

構成音の出現頻度に着目した楽曲の特徴解析

藤川 翔平^{1,a)} 西井 淳¹

概要: 例えば絵画では、使われている色や図形の使用順位とその頻度の関係にある一定の特徴があることが報告されている。音楽においても、より多く使う音とそうでない音を適度な比率で用いることは良い曲を作る上で重要と考えられる。そこで我々は、音楽を特徴付ける最小要素である構成音の出現および遷移の頻度に着目して楽曲のジャンルによる特徴を探った。その結果、構成音の出現頻度やその進行頻度が曲やジャンルの特徴をある程度表現していることを示唆する結果を得た。また、低出現、高出現構成音がどのような時間発展パターンに従って各曲内で出現するかを検討したところ、例えばベートーヴェンのピアノソナタ第29番の終楽章では低出現、高出現構成音が明確に区別されて計画的に楽曲中で使われており、その時間発展パターンがカオス性およびフラクタル性を示すことがわかった。一方で、ジャズ等他ジャンルの曲では全く異なる時間発展パターンも観察された。この結果は、力学的視点により構成音の使用頻度の時間発展を調べることで、楽曲の構成方法を分析・分類できることを示唆している。

An analysis of music based on the rank-frequency distribution of notes

SHOHEI FUJIKAWA^{1,a)} JUN NISHII¹

Abstract: Some studies have reported the existence of a consistent relation in the rank-distribution relation of some features in great pictures, e.g., color and shape. In writing music, some proper mixture of frequently used notes and others would be an important factor. We explored the relation between music genres and rank-distribution relation of notes and its progression. The result suggests that the relation shows the characteristic of music genre to some extent. Furthermore, we analyzed the temporal-development pattern of notes with lower and higher frequency and found that they are used systematically and shows a typical characteristics of chaos and fractals in some music pieces, e.g., the last movement of the Beethoven's piano sonata No.29. Different types of temporal-development pattern were also observed in other music pieces. These results suggest that we can classify and analyze the composition methods of music by respecting the rank-distribution relation from the view point of dynamical systems.

1. はじめに

ベートーヴェンはクラシック音楽、ビートルズはロック音楽といったように、楽曲は何かしらのジャンルに分類されている。そのジャンルを決める明確な区分はないが、我々は街なかで未知の音楽を聴いた場合でも、楽曲のジャンルをある程度同定することができる。では、クラシックらしさ、ロックらしさといった音楽のジャンルの特徴は、楽曲のどのような要素に現れるのだろうか。

音楽を特徴付ける要素にはメロディー、リズム、和声な

どがあり、いずれについても音楽理論の立場から多くの研究が行われてきている。しかし、メロディーや和声を決定する各構成音が楽曲中でどのような頻度で使われているかといった視点での研究はあまりされていない [1]。一方で、絵画などでは使われている色の頻度分布や [2]、絵の中で使われている四角形や円の大きさの順位と頻度の関係には、他の様々な自然現象でも観察されるようなある一定のパターンが存在する [3] といった報告がある。音楽においても、ジャズではテンションの高い和音が高頻度で使われる等、構成音の頻度も楽曲を特徴付ける重要な特徴のように思われる。そこで、本研究では楽曲中の各構成音の出現頻度やその進行、時空間パターンに注目して楽曲の特徴を検

¹ 山口大学大学院理工学研究科
Yamaguchi University

^{a)} fujikawa@bcl.sci.yamaguchi-u.ac.jp

表 1 主音判定の各正解率

手法	正解率 (%)
従来手法	79.17
LDA	97.92

討する。

2. データベース

本研究では、楽曲の特徴量抽出に MIDI データ 714 曲（ピアノソナタ：228 曲、クラシック：174 曲、ジャズ：106 曲、ロック：200 曲、雅楽：6 曲）を使用した。MIDI データは 128 個の音程を扱えるが、本研究では、オクターブ違いの音は同音とみなすことで [4]、128 個の音を 12 次元ベクトル（半音階ベクトル）に圧縮し、時間推移を以下で説明する楽曲の解析に用いた。

2.1 主音判定

2.1.1 解析手法

主音は、音階の起点であり、音階を代表する最も重要な音である。西洋音楽では、長調（major key）と短調（minor key）の 2 つの調性が使われることが多く、それぞれについて 12 の主音が存在するので、合わせて 24 の調が存在する。本研究では、長調の曲はハ長調へ、短調の曲はハ短調へ移調して、主音によらない楽曲の特徴解析に用いることにした。この移調を行うには、各曲の主音を判定する必要がある。

主音判定手法には様々なものがあるが、容易に考える手法としては、メロディの第 1 音や最終音を主音とみなす手法や出現頻度が最も多い音を主音とみなす手法が挙げられる [5]。しかし、全ての曲においてそれらが必ずしも主音であるという保証は無く、この単純なモデルでは実質的な正解率には到達しない。そこで、主音の他に属音の頻度も高いことを利用した次のような手法が提案されている [6]。各音の頻度をカウントし、上方完全 5 度の音程を持つペアの中で、最大頻度の音を主音とする。本研究では、この手法を主音判定における従来手法と呼ぶ。これに対し、本研究では線形判別分析（LDA: Linear Discriminant Analysis）を用いた主音判定を提案する。

LDA はクラス分類手法の 1 つであり、クラスと属性ベクトルの対の形式のデータが与えられたとき、クラス間の分散を最大化するような教師ありの次元削減法である。主成分分析などの次元削減法と違い、元のクラスを保ったまま次元を削減できることが大きなメリットである。本研究では、あらかじめ主音がわかっている曲の各音の出現頻度を訓練データとして与え、LDA を適用することで主音の判定を行った。

2.1.2 解析結果と考察

主音判定における従来手法と LDA による手法の有用性を比較した。データには各主音 4 曲の計 48 曲のクラシッ

ク音楽を使用した。テストデータにも同じ曲を使用した。比較結果を表 1 に示す。表 1 から LDA による主音判定の正解率が高いことがわかる。判定ミスをした曲は 1 曲のみで、その曲はショパンの「雨だれの前奏曲」であった。この曲は本来変ニ長調の曲であるが、変イ長調と判定されてしまった。これは、変イ音を 12 音中 25 % 強と極めて多く多用した曲であることが原因だと考えられる。

本研究における以下の解析では、曲の主音判定にこの LDA による手法を用い、ハ長調およびハ短調への移調を行った。

3. 構成音の出現頻度

3.1 解析手法

各ジャンルもしくは各曲中でどの構成音がどの程度使われているのかに着目することで、ジャンル毎（ピアノソナタ、クラシック、ジャズ、ロック、雅楽）の特徴を検討した。具体的には各ジャンルの全ての曲、もしくは各曲中に登場する全ての音の組み合わせ（半音階ベクトル）を抽出し、各半音階ベクトルの登場頻度をカウントした。その後、各曲の構成音の頻度を降順にソート、その分布（RFD: rank-frequency 分布）を曲線近似し、そのパラメータとジャンルとの関係を調べた。

RFD は次式により近似した [1]。

$$f(r) = \frac{K(R+1-r)^b}{r^a} \quad (1)$$

ここで、 $f(r)$ は順位 r の構成音の出現頻度である。 R は r の最大値で、 a, b は定数である。 K は次式で定義される。

$$K = \frac{\Gamma(b-a+2)}{\Gamma(1-a)\Gamma(1+b)} \quad (2)$$

ここで、 $\Gamma(x)$ はガンマ関数である。定数 a, b は、 $f(r)$ の値とデータ点との間のノルムが最小になるように、滑降シンプレックス法 [7] によって決定した。

3.2 解析結果と考察

図 1 は、各ジャンル毎における全ての曲から抽出した各構成音の出現頻度を示す。どのジャンルでも同じような形状のカーブを描き、構成音の使用頻度が低いと、急激に落ち込んでいることがわかる。ただし、クラシックおよびピアノソナタは、使用頻度の低い部分のカーブの形状が他ジャンルと比べなだらかであり、また、高順位の使用頻度が高い。そのため、他ジャンルに比べ近似曲線へのあてはまりがやや悪い。一方、雅楽、ロック、ジャズはそれぞれ非常に異なるジャンルに感じられるにも関わらず、その構成音の出現頻度は近似曲線でよく表現できていることがわかる。

図 2 は、各曲毎に決定した式 (1) の定数 a, b を 2 次元空間上に配置したものを示す。ジャズは b の値は比較的小さい範囲に、また雅楽は b が約 0.7 から 0.8 の範囲に分布しているのに対し、クラシックやロックでは広範囲に分布して

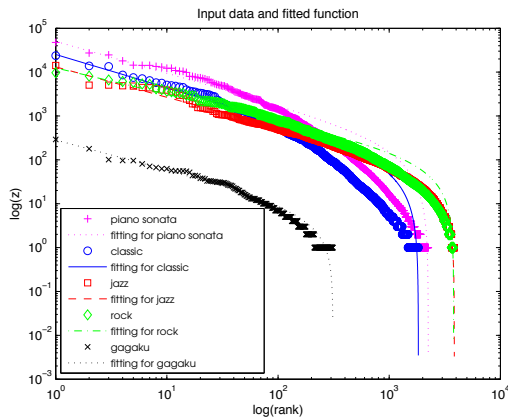


図 1 各ジャンルの構成音の出現頻度. 横軸は構成音の頻度順位であり, 縦軸は各構成音の出現頻度.

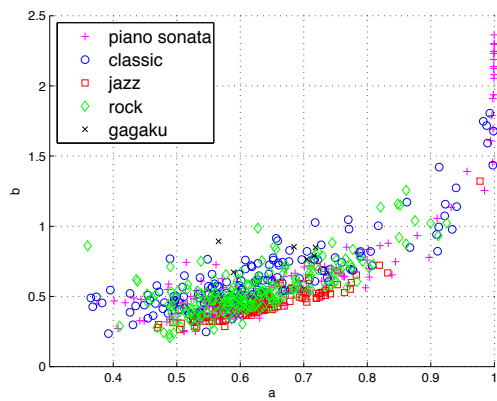


図 2 式 (1) のパラメータ a, b . 横軸が a を, 縦軸が b の各値を表す.

いることがわかる. このことから, ジャズや雅楽には構成音の使用頻度に暗黙のルールがあると考えられる. ピアノソナタも広い範囲に分布するが, b がやや低めの値であること, また, 近似曲線でのあてはまりが悪かったため, a, b を求めることができない曲 (図中で $a = 1$ となっているもの) が多かった点で特徴的である. 近似がうまくいかなかった曲は, 出現頻度第 1 位の構成音の頻度が他の構成音に比べ極端に高い曲, つまり同じ音を非常に多く使っている曲, 例えばショパンの「雨だれの前奏曲」などの曲であった. 以上のように構成音の頻度分布にはある程度ジャンルの特徴が現れる. また, 全ての曲の分布の様子から, a, b の値は全く独立ではなく, ジャンルにもあまりよらない拘束があることもわかる.

4. 構成音の進行頻度

4.1 解析手法

前節では個々の構成音の使用頻度に着目したが, 本節では各構成音の進行, すなわちどの構成音からどの構成音へ遷移しやすいかについて分析した. 具体的には, 構成音の遷移は単純マルコフ過程としてその遷移確率に注目し, ジャンル

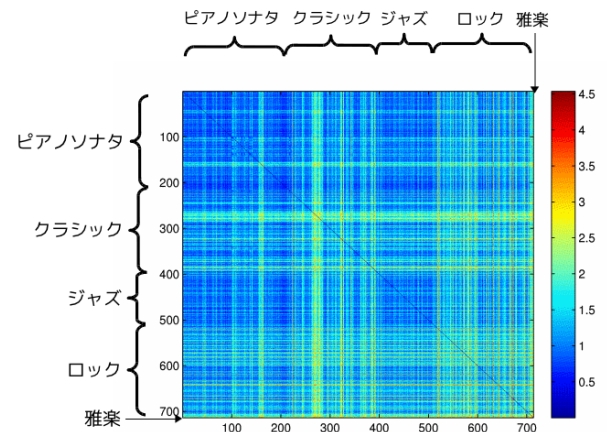


図 3 曲間類似度. 色が青いほど, 曲中の構成音の遷移確率が似ていることを示す.

ル毎の特徴を検討した. 各曲で登場する構成音の種類が T 個であるとき, t 番目に登場する構成音を $c_t (t = 1, \dots, T)$ とし, その曲が属するジャンル g に使われている構成音の使用頻度が高いものから順に, $S_g = \{s_1, s_2, \dots, s_N\}$ (N はジャンル g に存在する構成音の総数) とすると, 次式のような $N \times N$ の状態遷移行列が書ける.

$$P_{i,j} = P(c_{t+1} = s_j | c_t = s_i) (i, j = 1, 2, \dots, N) \quad (3)$$

本研究では, 状態空間 S_g を頻度が上位の 29 個とそれに一致しないその他の音 1 個の 30 個に絞って, 各曲の遷移確率 $P_{i,j}$ を計算し, この遷移確率から, 曲同士の類似度を計算した. 類似度には次式で定義される Bhattacharyya 距離 $BD(p, q)$ を用いた.

$$BD(p, q) = -\log \left(\sum_{x \in X} \sqrt{p(x)q(x)} \right) \quad (4)$$

視覚的に曲間の類似度を把握するため, 本研究では多次元尺度構成法 (MDS: Multidimensional Scaling) によって, 楽曲を類似度に基づき 2 次元空間上に配置した.

4.2 解析結果と考察

各曲間の距離行列を図 3 に示す. この図から, ピアノソナタとジャズは同ジャンル内の曲同士で類似度が高く, またピアノソナタ, ジャズ同士でも類似度が高い傾向にあることがわかる. 一方, クラシック, ロック, 雅楽は他ジャンル, 同ジャンルともに類似度が低いことがわかる.

距離行列に MDS を適用後, 2 次元空間上に配置したものを図 4 に示す. ピアノソナタ, クラシックは原点より右側に, ジャズ, ロックは左側に多く分布しており, ある程度ジャンル毎の特徴が表現されているように見える. ただし, この MDA による 2 次元空間への縮約は, mardia の第一基準, 第二基準ともに満足しておらず, 2 次元空間上に曲を配置することは困難と判断された. すなわち, より高次元のマッピングを行うことでジャンル毎の特徴がより明確にな

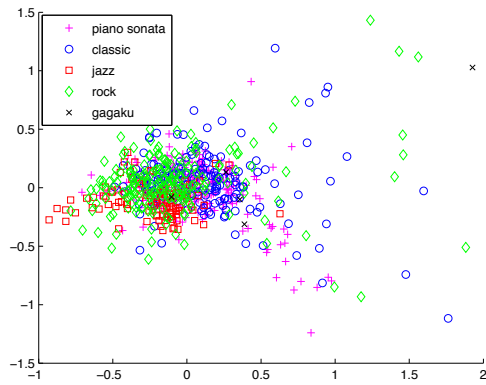


図 4 MDS 適用後の 2 次元空間上の曲の配置. 各曲の相対的な位置関係を表す.

る可能性がある.

5. 楽曲におけるカオス時系列解析

5.1 楽曲の時間発展パターン

前節までの結果は、構成音の出現頻度やその進行頻度が曲やジャンルの特徴をある程度表現していることを示唆する. そこで本研究ではさらに、低出現構成音と高出現構成音がどのような時間発展パターンによって各曲内で出現するかを検討した. ここで、低出現構成音とは、楽曲中で使われている構成音の平均出現頻度未満の構成音のことである.

半音階ベクトル図と構成音の出現頻度に着目した時間発展パターンの例を図 5 に示す. 半音階ベクトル図は横軸が時間、縦軸が半音階ベクトルを表し、赤色は高頻出構成音、黄色は低頻出構成音、青色は音が鳴っていない音階を示す. 時間発展パターンは横軸が時刻 t (厳密には t 個目に登場する構成音)、縦軸は時刻 t における構成音が何回目の登場であるかを示す. すなわち、落差が大きいほど何度も出現した構成音の後に稀な構成音が登場した、もしくはその逆ということになる. 図 5 で示されるように、ジャズやロックでは高出現構成音と低出現構成音が規則的に観察されることが多い. 一方で、ピアノソナタやクラシックではこのような周期性はあまり見られない.

5.1.1 楽曲におけるカオス性

リターンマップは、ある一定の遅れ時間 τ と、埋め込み次元 d を用い、観測時系列データ $x(t)$ を次式のような空間ベクトル v_t に再構成することを考える.

$$v_t = (x(t), x(t+\tau), \dots, x(t+(d-1)\tau))^T \in \mathbb{R}^d \quad (5)$$

図 6, 7 は $\tau = 1, d = 2, x(t)$ を時刻 t における構成音の時刻 t までの出現頻度として構成したベートーヴェンの「ピアノソナタ第 29 番第 4 楽章」と、ジャズの「Take the A Train」のリターンマップをそれぞれ示す. これらの図から、全体が部分で部分であるというフラクタル構造が見て取れる. 図 6 の各リターンマップをみると、序盤

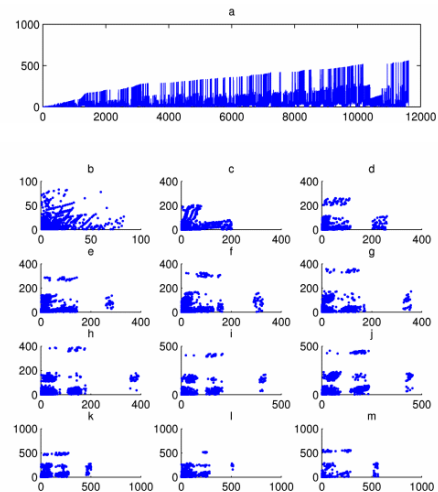


図 6 ベートーヴェンのピアノソナタ第 29 番第 4 楽章. (a) は構成音の出現頻度に着目した時間発展パターン. (b) から (m) はリターンマップ. 横軸は時間 t における構成音の頻度, 縦軸は時間 $t+1$ における構成音の頻度を表す. ここで、時間 t は曲中で構成音が切り替わったとき $t \rightarrow t+1$ となり、 $t = \{1, 2, \dots, T\}$ で最大値 $T = 11618$ である. (b)~(m) はそれぞれ、時空間パターンの $t = 1$ から 1000, $t = 1001$ から 2000, $t = 2001$ から 3000, $t = 3001$ から 4000, $t = 4001$ から 5000, $t = 5001$ から 6000, $t = 6001$ から 7000, $t = 7001$ から 8000, $t = 8001$ から 9000, $t = 9001$ から 10000, $t = 10001$ から 11000, $t = 11001$ から 11618 のリターンマップを示す.

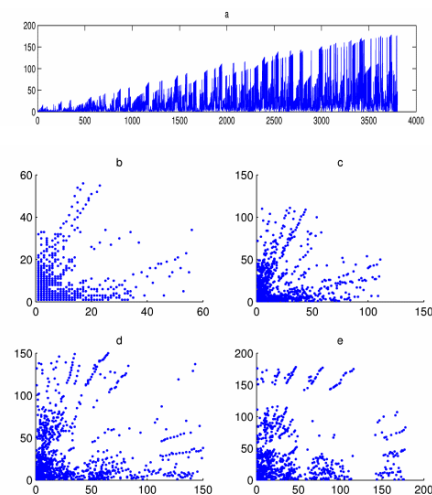


図 7 Take the A Train. (a) は構成音の出現頻度に着目した時間発展パターン. (b) から (e) はリターンマップ. 横軸は時間 t における構成音の頻度, 縦軸は時間 $t+1$ における構成音の頻度を表す. ここで、時間 t は曲中で構成音が切り替わったとき $t \rightarrow t+1$ となり、 $t = \{1, 2, \dots, T\}$ で最大値 $T = 3803$ である. (b)~(e) はそれぞれ、時空間パターンの $t = 1$ から 1000, $t = 1001$ から 2000, $t = 2001$ から 3000, $t = 3001$ から 3803 のリターンマップを示す.

((b) から (c)) では原点付近に固まっていたものが、それ以降になると、 $(x_t, x_{t+1}) = (\text{低}, \text{低}), (\text{高}, \text{低}), (\text{低}, \text{高})$ といった場所にいくつかのクラスターを形成している. そして、低

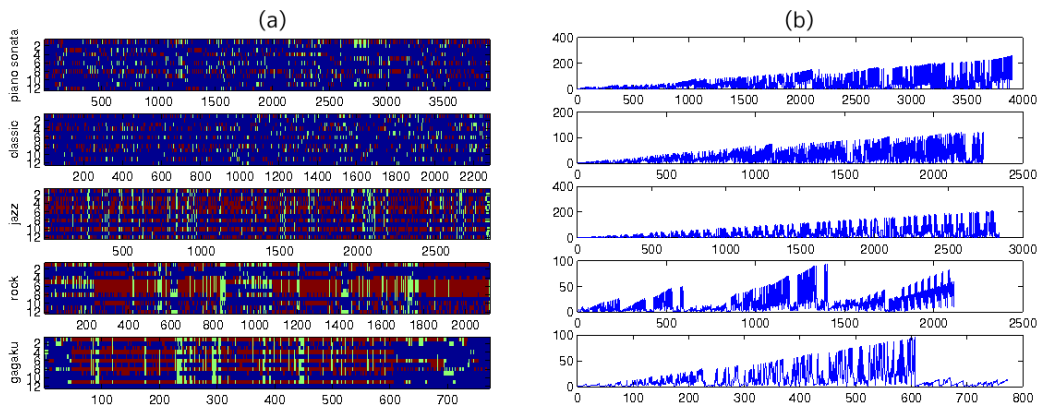


図 5 典型的な半音階ベクトル図 (a) と時間発展パターン (b). 楽曲は共に上からピアノソナタ (ベートーヴェン, 第 1 番 第 1 楽章), クラシック (アルビノーニ, アダージョ), ジャズ (サミー・フェイン, Alice in wonderland), ロック (Garbage, 13X Forever), 雅楽 (越天楽).

出現構成音が出現した後は低出現構成音もしくは高出現構成音が現れるが, 高出現構成音の後は必ず低出現構成音が現れて高出現構成音へ続けて遷移することはない, という事を意味する. すなわち, この曲ではよく使う音とあまり使われない音が明確に分離され, 計画的に楽曲中で利用されていることがわかる.

「Take tha A Train」の各リターンマップ (図 7) でも終盤 (e) にクラスタが確認できるが, 図 6 に比べ小さなクラスタが分散して存在し, 特有の遷移パターンがいくつも形成されていることがわかる. このクラスタはジャズにおける即興性を実現するための即興パターンを表現している可能性があると考えられる.

5.2 カオスの定量的評価

前節では, 楽曲の構成音の使用頻度の時間発展パターンに注目することにより, 例えばベートーヴェンのピアノソナタ第 29 番第 4 楽章ではフラクタル的特性が存在することを示唆する結果を得た. そこで, 様々な楽曲についてフラクタル性およびカオス性の存在の解析を試みた. カオスの特徴として, アトラクタの幾何学的な特徴としての自己相似性 (フラクタル構造) と, 力学的な特徴としての軌道不安定性とそれに起因して生じる長期予測不能性がある. これらの特徴は, それぞれフラクタル次元解析 [8] と, リャプノフスペクトラム解析 [9] により定量的に評価することができる. これら 2 つの特徴は必ずしも 1 対 1 に対応するとは限らず, 互いに独立している. すなわち, フラクタル構造を持ちながらも軌道不安定性を持たない系や, その逆も存在する.

ベートーヴェンのピアノソナタ第 29 番第 4 楽章の場合はリャプノフ指数が $\lambda = \{0.3080, -0.6386\}$ で, フラクタル次元が $D_L = 1.4822$ であった. 最大リャプノフ指数が正

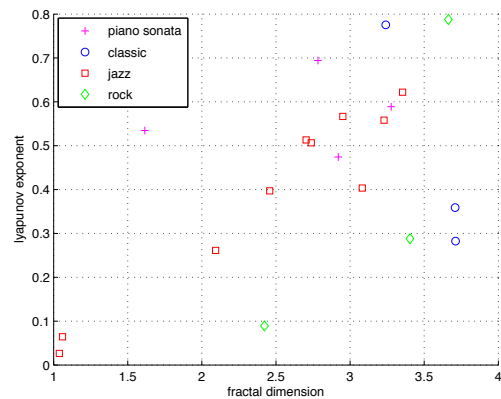


図 8 最大リャプノフ指数とフラクタル次元の関係. 横軸がフラクタル次元, 縦軸が最大リャプノフ指数を表す.

で, かつフラクタル次元も非整数値をとる結果となり, この曲がカオス性およびフラクタル性を有することがわかる.

更に, 各ジャンル 50 曲 (雅楽は 6 曲のみ) に絞り, リャプノフスペクトラムとフラクタル次元を計算した結果を図 8 に示す. この図は, 最大リャプノフ指数が正で, かつフラクタル次元が非整数値をとる曲のみを示している. 定量的にカオス性およびフラクタル性を持たないと評価された曲は除外した. この図から, カオス性を持つ曲数が多いジャンルはジャズだが (50 曲中 10 曲), フラクタル次元や最大リャプノフ指数の数値は小さく, 比較的その程度は弱い. 一方で, クラシック, ピアノソナタ, ロックはカオス性をもつ曲数こそジャズに比べ少ないが (クラシックでは 50 曲中 3 曲, ピアノソナタでは 50 曲中 4 曲, ロックでは 50 曲中 3 曲), 比較的強いカオス性を持つことがわかった. 具体的には, ベートーヴェンは「エリーゼのために」など 15 曲中 3 曲がカオスの特徴を持っていた. 一方で, ブラグミュラーの「25 の練習曲」は 25 曲中カオスの特徴を持つ曲は 1 曲のみであった. 雅楽はいずれの曲についてもカオス性

は認められなかった。

6. まとめ

本研究では、構成音の出現頻度と進行、時空間パターンに着目することで楽曲の属性別の特徴を探った。全く異なる場所で生まれ、多様に発展してきたジャンルに属する楽曲においても、構成音の出現頻度の分布は(1)式に近い形をとることがわかった。ただし、クラシックとピアノソナタにおいてはこの近似曲線からややはずれる傾向がある等ジャンルによる特徴も観察された。またジャズで使われる構成音やその進行の頻度については他ジャンルに比べ狭い範囲に分布しており、即興性を実現するために暗黙の拘束が存在することを伺える。さらに、低出現構成音と高出現構成音の出現パターンを調べた結果、楽曲においていくつかの時間発展パターンが存在すること、そしてカオス性およびフラクタル性を持つ楽曲が存在することがわかった。音楽においては何度も繰り返されるテーマとそれ以外の旋律、協和音と不協和音の組み合わせによって緊張感と安心感を生み出し、名曲では、その時間発展を巧みに設計することによって感動を生み出していると考えられる。本研究の結果は、構成音の使用頻度の時間発展パターンに着目した力学系的視点により、楽曲の構成方法を分析・分類できることを示唆しており、その結果は、感動を生む名曲の時間構造の理解の一助になると期待している。

参考文献

- [1] M. Beltran del Rio, G. Cocho, G. G. Naumis: *Universality in the tail of musical note rank distribution*, Physica A, 387, 5552-5560, (2008)
- [2] 鈴木伸子: 古大津絵の色彩に関する研究: 第1報 使用色の頻度特性について, 日本色彩学会誌 Vol.17 No.4, 242-248, (1994)
- [3] G. artinez-Mekler, R. A. Martinez, M. Beltran del Rio, R. Mansilla, P. Miramontes, G. Cocho: *Universality of Rank-Ordering Distributions in the Arts and Sciences*, PLoS ONE Vol.4, 1-7, (2009)
- [4] R. N. Shepard: *Circularity in judgments of relative pitch*, The Journal of The Acoustical Society Of America Vol.36, No.12, 2346-2353, (1964)
- [5] 吉野巖, 阿部純一: メロディの調を認定する過程の計算モデル, 情報処理学会研究報告, 23-30, (1993)
- [6] 宮下芳明, 西本一志: MIDI データからのクロマプロファイルの抽出と分析, 情報処理学会研究報告, 97-101, (2003)
- [7] <http://sysplan.nams.kyushu-u.ac.jp/gen/edu/Algorithms/DownHillSimplexAlgorithm/DownHillSimplexMethod.html>.
- [8] 合原一幸: カオス時系列解析の基礎と応用, 産業図書, (2000)
- [9] M.Sano and Y.Sawada: *Measurement of the Lyapunov Spectrum from a Chaotic Time Series*, Physical review letters, 1082-1085, (1985)