開口端での体積速度には影響を与えないので、開口端の体積速度 U_i と同じ値を持つ制御電流源により駆動されるものとして考えることができる。このように放射過程を2つの部分からなるものとして考えると、その特性のモデル化と音響計測による特性評価が比較的容易に行える。以下では、放射インピーダンスと伝達インピーダンスの具体的な周波数特性について述べる。

2.1 放射インピーダンス

声道モデルの終端条件として、理想的な開放、すなわち $Z_r=0$ として開口端での音圧を0とする簡化も行われるが、この近似が成立するのは開口端の大きさが波長に比べて極めて小さい場合のみで、口唇の開口サイズと音声帯域の音波に対しては適当な近似ではない。実体に近いモデルとしては、適当なバツフルを有するピストン振動板からの音波の放射が考えられている^{1,2)}。粒子速度が開口端面上において一様であると仮定すると、放射面をピストン振動音源と同一とみなすことができる。振動板には、発生した音圧 $P_s(\mathbf{r}_s)$ による反作用力が働く。ここで、 \mathbf{r}_s は振動板面上の位置で、ピストン面上で一様な振動速度 V_i を仮定しても、振動板上の音圧は一様にはならないことを

注意しておく。放射インピーダンス Z_r は、ピストン振動板が与える体積速度 U_i と振動板面上の平均音圧 $P_{s,av}$ の複素比として定義される。

$$Z_r = \frac{P_{s,av}}{U_i} = \frac{\iint_S P_s(\mathbf{r}_s) dS}{S^2 V_i} \quad (3)$$

ここで、 S は振動板の面積である。ピストン振動板面上の平均音圧 $P_{s,av}$ を声道モデルの終端音圧 P_l として考える。式(3)は、ピストン振動板とバツフル面の形状を与えることで計算することができ、図2のように半径 a の円形ピストン振動板が無限平面バツフルに設置された場合には、次のように求められる³⁾。

$$Z_r = \frac{\rho c}{\pi a^2} \{R_r + jX_r\} \quad (4)$$

$$R_r = 1 - \frac{J_1(2ka)}{ka} \approx \frac{(ka)^2}{2}$$

$$X_r = \frac{H_1(2ka)}{ka} \approx \frac{8ka}{3\pi}$$

ここで、 ρ ：空気密度、 c ：音速、 k ：波長定数(= ω/c 、 ω ：角周波数)である。 $J_1(\cdot)$ 、 $H_1(\cdot)$ は第1種 Bessel 関数、Struve 関数で、次のように展開できる⁴⁾。

$$J_1(2ka) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!(n+1)!} (ka)^{2n+1} \quad (5)$$

$$H_1(2ka) = \sum_{n=0}^{\infty} \frac{(-1)^n}{\Gamma(n + \frac{3}{2})\Gamma(n + \frac{5}{2})} (ka)^{2n+2} \quad (6)$$

$\Gamma(\cdot)$ はガンマ関数である。図3に R_r 、 X_r の特性を示す。式(4)の近似は $ka \ll 1$ の場合で、低い周波数域では、抵抗成分は周波数の2乗に、リアクタンス成分は周波数に比例して増大する。なお、高い周波数域 ($ka \gg 1$) では、 $R_r \approx 1$ 、 $X_r \approx 2/(\pi ka)$ として近似されるが、そのような周波数範囲は音

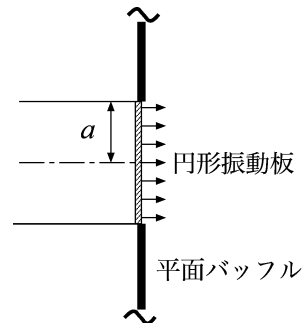


図2 平面バツフルにセットされた円形ピストン振動板

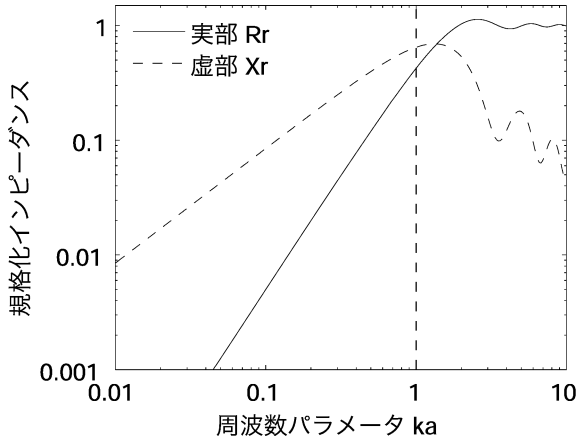


図3 平面バツフルに設置された円形ピストンの規格化放射インピーダンスの実部 (R_r) と虚部 (X_r). 規格化定数は特性インピーダンス $\rho c / \pi a^2$. k 波長定数, ρ 空気密度, c 音速, a ピストン半径.

声帯域よりも相当に高く、一様振動するピストンからの放射とする仮定も妥当ではない。 $a=1.5$ cm ($S=7.1$ cm²) の場合、 $ka=1$ に対応する周波数は、約 3.7 kHz となるので、通常の発話の場合には ka の上限は 1 程度までとなる。

声道終端に放射インピーダンスが接続された場合は、開口端で次の関係が成立すると考えることができる。

$$P_i = Z_{c,n}(U_i^+ + U_i^-) \quad (7)$$

$$U_i = U_i^+ - U_i^- \quad (8)$$

$Z_{c,n}$ は開口端となっている音響管の特性インピーダンス ($=\rho c / \pi a^2$), U_i^+ , U_i^- は、放射面へ入射する体積速度の前進波と反射する後進波を表す。式(1)を用いると、終端での反射係数 r_i は、

$$r_i = \frac{U_i^-}{U_i^+} = \frac{Z_r - Z_{c,n}}{Z_r + Z_{c,n}} \quad (9)$$

となる。図4は r_i の周波数特性を描いたものである。周波数の増加とともに終端での反射係数の大きさが小さくなっていくことが分かる。

$ka \ll 1$ の条件の下に、式(4)を代入すると、

$$r_i \approx -\frac{1 - j(8ka/3\pi)}{1 + j(8ka/3\pi)} \quad (10)$$

となるので、反射による位相変化 θ_i は、

$$\theta_i \approx \pi - \frac{16ka}{3\pi} \quad (11)$$

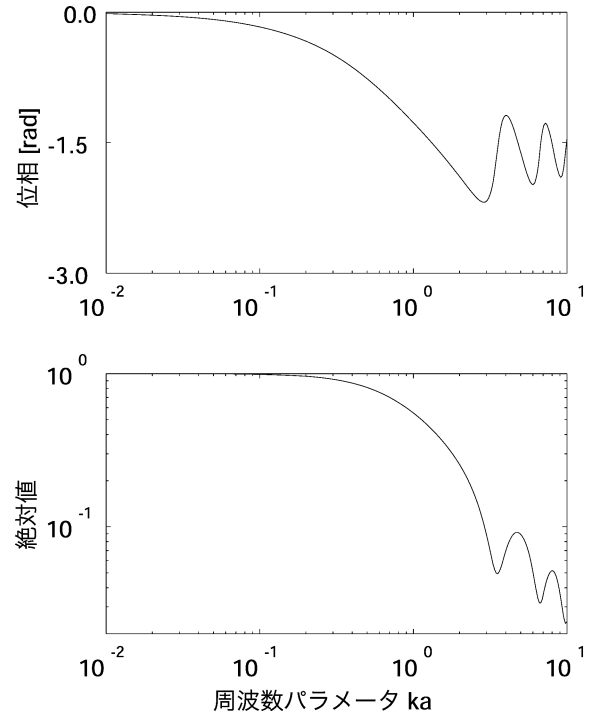


図4 平面バツフルに設置された円形開口面の反射係数特性. k 波長定数, a 開口半径. 位相特性は π だけシフトして描かれている (0 Hz で反射係数の値は -1 となる).

である。理想的開放の場合 ($r_i = -1$) よりも、 $16ka/3\pi$ だけ位相がさらに遅れることになる。この位相遅れを長さ Δl の管内を往復伝搬する場合の位相遅れ ($=2k\Delta l$) とみなすと、

$$\Delta l = \frac{8a}{3\pi} \quad (12)$$

と書ける。放射インピーダンスに起因するこのような管の延長効果は、開口端補正と呼ばれる。 $ka \ll 1$ が成立しない場合は、開口端補正は周波数の増加に伴い漸減する。図5に式(9)の位相遅れから求めた開口端補正を示す。放射インピーダンスは、最も基本的な音響特徴量で開口形状とバツフルの影響を受ける。頭部を球バツフルで近似するとより実体に近い幾何学的な形状となるが、平面バツフルの場合と比較して放射インピーダンスの変動はあまり大きくない²⁾。開口部を円形ではなく、同じ面積を有する正方形として考えてもほぼ同じ特性を示すものの、長短辺の比を大きくすると放射インピーダンスの実部、虚部ともに減少する傾向がある⁵⁾。また、同じ面積を有する楕円と方形を比較すると、長短軸 (辺) の比が同程度の場合は、放射インピーダンスの差異はあまりない⁶⁾。

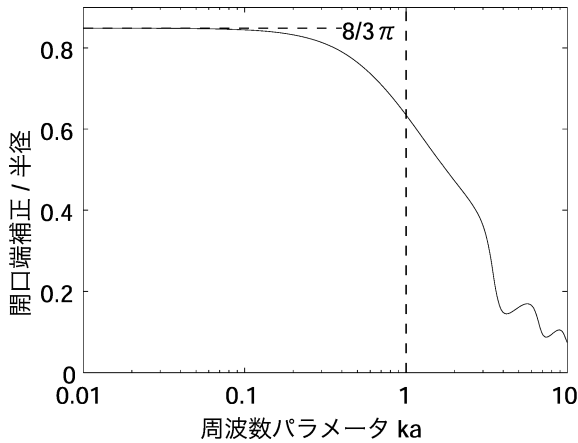


図5 円形ピストンの放射インピーダンスから求められる開口端補正. ピストン半径 a で規格化.

2.1.1 口唇開口端の反射係数の特徴

実際の口唇部はくさび形のような3次元的な広がりを有しており、波面の進行とともに音波の放射と声道への反射が連続的に生じるので、明確な声道の終端位置を決めることができない。このため声道長や放射面の断面積の定義があいまいとなり、1次元モデルを構成する上で大きな問題となる。図6のように均一音響管を口唇部の3次元的な物理模型（レプリカ）に接続し、均一管の終端（参照面）において、式(9)の反射係数が定在波法⁹⁾により測定されている^{9,10)}。参照面より先に、適切な断面積と長さを有する1区間の音響管を接続することで図7のようにレプリカで実測された反射係数の周波数特性を良好に近似することができる。このとき、口唇を表す管の断面積は実際の口唇を正面から見たときの開口面積によってほぼ近似できるものの、その長さについては口唇の幾何学的形状とよい相関は得られていない¹⁰⁾。

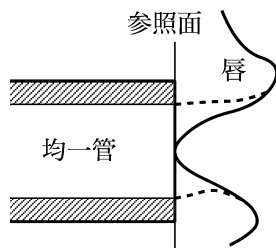


図6 口唇部における反射係数測定. 均一管の終端を参照面として、この面上の反射係数を測定する.

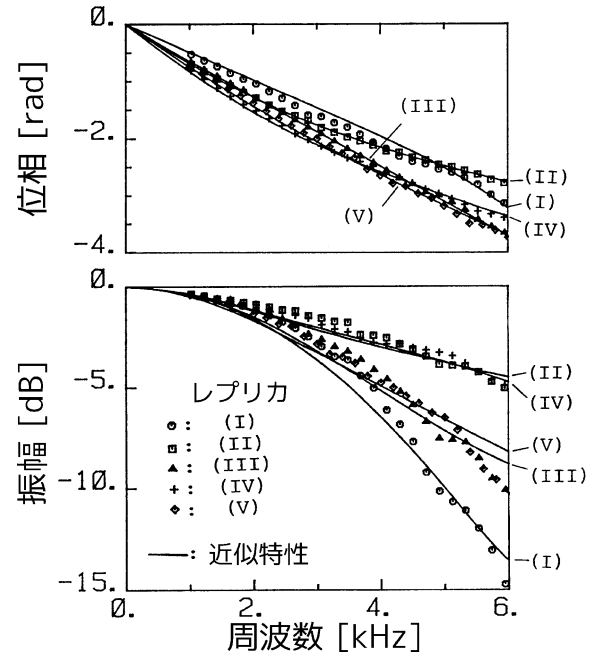


図7 口唇端反射係数の周波数特性. シンボル: /a/(I, III, IV, V), /u/(II)発話時のレプリカを用いた測定値, 実線: 付加的な1区間による近似特性. 位相特性は π だけシフトして描かれている (0 Hzで反射係数の値は-1となる).

2.2 伝達インピーダンス

放射された音波の指向特性は、周波数、開口形状、バツフルの影響を受けるものの、通常知覚する音声は、話者の前方で聴取するときには観測方向の違いで大きな変化はない。そのため、1次元モデルの放射過程として用いるときには、伝達インピーダンスのおおよそ周波数特性が分かればよい。平面バツフルに設置された振動板から放射された音波が、前方の位置 \mathbf{r} に発生させる音圧 $P_r(\mathbf{r})$ は次のレイリー積分により求められる。

$$P_r(\mathbf{r}) = \frac{j\rho ck}{2\pi} \iint_S \frac{V_l}{r'} e^{-jkr'} dS \quad (13)$$

V_l は振動板の振動速度、 r' は振動板上の微小面積要素から位置 \mathbf{r} までの距離である。積分は振動板面上で行う。半径 a の円形ピストン振動板の中心軸上の音圧は、振動板からの距離 z のとき次のようになる⁵⁾。

$$P_r(z) = 2\rho c V_l \sin\left\{\frac{ka}{2}(\sqrt{d^2+1}-d)\right\} \times \exp\left\{\frac{j}{2}\left(\pi - ka(\sqrt{d^2+1}+d)\right)\right\} \quad (14)$$

ここで、 $d = z/a$ としている。 $d \gg 1$ のとき、振動

板の体積速度 $U_l = \pi a^2 V_l$ を用いると,

$$P_r(z) \approx 2j\rho c V_l \sin\left(\frac{ka}{4d}\right) e^{-jkz} \approx \frac{j\omega\rho U_l}{2\pi z} e^{-jkz} \quad (15)$$

となる。これは $2U_l$ の体積速度を有する点音源が自由空間に置かれたときに z だけ離れた遠方点に発生する音圧に等しい。無限平面パツフル上の円形振動板モデルの場合は U_l の体積速度を有する振動板から自由空間の半分に放射されるので体積速度源の大きさに 2 倍の違いが生じる。式(15)より、距離 z の位置までの伝達インピーダンスは、

$$Z_t = \frac{P_r(z)}{U_l} = \frac{j\omega\rho}{2\pi z} e^{-jkz} \quad (16)$$

となる。距離に反比例する減少と位相項は、伝達インピーダンスの周波数特性に影響を与えないので、 Z_t の相対振幅特性は適当な係数 α を用いて次のように表される。

$$|Z_t| = \alpha\omega \quad (17)$$

従って、伝達インピーダンスは周波数に比例して増大する特性を持ち、6 dB/octave の 1 階微分特性として近似することができる。実際の発声状況を模擬するように等身大のマネキンを用いた測定では、30 cm³/s の一定体積速度による駆動に対して、250 Hz から 4 kHz の範囲で、口唇より 12 インチ前方では音圧変化が約 7 dB/octave であることが示されている⁷⁾。

2.3 開口端と遠方点における音圧と粒子速度

開口端での音圧 P_l と遠方点での音圧 P_r の大きさは、式(1), (2)より、

$$\left|\frac{P_l}{P_r}\right| = \left|\frac{Z_r}{Z_t}\right| \quad (18)$$

となる。開口端の音圧は、遠方での音圧と比較して、 $|Z_r/Z_t|$ 倍の周波数特性の変化があることを表している。式(4), (17)を用いて ka に対する $|P_l/P_r|$ の特性を描いたものを図 8 に示す。 $ka \ll 1$ の周波数域では、開口部と遠方点の音圧の周波数成分はほぼ等しいが、 ka の増加にともない開口部では遠方点よりスペクトルの成分が相対的に減少する。また、遠方での粒子速度を V_r とすると、遠方点での音波伝搬はほぼ平面波となるので、 $V_r = P_r/(\rho c)$ である。従って、式(1), (2), (17)より、

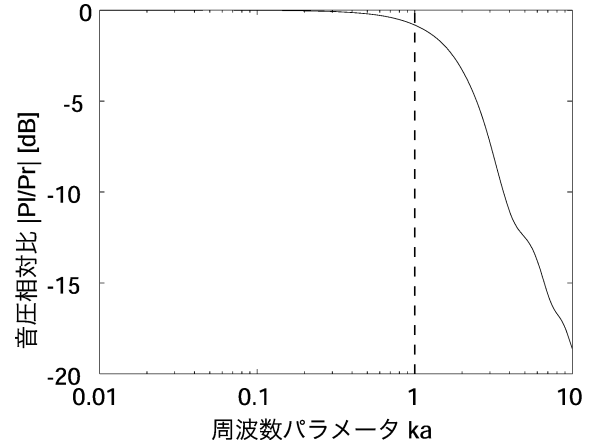


図 8 開口端音圧 P_l と遠方点音圧 P_r の比。最大値を 0 dB とした相対値。

$$\left|\frac{V_l}{V_r}\right| = \frac{\rho c}{\pi a^2} \frac{1}{|Z_t|} \propto \frac{1}{\omega} \quad (19)$$

となる。

指向性を有するマイクロホン（音圧傾度マイクロホン）は、粒子速度に比例する信号成分が出力されるため、指向性マイクロホンを用いて口唇近傍で音声を収録すると遠方で収録したときより低域が強調される。このような効果は近接音場効果として知られている。

3. 3次元放射音場の特徴と計算モデル

前節までに述べた放射過程のモデルは、声道や口唇部分の断面の寸法が音波の波長に比べて十分に小さいという平面波伝搬の仮定に基づいてる。母音の音韻性を表すにはおおよそ 3 kHz 程度までの周波数までが必要であり、このような仮定は十分に妥当なものと考えられている。一方、摩擦子音のように高い周波数成分を含む音声では、平面波伝搬の仮定は適当ではない。また、音声に含まれる個人性の特徴の一部は比較的高い周波数域にも含まれることから、高度な音声情報処理のためには、3次元的な音波伝搬を考慮したモデル化が必要となる。以下では、3次元音場の特徴計測の例とその計算モデルについて述べる。

3.1 矩形音響からの放射音圧分布の測定

平面波以外のモードを伴って音波が放射される場合の放射音場が測定されている^{11,12)}。図 9 は、(a) のように矩形断面を有する長さ 60 cm の音響管

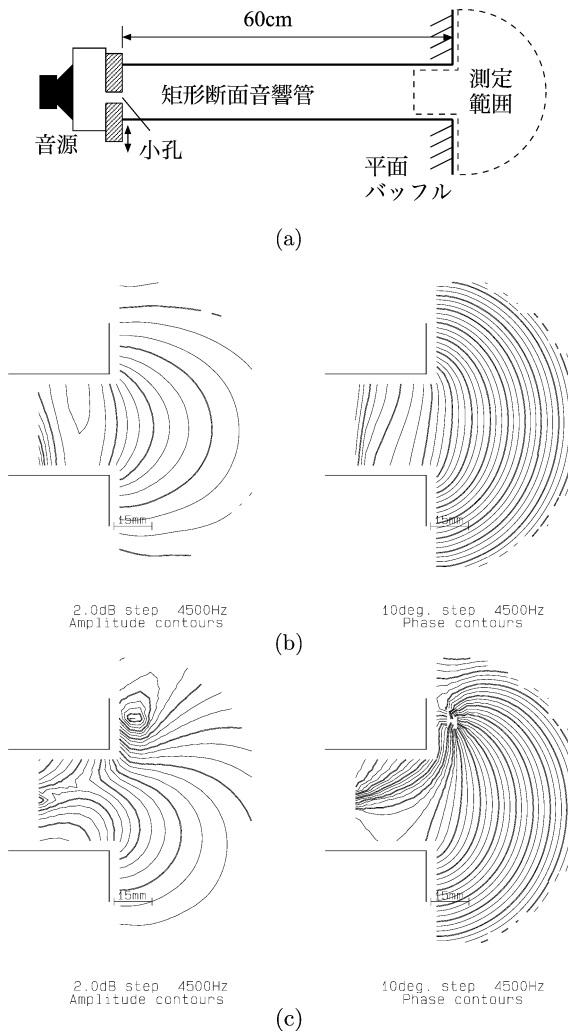


図9 バッフルにセットされた矩形音響管（断面寸法 40 mm×15 mm，長さ 40 cm）からの放射音場（4.5 kHz）. (a)測定の概略図. 小孔を開けた板の移動により高次モードの励振状態を変化させる. (b)平面波のみが放射される場合. (c)平面波と高次モードが放射される場合. (b), (c)の左図は振幅分布を右図は位相分布を示す.

の一端をバッフル面上に開口させ，他端に直径 4 mm の小穴を介してスピーカを接続して，4.5 kHz における開口端付近の音場を測定した結果である¹¹⁾. (b)は，管の中央に小孔を配置して，平面波のみを伝搬させた場合[†]，(c)は矩形管の角に小孔を配置し，奇数次の最初の高次モード（しゃ断周波数 4.3 kHz）も伝搬させた場合を表している．周波数が同じであっても駆動位置を変えることで矩形音響管内部の高次モードの励振状態が変化する．外部空間に音圧の低下する点とその周りで位

相回転が見られる．位相回転の中心では位相が不連続となるため，その点での音圧は完全に 0 となる．高次モードが伝搬モードとなると，平面波成分と高次モード成分の両者が放射され，開口形状が同一であっても放射域に生じる音場は複雑なものとなる．このような状態では，放射される音響パワーは，ピストン振動板モデルから求められる放射パワーと大きく異なる．高い周波数域では声道の共振特性を定める放射過程のモデルについても検討が必要となる．

3.2 3次元音場の表現

3.2.1 基本式と波動方程式

3次元音場では，音圧 p と粒子速度 \mathbf{v} に関する連続の式と運動方程式は次のようになる．

$$\frac{\partial p}{\partial t} = -K \nabla \cdot \mathbf{v} \quad (20)$$

$$\nabla p = -\rho \frac{\partial \mathbf{v}}{\partial t} \quad (21)$$

ここで， K は体積弾性率， ∇ は $\mathbf{e}_x, \mathbf{e}_y, \mathbf{e}_z$ を x, y, z 方向の単位ベクトルとして，次の演算子ベクトルを表す．

$$\nabla = \mathbf{e}_x \frac{\partial}{\partial x} + \mathbf{e}_y \frac{\partial}{\partial y} + \mathbf{e}_z \frac{\partial}{\partial z} \quad (22)$$

速度ポテンシャル ϕ を

$$\mathbf{v} = -\nabla \phi \quad (23)$$

により定義すると，

$$p = \rho \frac{\partial \phi}{\partial t} \quad (24)$$

となるのは 1次元の場合と同じである．これらの関係を式(21)に代入すると， ϕ に関する波動方程式が得られる．

$$\nabla^2 \phi = \frac{1}{c^2} \frac{\partial^2 \phi}{\partial t^2} \quad (25)$$

ただし， ∇^2 は直交座標系では次の演算を表す．

[†] 中央に音源があるので励振されるのは偶数次のモードのみ．最初の偶数次の高次モードのしゃ断周波数は 8.7 kHz なので管内は平面波のみが伝搬する．

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (26)$$

時間因子を $e^{j\omega t}$ として、 $\phi = \Phi e^{j\omega t}$ とすると、空間分布 Φ に関するヘルムホルツ方程式が得られる。

$$\nabla^2 \Phi + \left(\frac{\omega}{c}\right)^2 \Phi = 0 \quad (27)$$

また、音圧、粒子速度の空間分布を P, \mathbf{V} とすると、 $p = \text{Re}\{P e^{j\omega t}\}$ 、 $\mathbf{v} = \text{Re}\{\mathbf{V} e^{j\omega t}\}$ と書けるので、式(23)、(24)を用いて Φ から次のように求められる。

$$P = j\omega\rho\Phi \quad (28)$$

$$\mathbf{V} = -\nabla\Phi \quad (29)$$

3.3 矩形音響管に関する解と高次モード

式(27)は、円筒管や矩形音響管に対しては比較的容易に解を求めることができ、その内部音場を平面波と高次モードの和として表現することができる。矩形管は面積と縦横比を様々に選ぶことができるので、円筒管よりも自由度が高く、声道断面の形状近似として利用することが考えられる。

図 10 のように管軸を z 軸とした xyz 座標系を用いて波動方程式の解を求める。矩形管断面の縦、横寸法を L_x, L_y とし、 $x=0, L_x, y=0, L_y$ に剛壁とみなせる境界があるものとする。

管内部の速度ポテンシャル $\Phi(x, y, z)$ を位置 x, y, z に関して変数分離された関数 $\Phi_1(x), \Phi_2(y), \Phi_3(z)$ の積、

$$\Phi(x, y, z) = \Phi_1(x)\Phi_2(y)\Phi_3(z) \quad (30)$$

で表されるとすると、 $\Phi_1(x)\Phi_2(y)\Phi_3(z) \neq 0$ の場合には、式(30)を式(27)に代入すると次の関係が得られる。

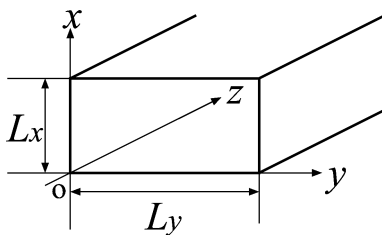


図 10 矩形音響管内の座標系

$$\frac{1}{\Phi_1(x)} \frac{d^2\Phi_1(x)}{dx^2} + \frac{1}{\Phi_2(y)} \frac{d^2\Phi_2(y)}{dy^2} + \frac{1}{\Phi_3(z)} \frac{d^2\Phi_3(z)}{dz^2} = -k^2, \quad k = \frac{\omega}{c} \quad (31)$$

$\Phi_1(x)$ に着目すると、

$$\frac{1}{\Phi_1(x)} \frac{d^2\Phi_1(x)}{dx^2} = -k^2 - \frac{1}{\Phi_2(y)} \frac{d^2\Phi_2(y)}{dy^2} - \frac{1}{\Phi_3(z)} \frac{d^2\Phi_3(z)}{dz^2} \quad (32)$$

となるが、式(32)は変数分離されているので、右辺は x, y, z に無関係な定数となる。これを γ_x^2 とおくと、 a_x, b_x を境界条件から定まる定数として、次の解が得られる。

$$\Phi_1(x) = a_x e^{-\gamma_x x} + b_x e^{\gamma_x x} \quad (33)$$

式(33)は、 $\Phi_1(x) = 0$ の場合にも式(27)を満たす。 $\Phi_2(y), \Phi_3(z)$ についても同様に次の解が得られる。

$$\Phi_2(y) = a_y e^{-\gamma_y y} + b_y e^{\gamma_y y} \quad (34)$$

$$\Phi_3(z) = a_z e^{-\gamma_z z} + b_z e^{\gamma_z z} \quad (35)$$

これらの解を式(27)に代入すると、次の関係が得られる。

$$\gamma_x^2 + \gamma_y^2 + \gamma_z^2 + k^2 = 0 \quad (36)$$

管軸方向 (z 方向) への伝搬定数 γ_z は、 γ_x, γ_y により表されるので、管壁での境界条件を満たす γ_x, γ_y に関する総和を取ることで、矩形音響管内部の速度ポテンシャル $\Phi(x, y, z)$ の一般解が得られる。

3.3.1 管壁の境界条件

ある $x-y$ 断面で、 y, z を固定し、音圧 P と粒子速度の x 方向成分 V_x を求めると、式(28)、(29)より、

$$P = j\omega\rho(a_x e^{-\gamma_x x} + b_x e^{\gamma_x x})\Phi_2(y)\Phi_3(z) \quad (37)$$

$$V_x = -\frac{\partial\Phi}{\partial x} = \gamma_x(a_x e^{-\gamma_x x} - b_x e^{\gamma_x x})\Phi_2(y)\Phi_3(z) \quad (38)$$

となる。管壁は剛壁なので、 $x=0$ では壁面に垂直な粒子速度は 0 となるから、 $a_x = b_x$ である。また、 $x=L_x$ においても $V_x=0$ なので、

$$\sinh(\gamma_x L_x) = 0 \quad (39)$$

より、次のようになる。

$$\gamma_x = j \frac{m\pi}{L_x}, \quad m=0, 1, 2, \dots \quad (40)$$

同様にして、 y 方向についても、

$$\gamma_y = j \frac{n\pi}{L_y}, \quad n=0, 1, 2, \dots \quad (41)$$

となる。 $m, n=0, 1, 2, \dots$ に対応する γ_z を γ_{mn} と書くと、式(36)より、

$$\gamma_{mn} = \sqrt{\left(\frac{m\pi}{L_x}\right)^2 + \left(\frac{n\pi}{L_y}\right)^2 - k^2} \quad (42)$$

となる。 γ_{mn} をモード (m, n) の伝搬定数と呼ぶ。モード $(0, 0)$ は平面波を表し、平面波以外のモードは高次モードと呼ばれる。モードのインデックスである m, n に関して可能な和をとると、矩形管内部の速度ポテンシャルが次のように表される。

$$\Phi(x, y, z) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} (a_{mn} e^{-\gamma_{mn} z} + b_{mn} e^{\gamma_{mn} z}) \Psi_{mn}(x, y) \quad (43)$$

ここで、 a_{mn}, b_{mn} は、 z 方向の境界条件から定まるモード (m, n) の複素振幅を表す定数である。 $\Psi_{mn}(x, y)$ は $x-y$ 断面上の速度ポテンシャルの空間分布形状を表す実関数で、固有振動姿態と呼ばれ、次のようにおいている。

$$\Psi_{mn}(x, y) = \frac{\cos(m\pi x/L_x) \cos(n\pi y/L_y)}{\sqrt{L_x L_y} \sigma_m \sigma_n} \quad (44)$$

ただし、 σ_m, σ_n は、 m, n が0のとき1、それ以外で1/2とする。これは、式(43)が $x-y$ 面上の分布を Ψ_{mn} で展開する表現となっているので、次の関係が成立するように係数を設定しているためである。

$$\int_0^{L_y} \int_0^{L_x} \Psi_{mn}(x, y) \Psi_{\mu\nu}(x, y) dx dy = \begin{cases} 1 & m=\mu, n=\nu \\ 0 & \text{上記以外} \end{cases} \quad (45)$$

従って、音圧 P と z 軸方向の粒子速度 V_z は、次のようなモード展開により表すことができる。

$$P(x, y, z) = j\omega\rho\Phi = j\omega\rho \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} (a_{mn} e^{-\gamma_{mn} z} + b_{mn} e^{\gamma_{mn} z}) \Psi_{mn}(x, y) \quad (46)$$

$$V_z(x, y, z) = -\frac{\partial\Phi}{\partial z} = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \gamma_{mn} (a_{mn} e^{-\gamma_{mn} z} - b_{mn} e^{\gamma_{mn} z}) \Psi_{mn}(x, y) \quad (47)$$

各モードについて、音圧と z 軸方向粒子速度の成分の比をとると、

$$Z_{cmn} = \frac{j\omega\rho}{\gamma_{mn}} \quad (48)$$

となる。 Z_{cmn} をモード (m, n) の特性インピーダンスと呼ぶ。なお、平面波については、 $\gamma_{00} = jk$ 、 $Z_{c00} = \rho c$ となる。

3.3.2 伝搬モードとエバネッセントモード

モード (m, n) の伝搬定数は、平面波以外では周波数の関数となる。伝搬定数の虚数部(位相定数)が0となる周波数をしゃ断周波数という。モード (m, n) のしゃ断周波数 $f_{c,mn}$ は、式(42)より、

$$f_{c,mn} = \frac{c}{2\pi} \sqrt{\left(\frac{m\pi}{L_x}\right)^2 + \left(\frac{n\pi}{L_y}\right)^2} \quad (49)$$

となる。しゃ断周波数より高い周波数では、伝搬定数は虚数となるので、管断面上で式(44)の分布を保ったまま、高次モードは z 軸方向に伝搬する。このようなモードを伝搬モードという。一方、しゃ断周波数より低い周波数では伝搬定数が実数(減衰定数)のみとなるので、 z 軸方向への位相変化はなく指数関数的に振幅が減衰する。このようなモードをエバネッセントモードといい、管の断面形状が変化するような部位で、境界条件を満足するために局所的に存在する成分と考えることができる。エバネッセントモードに対する特性インピーダンスは、式(48)より虚数となる。

3.4 高次モードを含む放射の表現

式(46), (47)において、加算するモードを適当な数で打ち切ると次のように表現できる。

$$P(x, y, z) \approx \Psi^T \{ \mathbf{D}(-z) \mathbf{a} + \mathbf{D}(z) \mathbf{b} \} \\ V_z(x, y, z) \approx \Psi^T \mathbf{Y}_c \{ \mathbf{D}(-z) \mathbf{a} - \mathbf{D}(z) \mathbf{b} \} \quad (50)$$

T は転置を表す。ここで、 $\mathbf{a}, \mathbf{b}, \mathbf{Y}_c, \mathbf{D}(z)$ はモードインデックス (m, n) を適当な順番で $(0, 0)$ から (p, q) まで並べて、次のように定義したベクトルと行列である。

$$\mathbf{a} = j\omega\rho \begin{bmatrix} a_{00} \\ a_{01} \\ \vdots \\ a_{pq} \end{bmatrix}, \quad \mathbf{b} = j\omega\rho \begin{bmatrix} b_{00} \\ b_{01} \\ \vdots \\ b_{pq} \end{bmatrix}, \quad \Psi = \begin{bmatrix} \Psi_{00}(x, y) \\ \Psi_{01}(x, y) \\ \vdots \\ \Psi_{pq}(x, y) \end{bmatrix}$$

$$\mathbf{Y}_c = \frac{1}{j\omega\rho} \text{diag}[\gamma_{00}, \gamma_{01}, \dots, \gamma_{pq}],$$

$$\mathbf{D}(z) = \text{diag}[e^{\gamma_{00}z}, e^{\gamma_{01}z}, \dots, e^{\gamma_{pq}z}]$$

\mathbf{a}, \mathbf{b} は管の両端境界条件から定める複素定数ベクトル、 \mathbf{Y}_c は各モードの特性インピーダンスの逆数 (モードの特性アドミタンスと呼ぶ) からなる対角行列、 $\mathbf{D}(-z)$ は伝搬モードの位相遅れ又は、エバネッセントモードの減衰を表す対角行列である。また、式(4)より、 \mathbf{E} を単位行列として、管断面 S 上の積分は、

$$\iint_S \Psi \Psi^T dx dy = \mathbf{E} \quad (51)$$

である。いま、 $\mathbf{P}(z), \mathbf{V}_z(z)$ をそれぞれ、モード音圧、モード粒子速度として次のように定義する。

$$\mathbf{P}(z) = \mathbf{D}(-z)\mathbf{a} + \mathbf{D}(z)\mathbf{b}$$

$$\mathbf{V}_z(z) = \mathbf{Y}_c \{\mathbf{D}(-z)\mathbf{a} - \mathbf{D}(z)\mathbf{b}\} \quad (52)$$

$\mathbf{P}(z)$ と $\mathbf{V}_z(z)$ の各成分を位置 z における電圧、電流に対応させると、各モードには独立した伝送線路が電氣的等価回路として対応する。エバネッセントモードに対しては、特性インピーダンスが虚数、伝搬定数が実数となる特殊な線路が対応するものとする。 $\mathbf{P}(z), \mathbf{V}_z(z)$ が定まると、音響管内の3次元音場分布は、

$$P(x, y, z) \approx \Psi^T(x, y) \mathbf{P}(z)$$

$$V_z(x, y, z) \approx \Psi^T(x, y) \mathbf{V}_z(z) \quad (53)$$

により求めることができる。1次元の等価回路である2端子対の縦続行列と同様に、 $z=0$ と $z=L_z$ における $\mathbf{P}(z), \mathbf{V}_z(z)$ の関係は、式(52)より次の縦続行列 $\mathbf{F}(L_z)$ で表現できる。

$$\begin{bmatrix} \mathbf{P}(0) \\ \mathbf{V}_z(0) \end{bmatrix} = \mathbf{F}(L_z) \begin{bmatrix} \mathbf{P}(L_z) \\ \mathbf{V}_z(L_z) \end{bmatrix} \quad (54)$$

$$\mathbf{F}(L_z) = \frac{1}{2} \begin{bmatrix} \mathbf{D}(L_z) + \mathbf{D}(-L_z) & \mathbf{Y}_c^{-1} \{\mathbf{D}(L_z) - \mathbf{D}(-L_z)\} \\ \mathbf{Y}_c \{\mathbf{D}(L_z) - \mathbf{D}(-L_z)\} & \mathbf{D}(L_z) + \mathbf{D}(-L_z) \end{bmatrix} \quad (55)$$

振動要素から自由空間への放射は式(13)により表されるから、開口面上の音圧 $P(x, y)$ は次のようになる。

$$P(x, y) = \frac{jk\rho c}{2\pi} \int_0^{L_y} \int_0^{L_x} V_z(x', y') \frac{e^{-jkr}}{r} dx' dy' \quad (56)$$

$$r = \sqrt{(x-x')^2 + (y-y')^2}$$

開口面を $z=0$ として、式(50)を式(56)に代入すると、

$$\Psi^T \{\mathbf{a} + \mathbf{b}\} = \frac{jk\rho c}{2\pi} \iint_S \Psi^T \mathbf{Y}_c \{\mathbf{a} - \mathbf{b}\} \frac{e^{-jkr}}{r} dx' dy' \quad (57)$$

左から Ψ を乗じて開口面 S 上で積分し、モード音圧 \mathbf{P} とモード粒子速度 \mathbf{V}_z を用いると、

$$\mathbf{a} + \mathbf{b} = \mathbf{Z}_r \mathbf{Y}_c \{\mathbf{a} - \mathbf{b}\}$$

$$\mathbf{P} = \mathbf{Z}_r \mathbf{V}_z \quad (58)$$

ここで、 \mathbf{Z}_r は次のように定義している。

$$\mathbf{Z}_r = \frac{jk\rho c}{2\pi} \iint_S \left(\iint_S \Psi \Psi^T \frac{e^{-jkr}}{r} dx' dy' \right) dx dy \quad (59)$$

式(59)より明らかなように、 \mathbf{Z}_r は相反性を有する対称行列である。 \mathbf{Z}_r の各要素はモード (m, n) の音圧とモード (μ, ν) の粒子速度のモード間の相互放射インピーダンス $Z_{mn, \mu\nu}$ を表しており、

$$Z_{mn, \mu\nu} = \frac{jk\rho c}{2\pi} \int_0^{L_y} \int_0^{L_x} \left(\int_0^{L_y} \int_0^{L_x} \Psi_{mn}(x, y) \Psi_{\mu\nu}(x', y') \frac{e^{-jkr}}{r} dx' dy' \right) dx dy \quad (60)$$

となる¹³⁾。1次元モデルで境界条件として用いた放射インピーダンス Z_r は、 $Z_{00,00}$ に開口面積を乗じたものである。開口部での反射過程は \mathbf{a} から \mathbf{b} への変換として $\mathbf{b} = \mathbf{R}_r \mathbf{a}$ として表すと、式(58)より、 $\mathbf{Z}_c = \mathbf{Y}_c^{-1}$ として次の関係が得られる。

$$\mathbf{R}_r = \mathbf{Z}_c (\mathbf{Z}_r + \mathbf{Z}_c)^{-1} (\mathbf{Z}_r - \mathbf{Z}_c) \mathbf{Z}_c^{-1} \quad (61)$$

\mathbf{R}_r により、放射端における反射に伴う各モードの振幅変化とモード変換が表される。これにより、声道の終端条件としての高次モードを含む放射過程を表現することができる。

3.4.1 計算例

図11に歯列咬合を模擬するように放射端の近くに障害物を配置した形状モデルの音圧分布を示す。音源は左上角に微小音源を配置している。1,

4, 8 kHzにおける音圧分布(振幅分布, 及び相対位相分布)を示す。(a)の各図は終端に無反射終端を与えた場合, (b)は式(59)の高次モードに対する放射インピーダンス行列を与えた場合である。この形状モデルでは, 最初の高次モードのしゃ断周波

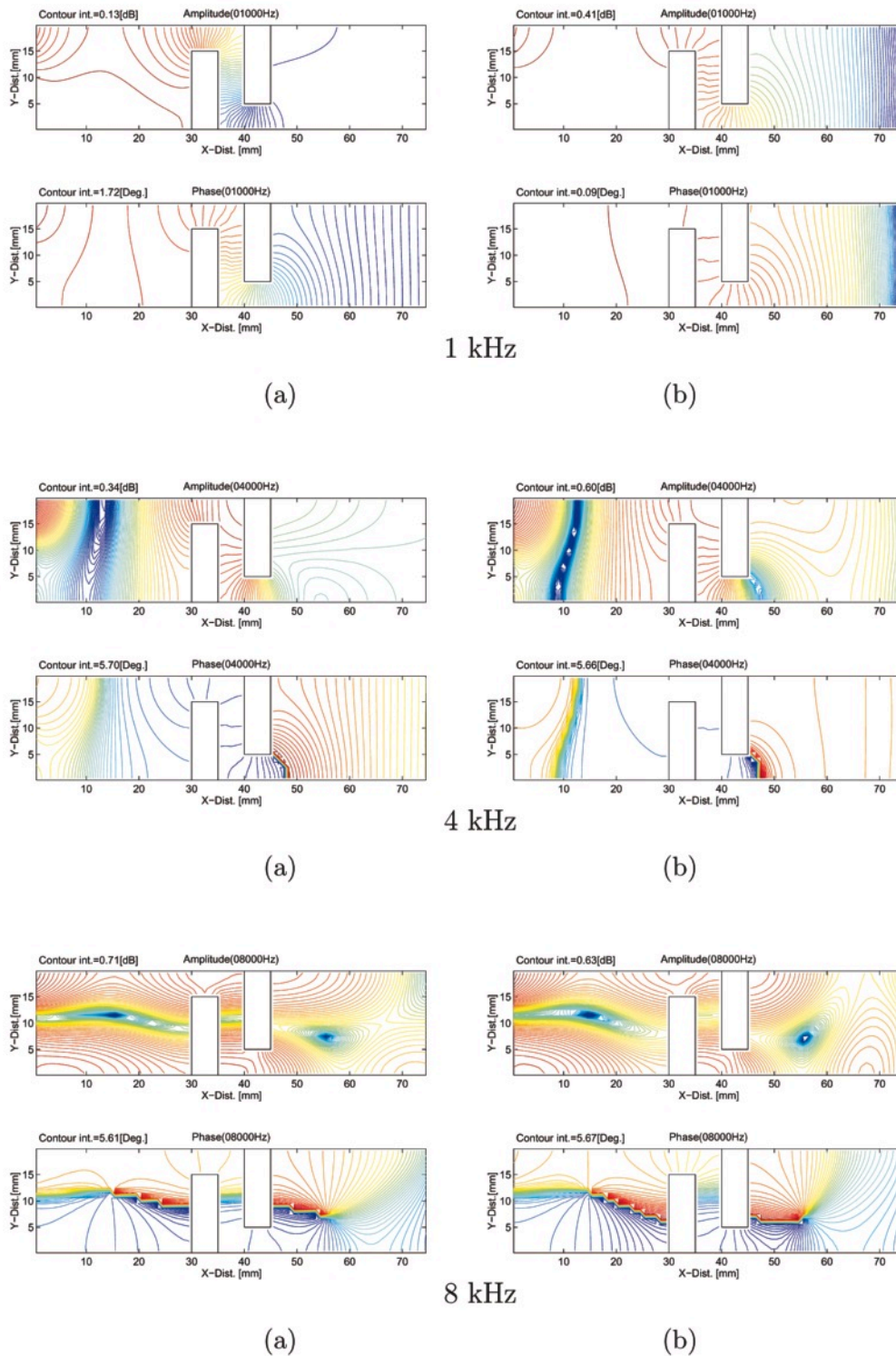


図11 歯列咬合を模擬したスリット間の1, 4, 8 kHzにおける音圧分布。(a)終端を無反射とした場合(b)高次モードの放射を考慮した場合。各図の上が振幅分布, 下が位相分布。

数は 8.57 kHz であるから、これらの周波数における音圧分布に寄与している高次モードは、全てエバネッセントモードである。特に障害物に挟まれた第 3 区間では、エバネッセントモードを考慮することで、上下方向への波面の伝搬が表現されている。1 kHz の結果は、終端インピーダンスの違いをよく表している。無反射終端の場合には、最終区間では反射波がないため振幅分布はほぼ一様であり、終端に近づく程、位相の等高線は等間隔に近くなる。放射インピーダンス行列を与えた場合には、放射端での反射のため振幅等高線が密になる。開口端でのモード結合により、開口面中央付近の音圧がやや高くなっていることが分かる。8 kHz では、最初の高次モードのしゃ断周波数に近くなるため、第 1 区間では、上下方向に強い定在波が存在するような音圧分布となっている。一般に周波数が高くなると放射に伴う反射波が減少し放射効率が増す。放射インピーダンス行列は、無反射終端条件に近づくから、8 kHz における(a)と(b)の音圧分布はかなり類似したものとなっている。また、第 1, 5 区間で位相回転が見られ、その中心では位相が不連続となるから、音圧振幅は 0 となる。このような現象は、矩形音響管からの放射音場に関する音圧分布計測で得られたものと同様である。

4. おわりに

本稿では、音声生成系における放射過程の特徴を述べた。また、放射特性に関する実測データをいくつか示し、3次元放射モデルについて説明した。最後の計算例に示したように、声道部分についても矩形音響管の構造に基づいてその音響特性を、高次モードをパラメータとして表現することができる。3次元的な取り扱いが必要な場合は、有限要素法 (FEM) などの数値計算手法を用いるのが通常であるが、1次元モデルの拡張として高次モードを用いて3次元音場をパラメトリックに表現するモデルは計算量が非常に少ない。このようなパラメトリックな計算手法は、詳細な音響特徴を解析できる FEM に比べると形状表現能力は低いものの、少ない計算量で発声器官の3次元形状に起因する効果、例えば高域に生じる多数の極や零の出現を表すことができる^{14,15)}。従来から用いられている1次元音響管モデルは、壁振動や空気の損失の効果を導入して実声道の近似性能を高

めているが、高次モードに対しても同様に損失項の導入が進められている¹⁶⁾。今後は、計算量が少ない特徴を活かして音声合成などのアプリケーションへの適用について検討していく予定である。

【参考文献】

- 1) G. Fant: *Acoustic theory of speech production*, Mouton, 1960.
- 2) J. L. Flanagan: *Speech Analysis Synthesis and Perception* (2nd Expnd ed.), Springer-Verlag, 1972.
- 3) P. M. Morse and K. U. Ingard: *Theoretical acoustics*, Chap.7, McGraw-Hill, 1968.
- 4) 森口繁一, 他: 数学公式Ⅲ, 岩波書店, 1960.
- 5) 早坂寿雄, 吉川昭吉郎: 音響振動論, 19章, 丸善, 1974.
- 6) X. Pelorson, P. Badin, K. Motoki, N. Miki and M. Plicque: On the radiation of sound at the lips during speech. Effects of lips geometry and of higher acoustical modes, Proc. 15th Int. Cong. on Acoustics, 4, pp. 497-500, 1994.
- 7) J. L. Flanagan: Analog measurements of sound radiation from the mouth, J. Acoust. Soc. Am., **32**, 12, pp.1613-1620, 1960.
- 8) JIS A 1405-1, 音響管による吸音率及びインピーダンスの測定 — 第1部: 定在波比法 (2007) [ISO 10534-1の翻訳].
- 9) K. Motoki, N. Miki and N. Nagai: Measurement of radiation characteristics using replicas of the lips, J. Acoust. Soc. Jpn. (E), **9**, 3, pp.123-130, 1988.
- 10) P. Badin, K. Motoki, N. Miki, D. Ritterhaus and M. -T. Lallouache: Some geometric and acoustic properties of the lip horn, J. Acoust. Soc. Jpn. (E), **15**, 4, pp.243-253, 1994.
- 11) K. Motoki and N. Miki: Effects of higher-order modes on sound radiation characteristics at mouth opening, Proc. 3rd Joint Meeting ASA and ASJ, 2pSC42, pp.895-898, 1996.
- 12) 元木邦俊: 放射過程におけるモード結合を考慮した音響放射パワーの計測法, 北海学園大学工学部研究報告, **25**, pp.355-370, 1998.
- 13) R. T. Muehleisen: Reflection, radiation, and coupling of higher order modes at discontinuities in finite length rigid walled rectangular ducts, Ph. D thesis, Pennsylvania State Univ., 1996.
- 14) K. Motoki and H. Matsuzaki: Computation of the acoustic characteristics of vocal-tract models with geometrical perturbation, Proc. 2004 Int. Conf. on Spoken Language Processing (INTERSPEECH2004-ICSLP), TuB602p.16, pp.521-524, 2004.

- 15) 元木邦俊, 松崎博季: 幾何学的な摂動を含む3次元声道モデルの音響解析法, 第22回ファジーシステムシンポジウム, 6A2-1, pp.11-14, 2006.
- 16) 元木邦俊, 松崎博季: 3次元声道モデルにおける高次モードの伝搬損失について, 日本音響学会秋季研究発表会, 2-P-33, pp.429-430, 2008.