

論文の内容の要旨

論文題目 下方倍音知覚の動力学

氏名 本條晴一郎

ひとは会話をするとき、あるいは音楽を聴くとき、音の高さを重要な情報として知覚している場合が多い。音高、つまり、ピッチは、空気の振動波形が周波数成分を一つしか含まない正弦波の場合にはその周波数と対応することが知られている。ところが、複数の周波数成分を含む複合音の場合には、周波数成分としては含まれていない低音が下方倍音として聞こえることが聴覚心理学実験から知られている。つまり、複合音の知覚においては、ピッチと周波数の対応関係は自明ではない。本研究は、下方倍音の知覚がいかになされているかを研究の対象とした。

本研究では、まず聴覚心理学実験を整理し、注目すべき問題を二つに分類した。一つは、下方倍音のピッチの値を求めるピッチシフトの問題である。聴覚系に入力される信号に含まれる全ての周波数成分を一定値シフトしたとき、下方倍音のピッチがどれくらいシフトするかを定量的に求めることが課題となる。もう一つは、下方倍音が聞こえないときの知覚の状態と、聞こえるときの知覚の状態がどのような関係にあるかという問題である。聴覚心理学実験においては、単一の周波数成分しか持たない純音からは下方倍音が知覚されないにも関わらず、ノイズの存在下では純音からも下方倍音が知覚されることが知られている。また、下方外音が知覚されるか否かには、個人差があることが見出されている。こちらの問題では、下方倍音を知覚できない分析モードと、知覚できる合成モードの間で、どのような仕組みで切り替わりが実現されるかが明らかにすべき課題となる。本研究では、以上のような分類をした上で、後者の問題を解明することを目的とした。

ピッチシフトの問題については、数理的な先行研究が存在している。そこで、分析モードと合成モードの切り替わりの問題を考える基盤として、まず、合成モードによって生み出されるピッチシフトの問題がどのように解決されているかを調べた。Cartwrightらは聴覚系が高調波成分を持つ自励振動子であると仮定し、その同期現象を

考察することにより下方倍音として知覚される周波数を定量的に説明した (J. H. E. Cartwright, D. L. González and O. Piro, Phys. Rev. Lett. **82**, 538 (1999)). 本研究では、Cartwright らの扱いが普遍的なものではなく、聴覚系の素子としての性質が個別的に反映していることを見出した。個別性は、自励振動子が周波数ロッキングを起こす際、入力信号を近似的に何次の高調波として捉えるかに表れていた。この個別性への注目により、聴覚系のデフォルト状態は下方倍音が知覚されない分析モードではなく、Cartwright らが扱う合成モードであることが見出された。このことから、分析モードと合成モードの切り替わりを考えるためには、「入力が複合音の場合、なぜ下方倍音が知覚されるか」ではなく、「入力が純音の場合、なぜ下方倍音が知覚されないか」が問われるべき問題であることがわかった。

本研究では、この問題に答えるため、デフォルト状態で合成モードであり、適応によって分析モードを実現するモデルを構築した。

まず、聴覚心理学実験の結果と、Cartwright らの結果を検討することにより、合成モードを表す基本モデルとして、神経細胞の発火を表す FitzHugh-Nagumo モデルがふさわしいことを見出した。なぜなら、合成モードを実現するには、非線形自励振動子である必要があるが、入力の信号がないときには、音が知覚されるべきではないからである。このことから、入力の信号によっては振動系の性質を持つ興奮系が、合成モードを表す基本モデルとして適切であることがわかった。

次に、FitzHugh-Nagumo モデルの神経細胞モデルとしての成り立ちを吟味し直し、FitzHugh-Nagumo モデルの元になっている Hodgkin-Huxley 方程式に対してなされた修正を、FitzHugh-Nagumo モデルに取り組んでモデルの拡張を行った。導出したモデルは、

$$\begin{cases} \frac{dx}{dt} = c \left(y + zx - \frac{x^3}{3} \right) - I \\ \frac{dy}{dt} = -\frac{x - a - by}{c} \\ \frac{dz}{dt} = \gamma(\alpha x + \beta - z) \end{cases} \quad (1)$$

である。変数 x 、 y が興奮系として振る舞い、入力 I を受けるとともに非線形性 z が低くなる方向に適応する。 α 、 β 、 γ は定数である。第三式で、 $dz/dt = 0$ を考えると、 z は時間スケール γ で、 $z = \alpha x + \beta$ に緩和することがわかる。モデルは、周期刺激を受けたヤリイカの巨大軸索が、単一の大きなスパイクを発した後は小さなスパイクしか生み出さないという実験事実を踏まえ、同じ性質を満たすように導出された。この結果、モデルでは、ダイナミクスの振幅情報が系そのものの性質を変えするという構造を持つこととなった。

スパイクの大きさが減少するように適応するモデルは、適応することによって高次の周波数ロッキングを実現できなくなることが見出された。このことは、元々合成モードであった FitzHugh-Nagumo モデルが、入力信号を受けることによって、分析モードに適応することに対応する。つまり、導出したモデルは、分析モードと合成モードの切り替えを記述した数理モデルとなっている。

導出した数理モデルの性質を調べるため、適応した状態での非線形性の強さ、つまり、適応の弱さを表すパラメータ \tilde{Z} を導入し、 $\alpha = (1.0 - \tilde{Z})(1.2)$ 、 $\beta = \tilde{Z}$ としてモデルの性質に対するパラメータ依存性を調べた。その結果は、1 の通りであり、適応の速さ γ よりも、適応の弱さ \tilde{Z} がいかなる値をもつかが重要であることがわかった。

本研究ではさらに、末梢聴覚系の数理モデルを導出した。モデルの導出は、Hudspeth らによって発表された、ウシガエルの有毛細胞から得られた実験データに注目することによって行われた (P. Martin, A. D. Mehta, and A. J. Hudspeth. PNAS, **97**, 12026 (2000)). 有毛細胞にある hair bundle が示す力学的特性に注目し、ヌルクラ

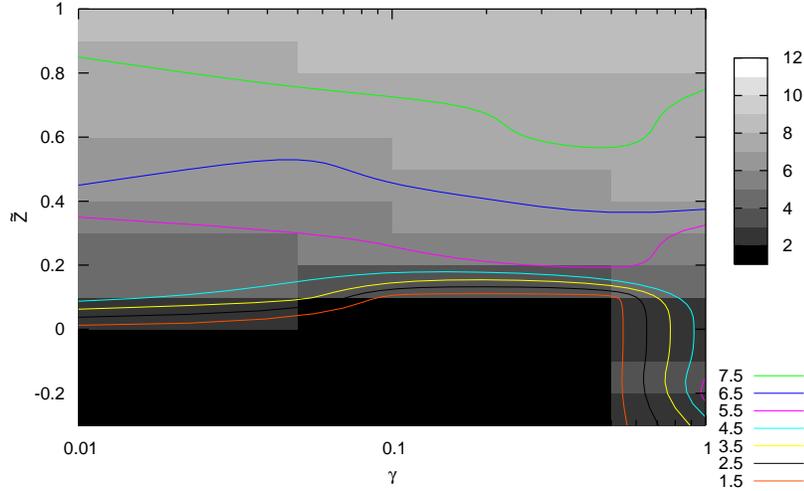


図 1: 入力として $I = A + A \sin(\omega t)$ を加えて ω を変化させたとき、ロッキングが実現する最大の次数を表す相図。
 $A = 1.5$ 。

インをデータから推定することによって、

$$\begin{cases} \frac{dx}{dt} = \epsilon_0 \left(-ax + \frac{bz}{1 + \exp(-\beta z(x - y))} - c + I \right) \\ \frac{dy}{dt} = \epsilon_1 (gx - y + h) \\ \frac{dz}{dt} = \epsilon_2 (k - ly - z) \end{cases} \quad (2)$$

が得られた。 ϵ_0 、 ϵ_1 、 ϵ_2 、 a 、 b 、 c 、 g 、 h 、 k 、 l は定数であり、 β は逆温度 $1/(k_B T)$ を表す。

導出されたモデルは、非線形性を 3 次の項までに限った場合、拡張した FitzHugh-Nagumo モデルと本質的に一致した。このことから、下方倍音の知覚は、聴覚系の様々な部位で行われている可能性が示唆された。

以上のように、本研究はで、ヌルクラインの形状変化を考えることにより、高次のロッキングが可能である合成モードの振動子が、低次のロッキングのみしか行うことのできない分析モードの振動子に適応して変化するダイナミクスを数理的に示した。このことにより、知覚の文脈依存性や個人差を考えるための新たな数理的枠組みが得ることができた。

本研究においては、生理的機構よりも知覚されるものを念頭において生理学的モデルを構築した。導出されたモデルでは振幅情報が重要な役割を果たしていた。知覚の文脈依存性や個人差を考える上では、動力学的の振幅により大きな注目を要する必要があることが示唆される。

導出されたモデルでは、適応の弱さ \bar{z} が、いかなる値をもつかが重要であった。 \bar{z} が個人によってどのように異なるのか、あるいは、注意付けやノイズの効果によってどのように変化するかについては、未検討のままである。これらの問題については、今後の課題である。