

螺旋と極座標による音高系列の表示方法の提案

竹渕 瑛一^{1,a)} 梶並 知記² 徳弘 一路³ 速水 治夫³

概要: 本論文では, 螺旋と極座標によって周波数スペクトルのうち, 音階と音高の成分を一瞥するための表示方法を提案する. 人間の蝸牛基底膜は螺旋状に巻かれており, 音の周波数によって反応する部位が異なる. 蝸牛基底膜が反応する部位は, 中心から 1/4 回転するごとに反応する周波数が倍になることがわかっている. 蝸牛基底膜の性質を参考にし, 音高成分を螺旋状に提示するシステムを実装した. 1 周ごとに 1 オクターブ上昇するように音階の成分を配置したところ, 和音を演奏した時に音階上の倍音の分布を読み取りやすくなることがわかった. 音階を螺旋状に配置することで, 音楽と倍音の関係性について議論しやすくなると考えられる.

Proposal of the display method of pitch sequences by helix and polar coordinates.

Eiichi Takebuchi^{1,a)} Tomoki Kajinami² Ichiro Tokuhiro³ Haruo Hayami³

1. はじめに

本論文では, 音高系列を極座標系から螺旋状に表示することで, 音階ごとの音高成分の分布を一瞥する手法を提案する.

周波数スペクトルは, 横軸が周波数, 縦軸に振幅を取るグラフが多く用いられている. 音楽情報処理では, 音高の成分を知りたいことがあるため音高系列 (Pitch Sequence) が用いられる [1]. 音高系列は, 周波数成分のうち, 音高周波数成分の系列である.

また, 西洋音楽の音楽理論に基づいて議論を行うには, 音階ごとの成分が極めて重要である. 音階ごとの成分を知るために, Chroma Vector[2] による音階系列 (Pitch Class) が用いられる.

本論文では音高成分を螺旋状に表示する方法を提案する. 提案手法は, 人間の蝸牛基底膜が音高を感じる部位が低い周波数から螺旋状に分布していることに着目する [3].

文献 [3] では, 低い周波数から 1 オクターブごとに 90 度ずつ回転し, 基底膜の起点からの距離と周波数は比例し, 最終的に対数螺旋となっていることが報告されている. 従来の研究 [4][5] では, 音高成分の音階ごとの分布を説明するのに円筒状のバネがイメージされていた. 実際の人間の蝸牛と同じように, 感じる部位を螺旋で表示したほうが都合が良い場合もあると考えられる.

本論文では, 2 章で音楽と螺旋の関連研究について述べ, 3 章で音高系列を極座標系で取り扱い, 周波数を複素数で扱う方法について提案する. 4 章では, 提案手法を試作システムで実装し, 周波数成分と音高系列の場合での表示方法を比較する.

2. 関連研究

西洋音階では, C,C#,D,...,A#,B と音階が続くと, さらにまた C,C#,D,...,A#,B と続く. 逆もまた同じである. 音階は輪のように循環することから, 音階を円として議論されている [4][6]. しかし, 音楽は音階だけでなく音高も重要であり, 円だけで音高を表現できない.

音階が円であるならば, 音高を円柱の高さとし, 音階と音高はコイル状であると議論する研究がある [7][8]. 音階と音高を円柱として議論するのは合理的であるが, 12 平均律から導出された音高の基本周波数と円柱の高さは一致し

¹ 神奈川工科大学大学院博士後期課程
Doctor's Course of Kanagawa Institute of Technology
² 岡山理科大学
Okayama University of Science
³ 神奈川工科大学
Kanagawa Institute of Technology
a) nanashi4129@gmail.com

ない。人間の心理学的感性に近づけるため、円錐での表現が報告されている [9]。これらの研究は、円と音高による階層構造として議論されている。一方で、どのように音高成分を音階に対応させるかが議論されていない。円柱では確かに音階に紐付いた音高成分を俯瞰できるが、円柱であることは一方の視点の反対側の成分は不可視となる。音高成分全体を俯瞰することができない。

一方で、人間の蝸牛の聴覚特性と音高の対応を対数螺旋で表現した報告がある [3]。蝸牛基底膜の周波数特性は1オクターブごとに90度ずつ回転し、 $2^{1/12}$ の周波数比で完全に再現できる。しかし、その聴覚特性図では同一角度上にある音高成分は4オクターブも離れているため、すべての音高成分を俯瞰するには読み取りづらい問題がある。

本論文で提案する手法では、文献 [3] の報告をもとに、1周で1オクターブ離れるような螺旋を描く方法を提案する。螺旋による表示では、螺旋の角度が音階に対応するとき、原点からの距離が周波数を意味する。従来手法と比較すると、音高成分と音階成分の双方を一瞥することができる。

3. 音高・音階・音程における複素数の導入

3.1 音高と音階の螺旋表示

螺旋で表示するには極座標系の導入が必要である。本節では、音高と音階をそれぞれ極座標上の距離と角度に対応させるための方法について述べる。

音高にはそれぞれ基本周波数が定められている。音高の基本周波数を式1に示す。

$$f(n) = 440 \cdot 2^{\frac{n-69}{12}} \quad (1)$$

$f(n)$ は音高の基本周波数、 n は音高のノート番号である。ノート番号は $n \in [0, 127], n \in \mathbb{R}$ を取る。440 と 69, 12 は、基準ピッチの周波数とそのノート番号、音階の数である。音高にはそれぞれノート番号が振られている。一般的にノート番号は整数であるが、本論文では実数として扱う。

式1を見て分かる通り、音高の基本周波数は指数関数である。原点からの距離を対数周波数、音階を角度として極座標で表すと等間隔螺旋の曲線となる。式2~4に音高の基本周波数と音階を極座標とした等間隔螺旋を示す。

$$r(n) = \log_2 \frac{f(n)}{f(0)} \quad (2)$$

$$\theta(n) = \frac{2n\pi}{12} \quad (3)$$

$$Q(r(n), \theta(n)) = P(r(n) \cos \theta(n), r(n) \sin \theta(n)) \quad (4)$$

$r(n)$ は原点から基本周波数に対応した対数距離であり、 $r(n) \in (0, \infty]$ を取る。 $\theta(n)$ は音階を表す角度であり、

$\theta \in [0, 2\pi]$ を取る。 Q は極座標、 P は直交座標である。 $r(n)$ は $f(0)$ を原点としているが、場合によっては $f(0)$ を無視してもよい。式2の単位は対数周波数の等間隔螺旋を示し、周波数に戻せば対数螺旋となる。

対数螺旋は拡大・縮小しても同じ形の螺旋を描くことが特徴である。しかし、高い音高になるほどその基本周波数は指数関数的に大きくなる。表示を高音に合わせてと低音が見えず、それぞれ逆もまた起こり得る。指数関数であることから、基本周波数の対数を取ることで等間隔螺旋に変換できる。

3.2 複素数の導入と音程による変化

音高と音階を極座標で表せるようになったため、複素数で極座標を表すことについて本節では考える。

あるノート番号の複素数を式5に示す。

$$z(n) = r(n)(\cos \theta(n) + i \sin \theta(n)) \quad (5)$$

n は任意のノート番号である。式5は極座標系における複素数の一般的な定義である。

音高と音階を複素数で表せるならば音程も表せる。基準となるノート番号が n のとき、その変化量を Δn としたとき、音程の変化は $n + \Delta n$ で表せる。すなわち、基準音高 n に対する音程は Δn である。

音程による音高の変化量を式6、音程による音階の変化量を式7に示す。

$$r'(n)dn = \lim_{\Delta n \rightarrow 0} \frac{r(n + \Delta n) - r(n)}{\Delta n} \quad (6)$$

$$\theta'(n)dn = \lim_{\Delta n \rightarrow 0} \frac{\theta(n + \Delta n) - \theta(n)}{\Delta n} \quad (7)$$

式6及び式7より、音高と音階からそれらの変化量を定義した。 $z(n + \Delta n)$ と置くことで、音高及び音階から音程だけ変化できる。音程は相対的であり、基準音高がなければ音程による変化を表せないため、 $z(\Delta n)$ は定義できない。従って、基準音高と音程の関係は式8のように表せる。

$$127 \geq \forall n \geq 0, \exists \Delta n \in \mathbb{R} \text{ s.t.} \quad (8)$$

$$\forall z(n) \in \mathbb{C}, n + \Delta n \Rightarrow z(n + \Delta n)$$

2つの音高から音程を求めるには、仮に2つの音高を n と n' と置くと、2つの音高から求まる音程は $n' - n = \Delta n$ であり、 $z(n') = z(n + \Delta n)$ となる。音程は上昇音程と下降音程の2種類があるため、 n と n' の間で大小の関係はない。

3.3 等間隔螺旋によるスペクトル表現

周波数成分を等間隔螺旋状に表示することを考える。単

に原点からの距離を周波数として捉えると、ただ放射状に周波数スペクトルを表示しただけであり、どの周波数と周波数成分が音階に対応しているのか非常にわかりづらい。そこで、提案手法では、周波数成分をベルト状の長いロープと捉え、原点から徐々に巻くことを提案する。

音響信号が与えられたとき、そのフーリエ変換を式 9、角周波数からノート番号と複素数への置き換えを式 10 に示す。

$$X(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x(t)e^{-i\omega t} dt \quad (9)$$

$$X(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x(t)e^{-i2\pi|z|t} dt \quad (10)$$

X は周波数スペクトル、 x は音響信号、 ω は角周波数、 t は時間である。角周波数は $\omega = 2\pi f$ であり、複素数のノルム $|z|$ が f にあたる。

4. 螺旋表示の実装

等間隔螺旋を表示するために、本研究では Siv3D を利用した [10]。開発環境は、Visual Studio 2015 であり、Siv3D のバージョンは、March2016 である。

録音には、PreSonus の AudioBox USB、エレキギターは Fender USA Stratocaster を利用した。サンプリング周波数は 44.1kHz、高速フーリエ変換の分析窓幅は 8,192 点である。窓関数にはハミング窓を利用した。

図 1 は、エレキギターで C メジャーコードを弾いた時の音響信号と周波数スペクトルである。図 2 は、螺旋で表示した音高成分である。

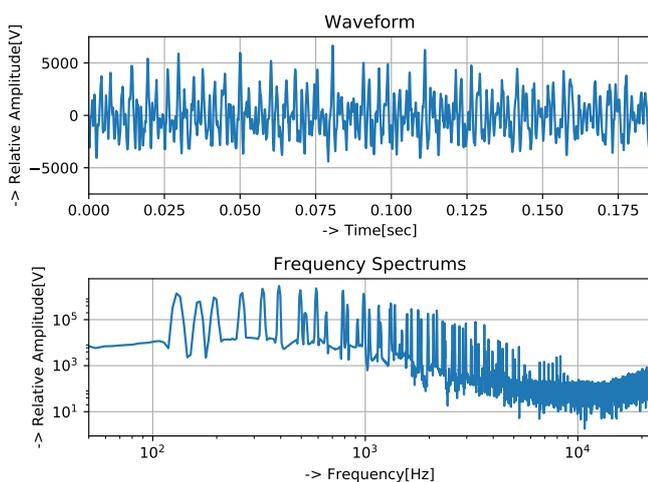


図 1 C メジャーコードの音響信号と周波数スペクトル

螺旋で表示することで、どの音階にどれだけの音高成分が分布されているのかがわかる。C メジャーコードの構成音は C, E, G であり、図中に音高成分として分布されていることがわかる。また、B は E の 3 倍音（完全五度）かつ G の 5 倍音（長三度）であるため、周波数スペクトルだ

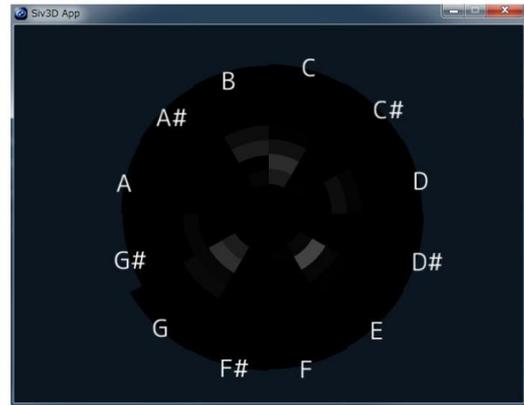


図 2 C メジャーコードの音高成分の螺旋表示

けではわかりにくい成分が表示されている。

5. おわりに

本論文では、音高系列を螺旋状に表示することで、音高成分と音階成分を一瞥するための表示方法について提案した。

音楽理論を集合論や組合せ論として論じる分野としてピッチクラス・セット理論 [11] があり、音階を整数による巡回群、和音を音階の集合としてみなし、その組み合わせから音楽を論じている。この分野では、音響工学的な知見が導入されていないため、楽器が音を出す仕組みと調律の関係によって調和の仕方が異なることまで踏み込んでいない。

今後の課題は、ピッチクラス・セット理論を周波数レベルで整理することである。例えば、弦楽器ではインハーモニシティ [12] と呼ばれる弦の材料的特性によって倍音周波数が僅かに高くなる現象が知られている。ギターの音色を語るのに倍音という言葉が多く使われるが、その響きの良さが音楽とどのように関わりがあるのか研究することが課題である。

参考文献

- [1] 竹渕瑛一, 梶並知記, 徳弘一路, 速水治夫: 音高系列の標準偏差を用いた基本周波数のピーク分類と弾弦時刻・演奏区間の抽出, 情報処理学会研究報告. MUS, 音楽情報科学, Vol. 113, No. 16, pp. 1-5 (2016).
- [2] Bartsch, M., Wakefield, G. H. et al.: To catch a chorus: Using chroma-based representations for audio thumbnailing, *Applications of Signal Processing to Audio and Acoustics*, pp. 15-18 (2001).
- [3] Cheng, K., Cheng, V. and Zou, C.-H.: A logarithmic spiral function to plot a Cochleagram, *Trends in Medical Research*, Vol. 3, No. 1, pp. 36-40 (2008).
- [4] Bartsch, M. A. and Wakefield, G. H.: Audio thumbnailing of popular music using chroma-based representations, *IEEE Transactions on multimedia*, Vol. 7, No. 1, pp. 96-104 (2005).
- [5] Shepard, R.: *Music, cognition, and computerized sound*, chapter 15, pp. 187-194, The MIT Press (1999).

- [6] Harte, C., Sandler, M. and Gasser, M.: Detecting harmonic change in musical audio, *Proceedings of the 1st ACM workshop on Audio and music computing multimedia*, ACM, pp. 21–26 (2006).
- [7] Warren, J., Uppenkamp, S., Patterson, R. D. and Griffiths, T. D.: Separating pitch chroma and pitch height in the human brain, *Proceedings of the National Academy of Sciences*, Vol. 100, No. 17, National Acad Sciences, pp. 10038–10042 (2003).
- [8] Shepard, R. N.: Structural representations of musical pitch, *The psychology of music*, pp. 343–390 (1982).
- [9] Krumhansl, C. L.: The psychological representation of musical pitch in a tonal context, *Cognitive psychology*, Vol. 11, No. 3, pp. 346–374 (1979).
- [10] 鈴木 遼, 長 幾朗ほか: ゲームとインタラクティブメディアのプログラミングを容易にする C++ ライブラリの開発, 第 77 回全国大会講演論文集, Vol. 2015, No. 1, pp. 97–98 (2015).
- [11] Forte, A., 森あかね: 無調音楽の構造: ピッチクラス・セットの基本的な概念とその考察, 音楽之友社 (2012).
- [12] Fletcher, H., Blackham, E. D. and Stratton, R.: Quality of piano tones, *The Journal of the Acoustical Society of America*, Vol. 34, No. 6, pp. 749–761 (1962).