

LLMによる任意の周波数列からの音楽理論の生成

平井 辰典^{1,a)}

概要：音楽理論を生成するタスク「音楽理論生成」を提案し、LLMによって任意の周波数列から音楽理論を生成する手法を示す。現在広く用いられている音楽理論の多くは12平均律を前提としている一方、任意の周波数列を対象として音楽理論を生成する枠組みは提案されていない。一方で、近年のポピュラー音楽ではラップなどの微分音を含む表現が多く使われている。そこで本研究では、ラップのフロウから抽出した周波数をはじめとしたさまざまな周波数列を対象とし、周波数列と音響特徴量をLLMへ入力することで12平均律を前提としない音楽理論を生成する。さらに、生成された音楽理論をJSON形式で構造化し、演奏インタフェースへの応用例を示すことでAIによる音楽理論生成の可能性について検討する。

1. 背景

1.1 西洋音楽由来の音楽理論と12平均律の限界

現代の音楽制作において広く用いられている音名、音階、和音、調性などの概念は主に西洋音楽に由来する枠組みに基づいている。特に12平均律は楽器設計、記譜法、MIDIなどに深く組み込まれており、現代の音楽制作における暗黙の前提となっている。その結果、本来は異なる音高体系^{*1}や音響の特徴を持つ民族楽器や声などの表現であっても、制作環境上では12平均律に基づく音高で扱われやすい。本研究では、このような制作環境上の標準化が音楽的多様性を記述・創作する上での制約になり得る点に着目し、従来の音律とその上に構築された音楽理論の枠組みとは異なるメタな文脈からの音楽理論体系について検討する。

本稿では、12平均律による離散的な音高配置と、音階、和音、調性、機能と声などの西洋調性音楽に由来する概念を中心とした理論的枠組みを「(12平均律に基づく)従来の音楽理論」と呼ぶ。既存の音楽理論の中にはこの定義に当てはまらないものも多数存在するが、本研究では理論の基盤となる音高体系のみに焦点を置き、12平均律に基づく従来の音楽理論とは異なる音高体系を基盤とする新たな理論の構築について検討する。なお、本研究は、従来の音楽理論を代替する理論の実現という位置付けではない。これまであまり主流ではなかったゼンハーモニック的な微分音を対象とし、12平均律における従来の音楽理論に相当する概念を微分音についても検討することを目的とする^{*2}。

¹ 駒澤大学

Komazawa University, Setagaya, Tokyo 154-8525, Japan

^{a)} thirai@komazawa-u.ac.jp

^{*1} 音楽で使用する音高の集合と、音同士の関係を定めた体系

^{*2} 説明の都合上、「音階」や「和音」といった従来の音楽理論の上

1.2 従来の音楽理論では記述しにくい音楽実践

音楽理論は多くの音楽を記述し、分析し、創作するための有効な枠組みを提供してきた一方で、現代の音楽実践には、既存の枠組みだけでは記述が難しいものも少なくない。

たとえば、ラップにおけるフロウは、歌唱旋律のような明確な音高列としては捉えづらい表現により形成されることが多い。そのため、ラップのフロウを既存の旋律理論や和声理論の枠組みだけで記述しようとするとその音楽的特徴を十分に捉えられない可能性がある。特に声の基本周波数の連続的な変化や、音高と発話の中間的な表現は12平均律上の離散的な音名へと単純に変換することが難しい。

他にも、ノイズ音楽やアンビエント音楽では、明確な音高、和音、拍節、調性が中心的な役割を持たないことが多く、これらの音楽実践を従来の音階、和音、機能と声を中心とした理論によって説明することは困難である。

また、音楽の対象は従来の楽器音や歌声に限定されるものではない。Schaferは、自然音、生活音、機械音などを含む人間を取り巻く音響環境全体をサウンドスケープとして捉え、環境中の多様な音を聴取および記述の対象として位置付けている[1]。ミュージック・コンクレートでは、録音された環境音や物音を加工・編集し、それらを楽曲を構成する音響素材として用いる実践が行われてきた[2]。こうした背景のもと、鳥の鳴き声、環境音、機械音などの音響素材を音楽制作に取り込むサンプリングの手法は現代の音楽制作における重要な表現技法の一つとなっている。サンプリングに使用される音響素材には、人間の歌唱や楽器演奏とは異なる周波数構造や時間変化が含まれているが、そのような音に関して音階、和音、進行などの創作規則とし

に成り立つ用語を使用するが、本研究で生成するものは各概念に相当する新たな概念を指すものであることに注意されたい。

て一貫して体系化する理論は十分に確立されていない。

このように、現代の音楽実践においては、既存の音楽理論が有効に機能する領域がある一方で、その枠組みだけでは十分に捉えきれない音楽的対象も多く存在する。本研究では、これらの対象が持つ音響の特徴、とりわけ周波数構造から音高間の関係や創作上の規則を体系化することで、新たな音楽理論を構築する方法について検討する。特に本稿では、任意の周波数列を出発点として、音階、音程、和音、和音進行に相当する概念を生成することにより、既存の音楽理論では記述しにくい音響素材を、創作に利用可能な音楽理論として構造化することを目指す。

1.3 AIによる「音楽生成」から「音楽理論生成」へ

近年、AIによる音楽生成技術が急速に発展しており、Suno^{*3}やUdio^{*4}などのテキストによる指示から楽曲を生成するサービスの登場により、専門的な作曲技術や音楽制作環境を持たないユーザでも短時間で楽曲を生成できるようになっている。このような技術は作曲支援、楽曲の試作、BGM制作などさまざまな用途で有用であり、音楽制作の敷居を下げる可能性を持つ。

一方、現在の音楽生成AIの多くは、既存の楽曲データから学習した音楽的特徴に基づき、既存のジャンル、様式、音色、編成などに沿った楽曲を生成するものである。そのため、生成される楽曲は既存の音楽理論や制作環境の影響を強く受けていると考えられる。すなわち、現在の音楽生成AIは、既存の音楽的枠組みの内側における楽曲生成を高度化する一方、その枠組みの外側において発揮される創造性や、新たな音楽的表現を生み出す可能性については十分に検討されていない。

また、AIが楽曲制作の工程そのものを代行することは、人間の制作負担を軽減する一方、音を試行錯誤して新たな規則や表現方法を発見するような創作過程を縮小させる可能性もある。音楽創作の価値は完成した楽曲だけにあるのではなく、音の関係を探索し、演奏や作曲を通して新しい音楽的可能性を発見する試行錯誤の中にも存在するものである。AIが、人間の代わりに作曲をするためだけに使われるのではなく、人が新たな音楽を探索することを支援するために利用されることが望ましい。

そこで本研究では、AIによる生成の対象を楽曲そのものから音楽理論へと拡張する。具体的には、任意の周波数列を入力として与え、各音に名称や役割を付与するとともに、和音、和音進行、メロディパターンに相当する概念などをLLMによって生成する。これにより、入力された任意の周波数列を、演奏や作曲に利用可能な理論として体系化することを目指す。

2. 関連研究

2.1 音楽理論の成立と体系化

音楽理論は、各時代や地域における音楽実践を記述、分析、教育するための枠組みとして形成されてきた。西洋音楽においても、音律、旋法、対位法、和声、調性などの概念は、音楽様式や楽器、演奏環境の変化に伴って段階的に体系化されてきた。したがって、音楽理論は音楽を制作・生成するための規則であるだけでなく、既に存在する音楽実践から特徴や規則性を抽出し、言語化したものとして捉えることができる。

現代の音楽教育や音楽制作では、音階、和音、調性、機能と和声など、西洋調性音楽に由来する概念が広く用いられている。鍵盤楽器、五線譜、MIDIなどのツールは、12平均律による離散的な音高配置が標準的な枠組みとなっている。そのため、現代において一般に音楽理論として学習・利用される内容は、音楽理論の歴史的・文化的な多様性のすべてを包含するものではなく、12平均律に基づく音高体系と音楽実践を中心として体系化されたものといえる。

2.2 音楽理論を再体系化する試み

既存の音楽理論は、対象とする時代、地域、音楽ジャンルなどの違いにより、異なる用語や規則によって記述されてきた。特に、クラシック、ジャズ、ポップスでは、類似した音楽現象に対して異なる説明が与えられる場合がある。このような複数の理論体系を比較し、現代の音楽実践に適した形で再整理する試みの一つとしてSoundQuestによる「自由派音楽理論」が挙げられる[3]。

SoundQuestは、既存の各種音楽理論を対比しながら、メロディ、コード、リズムなどに関する概念を体系的に整理した音楽理論の情報サイトである。「自由派音楽理論」では、特定のジャンルに由来する規則を絶対的な禁則として扱うのではなく、それぞれの理論が成立した背景や適用範囲を考慮しながら複数の理論を横断的に利用する立場が採られている。これは、音楽理論を唯一の固定的な体系としてではなく、音楽実践に応じて選択・再構成可能な知識体系として捉える試みといえる。

他にも、既存の音楽理論を現代の教育環境に合わせて再編成する事例として、Open Music Theoryが挙げられる[4]。これらの試みは、既存の複数の音楽理論を整理し、それらの関係や適用範囲を明確化するものである。これらが既存理論の再整理を目的とするのに対し、本研究は既存理論が前提としない音高体系の理論を生成を目的とする。

2.3 機能と和声によらない理論

調性や機能と和声を持たない音楽に対しても、音高構造を体系的に記述・分析するための理論が提案されている。そ

^{*3} <https://suno.com/>

^{*4} <https://www.udio.com/>

の代表例として、ピッチクラス集合論が挙げられる [5], [6].

ピッチクラス集合論では、オクターブの違いを同一視し、12 平均律上の各音を 0 から 11 までの整数によって表現する。楽曲中に現れる複数の音は、順序を持たないピッチクラスの集合として扱われる。各集合について、含まれる音の数、音程の構成、移高や反転による関係、共通音、部分集合と包含関係などを調べることで、機能と和声や調性を前提とせずに音高構造を分析できる [5].

また、ピッチクラス集合に含まれる音程は、Interval-Class Vector (ICV) により表現できる。ICV は、音程類 1 から 6 までの各音程類が、集合内の音高対にそれぞれ何回含まれるかを示す 6 次元のベクトルであり、異なるピッチクラス集合の音程構成を比較する際に用いられる [6]. ピッチクラス集合論に基づく分析は、伝統的な和音名称や和声機能を適用しにくい音楽においても、反復する音高集合や集合間の関係を記述することを可能にする。

ピッチクラス集合論は、調性や機能と和声を用いない音楽についても、音高間の関係に基づく構造を体系的に記述・分析するための枠組みを提供しているが、基本的に 12 平均律における 12 のピッチクラスを前提としている。これに対し、本研究では、12 平均律に限定されない任意の数の音からなる周波数列を対象とし、その周波数および音高間の関係から音楽理論を生成することを目指す。

2.4 微分音音楽と非 12 平均律の音楽理論

12 平均律とは異なる音高体系を用いる音楽として微分音を使った音楽がある。微分音は、一般には 12 平均律の音高に一致しない音高または音程の音を指す。微分音音楽では 24 平均律や 31 平均律などの N 平均律、純正律、不均等な音程構造を持つ音階など、多様な音高体系が用いられる。

非 12 平均律の音高体系に基づく代表的な音楽理論・音楽体系として、Partch による純正律の音楽体系が挙げられる。Partch は、整数比に基づく 1 オクターブ 43 音の音高体系を構築し、その体系に対応する記譜法、作曲法、専用楽器を開発した [7]。この事例は、12 平均律とは異なる音高体系を、作曲、演奏、記譜まで含めて実践可能な音楽理論・音楽体系へと発展させたものである。

N 平均律に関しては、Wyschnegradsky が 24 平均律に基づく四分音和声について、音程、和音、声部進行などを体系的に論じている [8]。Blackwood は、1 オクターブを 13 から 24 までの異なる数に等分した音高体系を比較し、各体系における音階、和音、調性的な構造、記譜法などを検討している [9]。また、Fokker は 31 平均律に基づく音高体系と、それに対応する鍵盤楽器を提案し、31 平均律による演奏環境を構築した [10]。他にも微分音に関するさまざまな理論は専門家からアマチュアまで幅広い音楽家によって探求されている。これらの事例は、多様な音高体系に基づく新たな音楽的構造を構築し得ることを示している。

Sethares は、協和・不協和の知覚が音程だけでなく音色のスペクトルにも依存することを指摘し、音色に適した音階を設計する方法を提案している [11]。この研究は、音階をはじめとする音高体系を設計する際には、音程関係だけでなく音色のスペクトル構造も考慮すべきであることを示唆している。

既存の非 12 平均律理論が整数比、オクターブの等分数、音色と協和性などの設計原理に基づくのに対し、本研究は、ラップや鳥の鳴き声などから抽出した任意の周波数列とその音響特徴量を入力し、LLM によって音名、和音、進行などを生成する点に特徴がある。

2.5 AI による音楽生成

深層学習を用いた音楽生成研究では、旋律、和音、演奏、音響信号など、さまざまな音楽表現を対象とした生成手法が提案されてきた [12]。近年では、テキストによる説明を条件として音響信号を直接生成する手法も発展している。たとえば、MusicLM は、テキストによる記述から音楽を生成する階層的な生成モデルであり、入力された旋律を条件として、指定された音楽様式の音楽を生成することにも対応している [13]。MusicGen は、音楽信号を圧縮された離散トークン列として扱い、テキストによる説明や入力旋律を条件として音楽を生成する言語モデルである [14]。

これらの研究では、ジャンル、楽器編成、雰囲気、テンポ、旋律などを条件として、それらに対応する楽曲や音響信号を生成することが主な目的とされている。すなわち、AI が生成する対象は、主として楽曲や音響信号であり、音高体系や音楽理論そのものを生成対象とするものではない。

一方、音楽知識を取り入れた生成手法では、和音、リズム、カデンツ、楽曲構造などを制約や条件として利用し、生成結果の制御性や一貫性を高める試みが行われている。たとえば、DeepBach は、バッハ風の四声体コラールを対象とし、旋律、リズム、カデンツ、声部の一部などを制約として与えることで、その条件を満たす四声体コラールを生成できるモデルである [15]。これは、音楽理論や音楽構造に関する知識を、生成モデルを制御するための条件として利用する研究と位置付けられる。

これらの研究では、音楽理論や様式に関する知識は、主に楽曲生成を制御するための条件として用いられている。これに対し、本研究では、楽曲ではなく音楽理論そのものを LLM による生成対象とする点に特徴がある。

3. LLM による音楽理論生成

本章では、LLM を用いて任意の周波数列から音楽理論を生成する手法について述べる。音楽理論生成という新たなタスクを定義するため、本研究で対象とする音楽理論の範囲を明確化し、音楽理論をデータとして構造化する方法、入力データ、および音楽理論生成手順について説明する。

3.1 前提条件

本研究では、任意の周波数列を出発点として、その周波数列を用いた音楽の制作・演奏を支援する新たな音楽理論を生成することを目的とする。音楽理論には、音階、音程、和音、調性、旋律、リズム、形式、記譜法、楽器法など、多様な要素が含まれる。そこで、本研究では生成対象をあらかじめ限定する。

本研究では、音高体系に関する理論を主な生成対象とする。具体的には、各音に名称や役割を与えると同時に、音同士の関係、和音、和音進行、旋律パターンなどを生成対象とする。一方、リズム理論、楽式論、楽器法、音色設計、歌詞などは対象としない。生成された音楽理論を活用する際には、既存のリズム、拍節、形式、音色設計などの音高以外の音楽的要素と組み合わせて利用することを想定する。

3.1.1 本論文における「音楽理論」の定義

音楽理論という語は、歴史的、文化的、教育的な文脈に応じた多様な意味を持ち得る。特に、伝統的な西洋音楽理論では、音階、音程、和声、対位法、調性、形式、記譜法などが長い歴史の中で体系化されてきた。したがって、任意の周波数列から生成された規則群を、それらと同じ意味で直ちに音楽理論と位置付けることは適切ではない。

本論文では、この点を踏まえ、「音楽理論」を伝統的な西洋音楽理論と同等の歴史的・規範的体系としてではなく、任意の音高体系を作曲や演奏に利用するための最小限の構造化された規則群として操作的に定義する。すなわち、本研究で生成する「音楽理論」は、既存の音楽理論を置き換えるものではなく、未知または非標準的な音高体系を創作に利用するための便宜的な理論的枠組みである。

本研究で用いる「音名」、「和音」、「和音進行」、「カデンツ」といった語は、既存の音楽理論における概念をそのまま適用するものではない。本研究では生成された音楽理論を操作的に定義するため、既存の音楽理論と類似した役割や機能を持つ要素に対しては便宜的に既存の用語を用いる。

たとえば、本論文で「和音」と呼ぶものは、必ずしも従来の三和音や機能と声における和音と同一の概念ではない。本研究では、入力された周波数列に含まれる複数の音の同時発音によって形成される音のまとまりを「和音」に相当するものとして位置付ける。同様に、「カデンツ」に相当する概念も既存の西洋調性音楽における終止形や調性を直接再現するものではなく、生成された音高体系の内部で、中心性の形成、緊張形成、解決、終止感などに相当する機能を記述するための概念として便宜的に用いる。

将来的に、本研究で生成された音楽理論が発展した場合には、これらの概念を表す新たな用語についても検討されることを想定している。

このように本論文では、既存の音楽理論の用語を説明上の足場として用いながら、任意の周波数列に基づく新たな音楽理論を構成する。生成された理論は初期的な枠組みで

あり、演奏、作曲、聴取、および評価を通じて継続的に検証・修正されることを想定する。

3.1.2 生成対象

本研究では、任意の入力周波数列を作曲や演奏に利用するための音楽理論を生成対象とする。入力となる周波数列は、既存の音律に基づいて設計されたものである必要はない。具体的には、人が任意に設計した周波数列、ラップ音声の基本周波数列、鳥の鳴き声や環境音から抽出した周波数列に加え、純正律やN平均律など既存音律に基づく音高列を入力として利用することもできる。生成対象には、各音の名称、役割、音同士の関係、和音とその役割、和音進行、旋律パターンなどを含める。

LLM への入力には、周波数列のほか、音同士の不協和度に基づく音響特徴量を含める。これにより、言語モデルである LLM が音同士の組み合わせから生じる協和・不協和に関する情報を利用して音楽理論を生成できるようにする。これらの音響特徴量は LLM への入力前に算出し、生成された音楽理論を解釈・評価するための補助的な指標としても用いる。具体的な算出方法については後述する。

前述の通り本研究ではリズム理論、楽式論、楽器法、音色設計、歌詞などは生成対象とはせず、生成された音楽理論を実際の楽曲制作へ応用する際には、既存のリズムなどと組み合わせて利用することを想定する。

3.2 音楽理論の構造化

LLM は自然言語による説明生成に適している一方で、自然言語のみで記述された音楽理論は後続の処理やシステム実装には利用しにくい。本研究で生成する音楽理論は、人による作曲や演奏だけでなく、新たな演奏インタフェースの開発、自動作曲アルゴリズム、既存 12 平均律メロディのアレンジなどへの応用を想定している。そのため、生成された音楽理論を自然言語によって説明するだけでなく、機械可読な形式で構造化して表現する。

既存の音楽データ記述形式としては、MusicXML が広く用いられている [16]。MusicXML は、五線譜、音符、休符、音価、拍子、調号、パートなどを記述することに特化した形式であり、楽曲や楽譜の表現に適している。一方、本研究で必要となるのは楽曲そのものの記述ではなく、楽曲を作る基盤となる理論体系の記述である。特に、新しい音高体系における各音の名称や役割、和音の機能や進行などの情報は MusicXML の主たる記述対象ではない。

そこで本研究では音楽理論を JSON 形式で構造化する。JSON 形式では、各音や和音の情報などを後続の処理で参照しやすい形で整理して記述できる。また、各音や和音に対して、周波数、基準音からの cent 値、安定度、緊張度などの数値情報を対応付けて記録できる。安定度、緊張度については 3.4.2 で述べる方法により算出した初期値を記録する。

表 1 JSON による音楽理論表現の主要 4 構成要素

項目	内容
pitches	音の定義
chords	和音の定義
melodic_patterns	旋律パターン
chord_patterns	和音進行

LLM により生成される音楽理論は、表 1 に示す 4 種類の主要要素からなる JSON 形式で構造化する。また、各要素で用いる詳細項目を表 2 に示す。

pitches は、入力周波数列に含まれる各音を定義する要素である。各音には、一意に識別するための id、音名を表す name、周波数を表す frequency_hz、基準音からの音程を表す cents_from_reference を付与する。さらに、安定度、緊張度、役割、説明を記述することで、各音を単なる周波数値ではなく、音楽理論内で特定の機能を持つ音として扱う。周波数、安定度、緊張度は LLM への入力情報を保持し、その他の項目のデータは LLM によって生成される。

chords は、複数音の組み合わせとしての和音を定義する要素である。ここでの和音は、既存の三和音や機能と和声における和音をそのまま意味するものではなく、入力周波数列に含まれる複数の音を同時に用いる際のまとまりを指す。各和音には ID と名称の他に、構成音を示す pitch_ids、安定度、緊張度、役割、説明を記述する。これらの情報により、どの和音が中心的に用いられるか、緊張感を形成するか、終止感を持つかといった情報を扱うことができる。

melodic_patterns は、生成された音高体系に基づく旋律パターンを記述する要素である。旋律パターンは音 ID の系列として表現し、その種別や説明を付与する。これにより、接近、解決、緊張形成など、旋律的な用法に関する情報を記述できる。同様に、chord_patterns は和音 ID の系列として和音進行を記述する要素であり、中心性、緊張形成、解決、終止など、カデンツに相当する進行の機能を表すために用いる。

このように、音楽理論を JSON 形式で構造化することで、生成された理論を自然言語による説明にとどめず、システム実装や応用研究に利用可能な機械可読データとして扱うことができる。また、音名や説明のような言語的情報と、周波数、cent 値、安定度、緊張度のような数値的情報を同一の構造内で管理できるため、LLM による概念生成と外部プログラムによる計算処理を接続しやすい。さらに、JSON は項目の追加や拡張が容易であるため、将来的には、各音や和音の使用しやすさを表す preference、聴取実験による評価値を表す rating、特定のジャンルとの関係を表す genre といった要素を追加することも可能である。このように、本研究で提案する JSON 形式は、音高体系を中心とした最小限の理論記述から、人間による評価やその他のアノテーション情報を含む音楽理論へと発展させられる

拡張性を備えている。その一方で、音楽理論にはオクターブ等価性、転調、旋法、調号、リズムなど、本研究で扱わないこととしている多様な要素も含まれる。本論文では、対象を 3.1 で述べた音高体系に関わる基本的な項目に限定しており、これらの要素については、将来的に JSON 構造へ追加・拡張する可能性がある対象として位置付ける。

本研究で提案する構造化は既存の音楽理論にも適用できる。その例として、表 3 に 12 平均律に基づく従来の音楽理論を本研究の形式で構造化した例を示す。ここでは C メジャースケールの例のみを示しているが、理論全体を構造化する場合には膨大なデータ量となるため、複雑な理論を扱うためにはデータ構造のさらなる工夫が必要となる。

3.3 周波数列の決定

本研究では、任意の周波数列を音楽理論生成の出発点とする。入力周波数列は人間が設計した音律に限定されず、既存の音律や規則に基づいて設計されている必要はない。ラップのフロウ、鳥の鳴き声、環境音など、音楽理論には適用されていない音から抽出した周波数列も対象とする。

周波数列は、たとえば以下の方法によって決定できる。

- 作曲者や研究者が意図的に周波数列を設計する方法
- LLM により周波数列そのものを生成する方法
- ラップのフロウなど、歌声から周波数を抽出する方法
- 鳥の鳴き声や環境音などの音響素材から抽出する方法
- 不協和度や既存音律との差異などの評価指標を用いて機械的に探索する方法
- 12 平均律、純正律、N 平均律など、既存の音律を利用する方法

周波数列の決定方法は本研究の音楽理論生成手法とは独立であり、さまざまな方法を採用できる。例として LLM により周波数列を生成する場合、音数や基準となる周波数などをプロンプトで指定して入力周波数列を生成することになる（例：1 オクターブあたり 5 音で、440Hz を基準とした周波数列等）。ここで、入力する周波数列は、後に生成される音楽理論に大きな影響を与える。音同士の周波数間隔が狭い周波数列では、緊張感を伴う音程や音の組み合わせが構成されやすい。一方、純正音程に見られるような単純な周波数比を多く含む周波数列では、比較的安定した音のまとまりが構成されやすい。したがって、どのような周波数列を入力として採用するかは、生成される音楽理論の性質を左右する重要な要因である。ただし、本論文の主目的は、最適な周波数列を探索することではなく、任意に与えられた周波数列から音楽理論を生成する枠組みを提案することである。そのため、周波数列の評価や探索については本論文の対象外とし、音楽の創作上有用な周波数列をどのように選択するかについては今後の課題とする。

入力された周波数列は、基準音に対する相対的な音程として整理する。入力時にユーザが基準音が指定した場

表 2 JSON による音楽理論表現の詳細項目

JSON 項目	日本語名	対象	内容
theory_name	理論名	全体	生成された音楽理論の名称
id	ID	音・和音 旋律パターン・和音進行	各要素を一意に識別するための識別子
name	名称	音・和音 旋律パターン・和音進行	音名, 和音名, 旋律パターン名, 和音進行名
frequency_hz	周波数	音	各音の周波数 [Hz]
cents_from_reference	基準音からの cent 値	音	基準音からの音程 [cent]
pitch_ids	和音構成音 ID	和音	和音を構成する音の ID 集合
sequence	パターン・進行の系列	旋律パターン・和音進行	旋律パターンでは音 ID の列, 和音進行では和音 ID の列
type	種別	旋律パターン・和音進行	接近, 解決, 緊張形成, 終止など, 旋律パターンまたは和音進行の音楽理論上の機能
stability	安定度	音・和音	中心音, 終止音, 休止点として機能しやすい程度
tension	緊張度	音・和音	不安定感や解決への方向性を生み出す程度
role	役割	音・和音	各音または各和音が理論内で果たす機能
description	説明	音・和音 旋律パターン・和音進行	各要素の音楽理論上の意味・役割を自然言語で説明

表 3 12 平均律に基づく従来の音楽理論を JSON 形式で構造化した例 (C メジャースケール)

JSON 項目	従来の音楽理論における例	本 JSON 形式での表現例
pitches	C メジャースケールの構成音	C, D, E, F, G, A, B をそれぞれ name として定義し, 周波数, 基準音 (C) からの cent 値, 役割, 安定度, 緊張度, 説明を付与する. たとえば, C を中心音, G を安定した支配的な音, B を C へ向かう緊張を持つ音として記述する.
chords	C, Dm, Em, F, G, Am, Bdim などのダイアトニックコード	各和音を音 ID の集合として定義し, 役割, 安定度, 緊張度, 説明を付与する. たとえば, C を中心的な和音, G を緊張を形成して C へ解決する和音, F を展開を生む和音として記述する.
melodic_patterns	順次進行, 導音から主音への解決, 音階上行・下行など	音 ID の系列として旋律パターンを記述する. たとえば, B から C への進行を解決を表す旋律パターン, C-D-E-F-G を上行する旋律パターンとして定義する.
chord_patterns	I-IV-V-I, ii-V-I などの和音進行	和音 ID の系列として和音進行を記述する. たとえば, C-F-G-C を中心化, 展開, 緊張形成, 解決を含む進行として定義し, 終止感を持つ和音進行として説明を付与する.

合にはその音を基準音として扱い, 特に指定されていない場合には入力周波数列のうちの最低音が基準音として扱われる. その後, 各音について, 基準音からの cent 値が `cents_from_reference` として LLM により計算される. これにより, 周波数そのものだけでなく, 基準音に対する各音の相対的な位置が JSON 内に記述される. また, 基準音と 1 オクターブ以上離れている音がある場合には, 1 オクターブ以内に折り返して配置する. 本論文ではオクターブ等価性を音楽理論の要素として扱うのではなく, 入力周波数列を整理するための前提として位置付けている.

3.4 音響特徴量の算出

任意の周波数列から音楽理論を生成するためには各音や各和音の役割を検討するための補助的な指標が必要である. 本研究では, 音楽理論生成を LLM の推論だけに依存するのではなく, 計算可能な音響特徴量を外部プログラムによってあらかじめ算出し, LLM への入力情報として利用する. 具体的には, 入力周波数列に含まれる音同士の不協和度を算出し, その値から各音および各和音の安定度と緊

張度を暫定的に算出する. これらの値は, LLM が音名の生成, 役割の付与, 和音の構成, 進行や旋律パターンの生成に利用可能な補助的な数値情報として LLM へ入力する.

3.4.1 不協和度

本研究では, Sethares のモデル [11] に基づく不協和度を算出し, これを各音および各和音の安定度・緊張度を算出する基礎となる値として用いる. Sethares のモデルでは, 音程だけでなく音色のスペクトル構造も不協和度に影響する. ただし本研究では, 入力周波数列の違いによる影響を評価することを目的とするため, すべての音に共通の倍音構造を仮定し, 音色を固定した条件で不協和度を算出する.

不協和度の算出にあたって, まず, 2 つの周波数成分の周波数を f_1, f_2 とする. 各音は基音とその倍音から構成される複合音として扱うため, 2 音間の不協和度は, 両音に含まれるすべての周波数成分の組み合わせについて $d(f_1, f_2)$ を計算し, その総和として求める. このとき, スケーリング係数 s を式 (1) で定義する.

$$s = \frac{0.24}{0.021f_{\min} + 19} \quad (1)$$

また、2つの周波数成分間の不協和度 $d(f_1, f_2)$ を式 (2) で求める。

$$d(f_1, f_2) = \exp(-3.5s\Delta f) - \exp(-5.75s\Delta f) \quad (2)$$

次に、各音を倍音を含む複合音として扱う。第 k 倍音の周波数を kf 、振幅を $1/k$ とし、本研究では標準設定として8倍音までを用いる。2音 f_i, f_j の不協和度 D_{ij} は、両音に含まれるすべての周波数成分の組み合わせについて d を計算し、振幅に基づく重み付き総和として定義する。

$$D_{ij} = \sum_{k_1=1}^n \sum_{k_2=1}^n \min(a_{k_1}, a_{k_2}) d(k_1 f_i, k_2 f_j) \quad (3)$$

ここで、 a_k は $1/k$ である。

この計算を周波数列内のすべての2音の組み合わせに対して行い、不協和度行列 $\mathbf{D}$ を構成する。不協和度行列は、同じ音同士の値が0となる対称行列とする。本研究では、不協和度を LLM への独立した入力特徴量とはせず、安定度および緊張度を算出するための値として利用する。次節で不協和度行列から各音および各和音の安定度と緊張度を算出する方法について述べる。

3.4.2 安定度 (stability) と緊張度 (tension) の導入

本研究では、各音および各和音に対して、安定度と緊張度を算出し、LLM への入力特徴量とする。ただしこれらの値は、不協和度に基づき機械的に算出した初期値であり、音楽的な安定感や緊張感を直接表すものではない。そのため、生成された理論を実際の演奏、作曲、聴取評価へ適用した結果に応じて更新する余地を残しており、必要に応じて音楽理論自体も更新していくことを前提とする。

安定度および緊張度は、主音や調性的中心を仮定せず、単純に周波数列内の2音間不協和度のみに基づいて算出する。まず、不協和度行列 $\mathbf{D}$ の対角成分を除外し、非対角成分の最小値 $D_{\min}$ と最大値 $D_{\max}$ を用いて、式 (4) により0から1の範囲に正規化する。

$$R_{ij} = \frac{D_{ij} - D_{\min}}{D_{\max} - D_{\min}} \quad (4)$$

この正規化後の値 R_{ij} を、音 i と音 j の間の音程緊張度 T_{ij}^I として定義する。また、音程安定度 S_{ij}^I は、音程緊張度 T_{ij}^I の補数、すなわち $S_{ij}^I = 1 - T_{ij}^I$ として定義する。

各音の緊張度は、その音が他のすべての音と形成する音程緊張度の平均値として算出する。周波数列に含まれる音の数を N とすると、音 i の緊張度 T_i^N は式 (5) で表される。

$$T_i^N = \frac{1}{N-1} \sum_{j \neq i} T_{ij}^I \quad (5)$$

また、音 i の安定度 S_i^N はその補数として定義する。

$$S_i^N = 1 - T_i^N \quad (6)$$

和音の緊張度は、和音に含まれるすべての2音の組み合

わせに対する音程緊張度の平均値として算出する。和音 C に含まれる音の組み合わせ集合を P_C とすると、和音の緊張度 $T^C(C)$ は式 (7) で表される。

$$T^C(C) = \frac{1}{|P_C|} \sum_{(i,j) \in P_C} T_{ij}^I \quad (7)$$

また、和音の安定度 $S^C(C)$ はその補数として定義する。

以上の定義により、他の音との不協和度が相対的に小さい音や和音ほど安定度が高く、不協和度が相対的に大きい音や和音ほど緊張度が高くなる。和音の安定度と緊張度は、入力周波数列のすべての3音の組み合わせについてあらかじめ計算しておき、すべての和音候補の中から、LLM が理論に採用する和音を選定する。

本研究では、すべての音に共通の倍音構造を仮定することで周波数列の違いのみが安定度および緊張度に反映されるようにしている。一方、Sethares は音色のスペクトル構造と音高体系を対応付けるスペクトルマッピングを提案しており、音色に応じて協和的な音高体系が変化することを示している [11]。本研究では、音色を固定することで周波数列の違いのみに着目して音楽理論生成を行うが、将来的にはスペクトルマッピングを導入することで、音色に適した音楽理論を生成する枠組みへと拡張できる可能性がある。

3.5 音楽理論生成

本節では、LLM による音楽理論の生成手順について述べる。音楽理論の生成にあたり、LLM には、入力周波数列、各音と和音候補の安定度、緊張度などの音響特徴量とともに、表1および表2に示したJSON構造を与える。LLM はこれらの情報に基づき、`pitches`, `chords`, `melodic_patterns`, `chord_patterns` の各要素を生成する。

音楽理論の生成に際し、LLM が学習済みの既存の音楽理論に関する知識を利用することを許容する。たとえば、中心音、安定音、緊張音、解決、終止といった既存の音楽理論の概念を参考として用いることを想定している。一方で、入力周波数列の音を既存の音楽理論へ無理に対応付けることはしない。本研究で生成される音楽理論は、既存理論の名称や機能をそのまま再現するものではなく、既存理論の概念を参考にしつつ、入力周波数列に固有の音の役割や音同士の関係を記述することを目的とする。

安定度および緊張度は、3.4節で述べた方法により事前に外部プログラムで算出し、LLM への入力情報として与える。LLM はそれらの数値を参照し、音名の付与、役割の解釈、和音の命名、進行や旋律パターンの構成、および自然言語による説明を生成する。このように、数値計算は外部プログラムが担い、LLM はそれらを踏まえた概念的・言語的な音楽理論の生成を担うことで、両者の特徴を組み合わせた音楽理論生成を実現する。

LLM による音楽理論生成の手順は次の通りとなる。

- (1) 入力となる周波数列の決定
- (2) 周波数列内の各音ペアについて不協和度を算出
- (3) 不協和度に基づいて、各音および各和音候補の安定度と緊張度を算出
- (4) 周波数列, 安定度, 緊張度および生成対象となる JSON 構造を LLM に入力
- (5) LLM が各音の名称と役割, 和音, 和音の役割, 和音進行, 旋律パターンを生成
- (6) JSON 形式の音楽理論データとそれを説明する自然言語による理論記述が出力される

本手法では、最終的に入力周波数列に対応する機械可読な JSON 形式の音楽理論データと、人が理解しやすい自然言語による理論記述が得られる。自然言語による理論記述は、人が生成された音楽理論を理解するための情報として使われ、理論を説明するドキュメントとして参照することを想定している。このドキュメントの整形についても LLM への追加のプロンプト入力によって行える。

また、生成された JSON 形式の音楽理論データを LLM へ再入力することで、理論を解説する Web ページ、理論に基づく演奏インタフェースなど、用途に応じたさまざまな表現へ展開することも可能である。たとえば、各音や和音の試聴機能や、鍵盤型インタフェースによる試奏機能を備えた Web ページを生成できる。

このように、本手法では JSON 形式の音楽理論データを用いることで、人による理解とシステムによる利用の双方に適した音楽理論の生成・活用を実現する。

4. 生成された音楽理論

本章では、生成された音楽理論の一部を紹介する。はじめに、4.1 節では本研究において最初に生成された音楽理論 Gravitonal Theory を紹介する。その後、4.2 節以降では、それぞれ異なる方法で作成した周波数列を入力として音楽理論を生成した事例を示す。これらの事例を通じて、さまざまな対象から音楽理論を生成できることを示し、本手法の応用可能性を示す。生成された各理論には多数の音、和音、和音進行および旋律パターンが含まれるため、本章ではそれぞれの理論の特徴を表す要素を抜粋して示す。

4.1 初めて生成された音楽理論

4.1.1 Gravitonal Theory (5 音版)

本研究において 2026 年 5 月に初めて生成された音楽理論は LLM により “Gravitonal Theory” と名付けられた 5 音からなる音高体系を基盤とした理論である。この理論の生成は GPT-5.2 Thinking を用いて対話的に行われた。前述の不協和度を基に算出される安定度や緊張度については未実装の状態で生成された初期理論である。

Gravitonal Theory の入力周波数は LLM により生成した。440Hz の周波数を基準とした 5 音からなる周波数列

表 4 Gravitonal Theory における音名と役割

音名	周波数 [Hz]	役割
Ga	440.0	基準音・中心音として理論の重心を形成
Ve	495.0	Ga から上方へ向かう接近音・推進音
Li	528.0	緊張を形成し、解決先への動きを促す音
No	660.0	安定した到達点となる終止音・休止音
Zi	742.0	高域側の安定音として浮遊感を与える音

表 5 Gravitonal Theory における和音と役割

和音名	構成音	役割
C-GLN	Ga, Li, No	中心和音
B-VLN	Ve, Li, No	遷移和音
D-VNZ	Ve, No, Zi	緊張を含む浮遊和音

の生成を指示し、LLM が生成した値をそのまま採用した。安定度や緊張度のデータが入力されていない状態ではあるが、入力周波数列における 2 音の周波数比が単純な整数比に近いかどうかを基に LLM により各音や和音の役割が推論されている。

Gravitonal Theory は [440.0, 495.0, 528.0, 660.0, 742.0] という周波数列から生成されている。この入力周波数列から、LLM が音名と役割、和音とその役割、和音進行と旋律パターンを生成した。理論の名前も LLM が考案し、重力（中心性）に引っ張られる音楽理論という意味を込めて Gravitonal Theory と名付けられた^{*5}。

表 4 に、Gravitonal Theory における音名、周波数、役割を示す。本生成例では、440.0 Hz の音が基準音として位置付けられ、Ga と命名された。また、660.0 Hz の No および 742.0 Hz の Zi は比較的安定した音として位置付けられ、495.0 Hz の Ve および 528.0 Hz の Li は緊張形成や展開に関わる音として位置付けられた。

表 5 に、Gravitonal Theory における和音の例を示す。各和音には、構成音とともに、中心形成、緊張形成などに相当する役割が付与されている。また、5 音の中から 3 音の和音を構成する組み合わせは 10 通りあるが、Gravitonal Theory では 3 種類の和音のみが生成された。

以上のように、Gravitonal Theory では、入力周波数列に対して音名と役割を付与し、さらに和音を生成することで、5 音からなる音高体系を作曲や演奏に利用できる理論としてまとめている。ここで示した名称や役割は、LLM が生成した初期的な理論記述であり、今後の作曲、演奏、聴取評価を通じて更新されることを前提としている。

Gravitonal Theory の詳細については、本論文と同時に発表予定の論文 [17] に記述する。

4.1.2 12 音版 Gravitonal Theory

初めに生成した音楽理論 Gravitonal Theory では、安定度および緊張度がまだ考慮されていなかった。そこで、4.1.1 節の 5 音および 3 つの和音の安定度と緊張度を算出

^{*5} この理論では tonality は仮定されていない点には注意されたい。

表 6 Gravitonal Theory 12 音版を構成する音

音名	周波数 [Hz]	cent 値	役割
Ga	440.00	0.0	理論の基底音であり旋律や進行の出発点となる音
Gave	472.94	125.0	Ga から Ve へ向かう運動を開始する起動音
Ve	495.00	203.9	Li や No へ向かう遷移を生じさせる音
Li	528.00	315.6	高い緊張を持ち、No への運動を形成する媒介音
Lino-I	564.06	430.0	Li の緊張を維持しながら No 方向へ進む接続音
Lino-II	599.32	535.0	Li 領域と No 領域の間を結ぶ橋渡し音
Lino-III	631.30	625.0	No への接近と収束を強める前中心音
No	660.00	702.0	最も安定した中心音であり、主要な到達点となる
Nozi	708.62	825.0	No の安定性を保ちながら Zi へ展開する音
Zi	742.00	904.7	浮遊感を伴う高位の副中心音
Ziga-I	797.69	1030.0	Zi から上方の Ga へ向かう帰還運動を開始する音
Ziga-II	842.69	1125.0	上方の Ga への帰還を準備する再進入音

し、構成音を 12 音へ拡張するよう LLM に指示した。なお、これ以降の音楽理論生成事例では、GPT-5.5 を利用している。

その結果、表 6 に示すような 12 音版の Gravitonal Theory の構成音が得られた。このように、LLM は柔軟な指示に対応可能なため、生成済みの音楽理論を再構築することも可能である。12 音への拡張にあたっては、12 平均律の構成音の音高に近い音なるべく採用しないようにプロンプトで制約を加えている。表 6 の結果では、既存の 5 音の間に位置する音の名称が既存音名を結合させただけの単純なものであったり、Lino-I から Lino-III までは番号で区別されている点など、実用性を考えると扱いにくい生成結果となっている。一方で本手法では LLM を用いていることから、ユーザの要望に応じた生成結果の修正も可能であり、生成した理論をそのまま使うのではなく、人と AI の協調による音楽理論の確立の作業に繋げることが重要となる。

4.2 ラップのフロウから生成した音楽理論

著者が知る限り、ラップにおける旋律に相当する概念である「フロウ」を基に、音名、和音、進行などを含む音楽理論を構築した研究や事例は報告されていない。そこで、本手法における入力周波数列として、ラップ歌唱の周波数を分析し、使用頻度が高い音高を 7 つ選択してラップのフロ

表 7 ラップ 38 曲分のフロウから生成した音楽理論「双錨環流律」の主要要素の抜粋

種別	名称	周波数・構成	役割・特徴
音	ナル	133.50 Hz	低域の第一支点
音	ネリ	136.21 Hz	低域の第二支点
音	モラ	189.21 Hz	中心的な接続音
音	オリン	252.57 Hz	上域の収束点
和音	ナル・モラ軸	ナル-モラ-オリン	理論の中心和音
和音	裂錨節	ネリ-ザイル-モラ	展開的な緊張
和音進行	裂節融解	裂錨節 → ヴェヤ渡環 → ナル・モラ軸	低域側への解決
旋律パターン	双錨揺動	ナル → ネリ → ナル	安定領域の内部運動

ウからの音楽理論生成を試みた。ラップでは、大きなピッチ変化が頻繁には起こらないことも多く、特定の周波数帯にデータが集中する傾向がある。そこで、1 オクターブを 100cent 毎に区切り、各区分で最も出現頻度が高い音高の音を 12 個選択し、その中から上位 7 つの音を採用した。なお、すべての音高は、オクターブ等価性に基づき 1 オクターブ内に折り畳んだ。

本事例では、伴奏トラックの分析により各楽曲の調を特定し、ラップ歌唱を含む楽曲全体を共通の調に移調した上で、38 曲分のラップ歌唱のみを抽出し、フレーム毎の基本周波数の頻度を基に入力周波数列を構成した。この周波数列に基づいて生成した音楽理論「双錨環流律」の一部を表 7 に抜粋して示す。原曲の伴奏トラックはいずれも 12 平均律に基づいて制作されているが、双錨環流律の音高体系に基づいて伴奏トラックを再構成することで、ラップのフロウとの親和性が高まるかどうかを検証することは今後の興味深い研究課題である。

ラップのフロウには、音高の要素だけでなく声の強弱や発声のタイミング、リズムなどの要素も含まれる。これらの要素は、本研究で対象とする音高体系を中心とした音楽理論のみでは十分に記述できない。今後、音楽理論の生成対象を音高体系以外の要素にも拡張することで、ラップのフロウにみられる複合的な音楽表現を、より忠実に記述できる可能性がある。

4.3 鳥の鳴き声から生成した音楽理論

鳥の鳴き声を、人間の歌声や音楽と同様に音高や音程の観点から分析する研究が行われている。たとえば、Doolittle らは、チャイロコツグミのさえずりに含まれる音高間の関係を分析し、多くの歌型において、倍音列に対応する単純な周波数比が選択されていることを示した [18]。

一方、鳥の鳴き声が、人間の音楽における 12 平均律のような人が定めた既存の音高体系に基づいているとは限ら

ない。鳴き声に含まれる音高の分布や音高間の関係には、種や個体に固有の特徴が存在する可能性があり、それらを人間の既存の音高体系とは異なる枠組みによって記述できる可能性がある。本節では、カラスおよびウグイスの鳴き声から音楽理論を生成した事例を示す。

4.3.1 カラスの鳴き声からの音楽理論生成

カラスの鳴き声を収録した音声データから基本周波数を抽出し、その分布に現れるピークに基づいて代表周波数を選択した。分析には、Web 上でフリー素材として公開されているカラスの鳴き声の音声データを用いた [19]。

約 38 秒分の鳴き声のデータから基本周波数を抽出した。基本周波数の推定には `librosa.pyin` を用い、推定された各フレームのうち、基本周波数が未定義ではなく、有声音確率が閾値以上で RMS 振幅が一定以上のフレームのみを分析対象とした。さらに、局所中央値から大きく外れた基本周波数は外れ値として除去した。得られた基本周波数は、最小周波数を基準として 1 オクターブ内に折り畳んだ上で cent 値に変換した。

その後、cent 値のヒストグラムを作成し、ガウス平滑化を施した上でピーク検出を行った。検出されたピークのうち、出現数およびピークの明瞭さに基づいて最大 7 個の周波数を選択し、各ピーク近傍に含まれる基本周波数の中央値を代表周波数とした。

以上の処理で得られた入力周波数列を以下に示す。

[355.06, 376.18, 566.89, 583.50, 651.19] Hz

この周波数列を本手法に入力して音楽理論を生成した。生成された理論は「黒翼双錨律」と名付けられ、5 個の音、5 個の和音、5 個の旋律パターン、および 4 個の和音進行を含む。そのうち、理論の特徴を表す主要な要素を表 8 に抜粋して示す。なお、理論名を生成する際の参考情報として、入力周波数列がカラスの鳴き声に基づくものである旨は LLM に与えている。

4.3.2 ウグイスの鳴き声からの音楽理論生成

ウグイスの鳴き声を収録した音声データからも周波数列を取得し、音楽理論を生成した。鳴き声の音声データや周波数列の取得方法は 4.3.1 節と同様である。分析対象とした音声データの長さは約 20 秒分であった。

得られた入力周波数列を以下に示す。

[1038.64, 1159.12, 1228.05, 1286.13, 1451.99] Hz

この周波数列を本手法に入力して音楽理論を生成した。生成された理論は「翠啼双峰律」と名付けられ、5 個の音、5 個の和音、5 個の旋律パターン、および 5 個の和音進行を含む。そのうち、理論の特徴を表す主要な要素を表 9 に抜粋して示す。なお、理論名を生成する際の参考情報として、入力周波数列がウグイスの鳴き声に基づくものであることを LLM に与えている。

4.4 12 平均律の C メジャースケールの音高列からの音楽理論生成

C メジャースケールを構成する鍵盤上の白鍵 7 音に対応する周波数列を入力し、既存の音高体系を用いた場合に、どのような音楽理論が生成されるかを確認した。その結果、従来の C メジャースケールにおける機能関係とは異なる音や和音に中心的な役割を与えた音楽理論「中脊昇冠律」が生成された。

このような理論が生成された要因の一つとして、LLM への指示に、「既存の長調や短調、C メジャー、ドレミなどの名称に依存しない」という制約を設けていることが挙げられる。また、本手法では、Sethares のモデルに基づいて暫定的に算出した安定度および緊張度を入力情報として用いており、これらの数値に基づく場合、従来の C メジャーにおける主和音が必ずしも最も安定した和音として評価されなかったことも従来とは異なる理論が生成された要因と考えられる。

ただし、従来の音楽理論で使用される周波数列を入力した場合に、その理論が再現されることが必ずしも最善であるとは限らない。同一の音高集合であっても、従来とは異なる役割や関係性を与えることで、別の原理に基づく作曲や演奏を行うことが可能である。したがって、本手法により生成される音楽理論は、従来の音楽理論を否定または置換するものではなく、既存の音高集合に対する異なる解釈や、新たな音楽実践のための選択肢として位置付けることもできる。

5. 評価

本手法により生成された音楽理論をどのように評価するかは、本研究における重要な課題である。生成された理論については、実際に各音や和音を聴取した際の印象と、LLM によって付与された安定、緊張、中心性などの役割との整合性を調べることで、その妥当性の一部を検証できると考えられる。さらに、旋律パターンや和音進行の一貫性、理論の理解しやすさ、および作曲や演奏への利用可能性についても、人による評価が必要である。

一方、本手法では入力周波数列ごとに異なる音楽理論が生成されるため、評価対象となる理論や評価基準を一意に定めることは難しい。また、音楽理論を多数かつ短時間に生成できるのに対し、聴取、演奏、作曲などを伴う人手評価には相応の時間を要する。したがって、生成されたすべての理論を同じ方法で網羅的に評価することは現実的ではなく、評価対象の選定方法を含めた枠組みを検討する必要がある。

従来の音楽理論が長期にわたる音楽実践の歴史の中で形成・更新されてきたことを踏まえると、本手法により生成された理論についても、一度の生成によって完成したものとはみなすのではなく、その後の検討や実践に向けた初期案

表 8 カラスの鳴き声から生成した音楽理論「黒翼双錨律」の主要要素の抜粋

種別	名称	周波数・構成	役割・特徴
音	Krau	355.06 Hz	旋律の出発点となる低域の基準音
音	Vek	376.18 Hz	副次的な収束点として機能する安定音
音	Raan	566.89 Hz	低域と高域を跳躍によって結ぶ接続音
音	Ghir	583.50 Hz	旋律や和音に強い緊張を与える音
音	Moro	651.19 Hz	終止や帰還の中心となる最も安定した音
和音	Moro Nest	Vek--Raan--Moro	終止や休止に用いる最も安定した中心和音
和音	Raan--Ghir Knot	Raan--Ghir--Moro	解決直前の最大緊張を形成する和音
旋律パターン	Vek--Raan--Ghir--Moro Arc	Vek → Raan → Ghir → Moro	大きな跳躍と微細な音高変化を経て中心音へ収束

表 9 ウグイスの鳴き声から生成した音楽理論「翠啼双峰律」の主要要素の抜粋

種別	名称	周波数・構成	役割・特徴
音	Nagi	1038.64 Hz	低域の安定音であり、開始点や低位終止点となる音
音	Sayo	1159.12 Hz	低域から中央領域への接近を担う接続音
音	Teri	1228.05 Hz	最も緊張度が高く、進行の方向を変える支点となる音
音	Hane	1286.13 Hz	中央の緊張領域と高域の安定音を結ぶ音
音	Mine	1451.99 Hz	最も安定度が高く、上昇の到達点や終止音となる音
和音	双峰澄座	Nagi--Hane--Mine	二つの安定音を含む中心和音であり、開始や終止に用いる和音
和音	三啼裂	Sayo--Teri--Hane	中央の近接した三音により強い緊張を形成する和音
旋律パターン	翠枝昇り	Nagi → Sayo → Teri → Hane → Mine	緊張を段階的に高めながら高域の安定音へ到達する
和音進行	裂啼帰座	三啼裂 → 高枝導環 → 双峰澄座	強い緊張から高域への接近を経て中心和音へ解決する進行

として位置付けることが適切である。生成された理論を人が実際に聴取、演奏、および作曲に利用し、得られた評価や指摘を基に修正することで、より一貫性および実用性の高い理論へと更新できる可能性がある。このため、初期段階における評価結果は、理論の更新に伴って変化し得るものとして扱う必要がある。

以上の理由から、本稿では生成された音楽理論に対する人手評価を実施しておらず、今後の課題として位置付ける。今後は、人による評価結果を LLM への追加情報として与えて理論の修正や再生成に反映する、人間と LLM が協働した理論の評価・更新の枠組みについて検討する。

5.1 既存音律との比較

人手での評価に代えて、基礎的な数値評価を通じて生成された音高体系と既存音律との比較を行う。比較対象は、4.1 節から 4.3 節までに示した 5 つの生成理論、純正律、および 12 平均律である。ただし、ここでは各理論における音の役割、和音、旋律パターン、および和音進行の妥当性を比較するのではなく、理論を構成する音高の配置を音律として比較する。この音律同士の比較を通じて、本手法によって既存音律とは異なる音響的特徴を持つ多様な音高体系が構成できることを確認する。

音律の比較指標として、3.4.1 節で述べた Sethares のモデルに基づき、各音律に含まれる異なる 2 音のすべての組合せについて不協和度を算出し、その平均値を求める [11]。また、各音律の構成音を周波数の昇順に並べ、隣接する音同士の音程を cent 単位で算出し、その平均値および標準偏

差を算出する。

12 平均律および純正律については、C4 を 261.63 Hz とし、C4 以上 C5 未満の 1 オクターブに含まれる 12 音を対象として各指標を算出する。4.1 - 4.3 節の周波数列についても、同様に 261.63Hz を含む周波数帯における 1 オクターブ分の音についての数値を算出する。音律の比較結果を表 10 に示す。

平均不協和度に注目すると、今回比較した音高体系の中では、純正律が最も低い値を示した。一方、本手法では入力周波数列を任意に設定できるため、不協和度を選択基準として周波数列を探索することで、今回用いた純正律よりも平均不協和度の低い音高体系が得られる可能性もある。以上の結果から、本手法により生成された各音高体系はそれぞれ異なる音響的特徴を持ち、多様な音楽理論を構築するための候補となり得ることが示唆される。

5.2 評価における課題と展望

前述の通り、人手での評価は今後の課題である。本手法では、入力周波数列を変更することで多数の理論を生成できる一方、理論の良し悪しは単一の観点から評価することは困難であり、長期的な活用を伴う多面的な評価が不可欠である。作曲がしやすい理論であっても、演奏が容易ではない理論など、理論の評価は複数の観点から実施しなければならない。

一方、異なる音楽理論の音高体系に基づいて既存楽曲を編曲したり、新たな楽曲を制作したりすることが容易になれば、創作に利用できる音楽的選択肢を拡張できる可能性

表 10 生成された音高体系と既存音律の数値的比較

生成理論・音律	音数	平均不協和度	隣接音程平均 [cent]	隣接音程標準偏差 [cents]
Gravitonal Theory (5 音版)	5	0.305	226.17	99.73
Gravitonal Theory (12 音版)	12	0.289	102.27	18.14
双錨環流律 (ラップ)	7	0.382	183.83	155.72
黒翼双錨律 (カラス)	5	0.351	262.50	263.17
翠啼双峰律 (ウグイス)	5	0.405	145.00	55.90
純正律	12	0.260	98.93	20.19
12 平均律	12	0.263	100.00	0.00

がある。将来的に、一人一人が目的や嗜好に応じて独自の音楽理論を生成し、創作、演奏、および鑑賞に利用する環境が実現すれば、実践を通じて有用な理論が選択され、改良されていくことも考えられる。

音楽理論の良し悪しは、その理論に基づいた名演奏や名曲が実現するかといった後からの評価にも左右され得る。したがって、本手法の有効性を評価するためには、生成直後の理論に対する評価に加えて、理論の継続的な利用と更新を含む評価方法を検討する必要がある。今後は、人による聴取、作曲、演奏および理解の観点を組み合わせ、本手法に適した評価・更新の枠組みについても検討していく。

6. 音楽理論生成の応用事例

音楽理論を生成するだけではまだ音楽にはならない。その理論に基づく作曲や演奏ができて初めて価値が生まれることになる。任意の周波数列に含まれる各音に名称や役割を与え、音同士の関係、和音、および進行を体系化することは、既存の音高体系にとらわれない音を作曲や演奏に利用する上で有用である。したがって、本手法は、新たな音高体系を音楽創作へと接続するための基礎技術として位置付けられる。

本章では、本研究の応用事例として、生成された音楽理論に基づく作曲事例や既存楽曲の編曲、演奏のためのインタフェース、音楽理論生成のためのフレームワークを紹介する。

6.1 生成した音楽理論に基づく自動作曲

本手法で生成した音楽理論は、JSON による構造化と自然言語による説明文章により表現されている。そのため、これらを LLM に入力し、作曲条件を追加のプロンプトとして与えることで、その理論に基づく曲を生成させることができる。

LLM に Gravitonal Theory (5 音版) として生成済みの音名および和音体系を用いて、短い作曲例を生成するよう指示した。このとき、既存の長調・短調や機能と声に基づく作曲ではなく、Gravitonal Theory (5 音版) における音の役割と和音機能を反映した旋律および和音進行を作成するよう指定した。その結果、LLM は 4 小節からなる作曲

例 “No Drift I” を生成した。“No Drift I” の詳細については [17] にまとめている。

生成された音楽理論に基づく自動作曲は、LLM を用いる方法に限られない。JSON 形式で記述された各音の役割、音同士の関係、和音、旋律パターン、および和音進行を作曲規則として利用することで、既存のルールベース自動作曲手法にも適用できる。さらに、特定の生成理論に基づく楽曲データが十分に蓄積されれば、それらを学習データとした機械学習ベースの自動作曲手法への展開も考えられる。著者はこれまでに、微分音を含む旋律を対象とした生成モデルを提案している [20]。今後は、これらの手法を応用し、多くの人にとって耳慣れない微分音を使った作曲を支援する技術についても検討していく。

6.2 既存の 12 平均律メロディの生成された音楽理論に基づく編曲

既存のメロディを構成する各音高を、生成された音楽理論の構成音に置き換えることで、元の旋律的特徴を保持しつつ、異なる音高体系に基づくメロディへ変換できる。本事例では、既存メロディの各音高に対して、cent 値による音程距離が最小となる生成理論の構成音を割り当てた。

具体例として、カラスを題材とした童謡「七つの子」のメロディを、カラスの鳴き声から生成した音楽理論「黒翼双錨律」の構成音を用いて編曲した。その結果、原曲の旋律輪郭をある程度保持しながら、黒翼双錨律の音高体系に基づく新たなメロディを生成できた。一方、黒翼双錨律が 5 種類の音高から構成されるのに対し、「七つの子」のメロディには 6 種類の音高が含まれるため、原曲における異なる 2 種類の音高が、同一の構成音へ割り当てられる箇所が生じた。その結果、原曲に含まれる一部の音程関係が失われ、旋律の再現性が低下した。

この問題に対しては、生成する音楽理論の構成音数を増やすことで、音高の分布がより高解像度となり、元の旋律の再現度が向上する可能性がある。ただし、原曲への忠実度と、生成された音楽理論固有の音高的特徴との間にはトレードオフがあるため、編曲の目的に応じて割当方法を調整する必要がある。

既存の楽曲を長調から短調へ編曲するような実践は広く

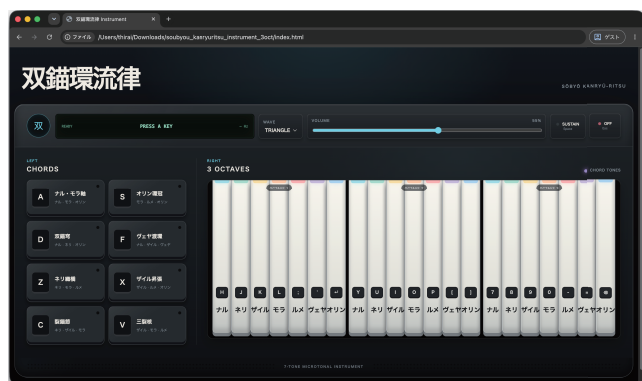


図 1 双錨環流律の演奏インタフェース

行われているが、本手法を用いることで、既存の長調・短調とは異なる音高体系に基づいて楽曲を再構成できる可能性が示された。

6.3 新しい音楽理論に基づく音楽演奏

著者らはこれまでに、微分音を演奏可能な XR 鍵盤楽器 [21] および XR 弦楽 [22] や、微分音を容易に演奏するためのインタフェース [23] を提案してきた.

本事例では、これまでに提案した演奏インタフェース [23] の設計思想を双錨環流律に適用し、同理論に基づく演奏インタフェースを実装した。実装に際しては、双錨環流律についての JSON ファイルとインタフェースに求める機能を LLM に与え、ブラウザ上で動作する演奏インタフェースとして、プログラムのコードを生成した。

実装したインタフェースを図 1 に示す。このインタフェースでは、PC キーボードの右手側に割り当てたキーを用いて、双錨環流律を構成する 7 種類の音を 3 オクターブにわたって演奏できる。また、左手側に割り当てたキーを用いることで、理論内で定義された和音を一つのキー操作で演奏できる。生成された音楽理論を利用者自身が試奏できる状態にすることは、各音や和音の響きと、理論上付与された役割との関係を体験的に理解する上で有用である。さらに、未知の音高体系の音を実際に演奏することで、その理論を用いた作曲や演奏に関する新たな発想を促す可能性がある。

本章で示した自動作曲，既存メロディの編曲，および演奏インタフェースには，それぞれ独立した研究課題が残されている．対象とする音楽理論が変化すれば，自動作曲，楽曲分析，編曲，演奏支援，記譜，およびインタフェース設計に求められる方法も変化すると考えられる．このように，生成理論に応じた演奏インタフェースの設計自体も，今後の研究課題となる．

6.4 音楽理論生成フレームワーク

本論文で述べた音楽理論生成手法は，各処理が独立しており，現状では，周波数列の入力から音楽理論の生成，各

音や和音の試聴，および演奏インタフェースの生成までを，一貫したユーザインタフェース上で実行できる状態にはない．そこで今後は，ユーザが直接入力した周波数列，またはアップロードした任意の音素材から抽出した周波数列を基に，音楽理論の生成から，6.3 節で述べたような演奏インタフェースの生成までを一貫して行えるフレームワークの開発に取り組む．現在開発中のプロトタイプを図 2 に示す．図 2 のプロトタイプは，一部のみが実装されている状態である．

このプロトタイプにおける最も重要な機能として、LLMとの対話を通じて音楽理論を更新する機能が挙げられる。今後、ユーザがフィードバックを与え、その内容を基にLLMが理論を修正または再生成することで、人とLLMの協働により理論を段階的に洗練させていけるシステムとすることを目指している。

著者はこれまでに、ユーザがオリジナルの音律を作成するインタフェース [24] を提案している。今後、このような研究成果も取り入れていくことで、既存の音律や音楽理論の枠組みにとらわれない音高体系の設計から、音楽理論の生成、作曲、試聴、および演奏までを一貫して支援する、統合的な音楽創作・演奏支援技術の実現を目指している。

7. まとめ

本稿では、LLMを用いて任意の周波数列から音楽理論を生成する手法について紹介した。本手法では、入力周波数列と、安定度、緊張度などの数値情報を LLM に与え、各音の名称と役割、和音、旋律パターン、および和音進行を生成する。また、ラップのフロウ、鳥の鳴き声から抽出した周波数列を入力とした生成事例を示すとともに、生成された理論に基づく自動作曲、既存メロディの編曲、および演奏インタフェースへの応用可能性を示した。

LLM の出力内容にはハルシネーションが含まれたり、同じ入力に対しても試行ごとに異なる結果が生成されるなどの問題もあり、初期の出力結果への信頼度は十分とは言えない。LLM には耳がないため、LLM は生成された各音や和音の響きを実際に聴取して評価しているわけではない。そのため、生成された理論は、人による聴取、演奏、作曲、および分析を通じて検証・更新される初期案として位置付ける必要がある。生成された音楽理論の初期案を音楽に繋げる役割は人間が担うことになる。生成された理論は未完成なものであり、人による評価の上での更新が不可欠である。

既存の音楽理論が長期にわたる実践を通じて形成・更新されてきたように、生成された理論についても、継続的な利用と評価を通じて洗練することが重要である。そのような観点では、音楽理論には完璧な形というものは存在せず、音楽という膨大な表現領域の一部を便利に記述するための道具とも見做せる。音楽理論の生成というタスクは、

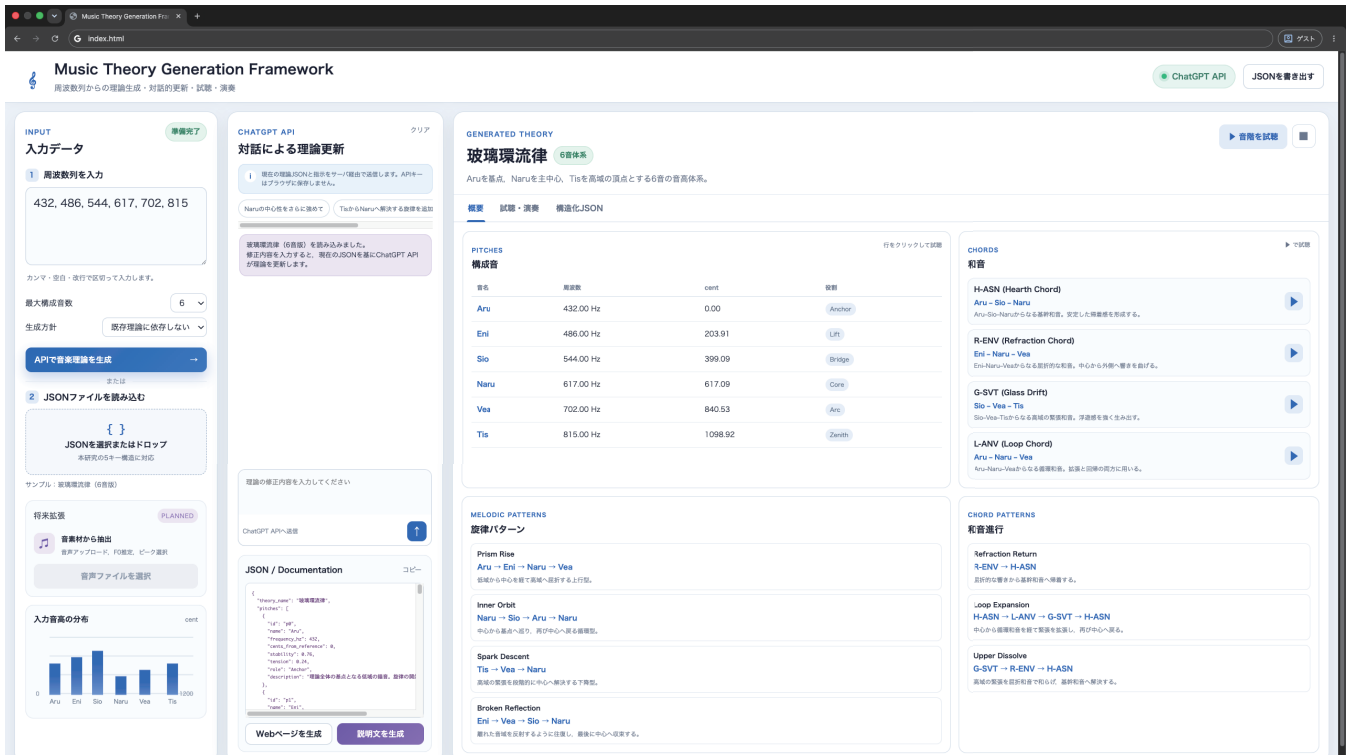


図 2 音楽理論生成フレームワークのプロトタイプ. 最も左のパネルに周波数列を入力し, 右側のパネルで生成された音楽理論を確認でき, 中央のパネルで LLM と対話的に音楽理論が更新できる (一部機能については未実装)

これまで音楽理論が網羅できていなかったような音楽・音高体系にまで理論的な土台を提案しうる技術であると考えている。

将来的に人間の感性に近い聴覚を持った AI が出現したとしても, 音楽を記述するルールのすべてを AI が担うような未来は考えにくい。生成された理論を採用するか, どのように修正し, どのような音楽に利用するかは, 最終的に人による判断と実践に委ねられる。音楽は人が主体として創造するものであり, AI はあくまでもその創造力を発揮するための支援をしていくものと考えている。

今後の重要な研究課題の一つとして, 12 平均律とは異なる歌唱や発声に対して, その音高的特徴に適合する音楽理論を生成し, 伴奏に適用することの効果を検証することが挙げられる。他にも多くの派生研究が考えられ, 本研究が, 既存の音高体系に限定されない新たな音楽情報処理研究を展開するための出発点となることを期待する。

謝辞 本研究の一部は JSPS 科研費 JP23K17023 の助成を受けたものである。

参考文献

- [1] Schafer, R. M.: *The Tuning of the World*, Alfred A. Knopf, New York (1977).
- [2] Schaeffer, P.: *Traité des objets musicaux: Essai interdisciplinaires*, Éditions du Seuil, Paris (1966).
- [3] 吉松悠太: SoundQuest - 新しい音を見つけよう, <https://soundquest.jp/>. 参照 2026-07-23.
- [4] Gotham, M., Shaffer, K., Hughes, B. and Moseley, B.: *Open Music Theory*, VIVA Open Publishing (2021). Version 2, open educational resource.
- [5] Forte, A.: *The Structure of Atonal Music*, Yale University Press (1973).
- [6] Straus, J. N.: *Introduction to Post-Tonal Theory*, W. W. Norton & Company, New York, 4 edition (2016).
- [7] Partch, H.: *Genesis of a Music: An Account of a Creative Work, Its Roots and Its Fulfillments*, Da Capo Press, New York, 2 edition (1974).
- [8] Wyschnegradsky, I.: *Manual of Quarter-Tone Harmony*, Underwolf Editions, Brooklyn, NY (2017). Originally published in 1932.
- [9] Blackwood, E.: *The Structure of Recognizable Diatonic Tunings*, Princeton University Press, Princeton, NJ (1985).
- [10] Fokker, A. D.: Equal Temperament and the Thirty-One-Keyed Organ, *The Scientific Monthly*, Vol. 81, No. 4, pp. 161–166 (online), available from <http://www.jstor.org/stable/21995> (1955).
- [11] Sethares, W. A.: *Tuning, Timbre, Spectrum, Scale*, Springer, London, 2 edition (2005).
- [12] Briot, J.-P., Hadjeres, G. and Pachet, F.: *Deep learning techniques for music generation*, Springer (2020).
- [13] Agostinelli, A., Denk, T. I., Borsos, Z., Engel, J., Verzett, M., Caillon, A., Huang, Q., Jansen, A., Roberts, A., Tagliasacchi, M., Sharifi, M., Zeghidour, N. and Frank, C.: MusicLM: Generating Music From Text, *arXiv preprint arXiv:2301.11325* (2023).
- [14] Copet, J., Kreuk, F., Gat, I., Remez, T., Kant, D., Synnaeve, G., Adi, Y. and Défossez, A.: Simple and Controllable Music Generation, *Advances in Neural Information Processing Systems*, Vol. 36 (2023).

- [15] Hadjeres, G., Pachet, F. and Nielsen, F.: DeepBach: A Steerable Model for Bach Chorales Generation, *Proceedings of the 34th International Conference on Machine Learning*, pp. 1362–1371 (2017).
- [16] Good, M.: MusicXML: An Internet-Friendly Format for Sheet Music, *Proceedings of XML 2001*, Boston, MA (2001).
- [17] 平井辰典, 駒形太郎: AI 生成音楽理論: Gravitonal Theory, 情報処理学会研究報告音楽情報科学(MUS), Vol. 2026-MUS-147 (2026).
- [18] Doolittle, E. L., Gingras, B., Endres, D. M. and Fitch, W. T.: Overtone-based pitch selection in hermit thrush song: Unexpected convergence with scale construction in human music, *Proceedings of the National Academy of Sciences*, Vol. 111, No. 46, pp. 16616–16621 (2014).
- [19] Killy: 効果音ラボ, <https://soundeffect-lab.info/>. 参照 2026-07-23.
- [20] Hirai, T.: Application of a Deep Generative Model to Microtonal Music, *Proceedings of 50th International Computer Music Conference (ICMC 2025)*, pp. 521–528 (2025).
- [21] Hirai, T., Topliss, J. and Piumsomboon, T.: XR Musical Keyboard: An Extended Reality Keyboard with an Arbitrary Number of Keys and Pitches, *Proceedings of the International Conference on New Interfaces for Musical Expression*, pp. 40–45 (2025).
- [22] 安引海翔, 長川原颯人, 青柳西藏, 服部哲, 平井辰典: 微分音コードが発音可能な XR 弦楽器システムの構築と評価, 情報処理学会研究報告音楽情報科学 (MUS), Vol. 2025-MUS-143, No. 73, pp. 1–7 (2025).
- [23] 中原 丈, 平井辰典: 12 平均律と微分音単音を統合したリアルタイム演奏支援システム, 情報処理学会研究報告音楽情報科学 (MUS), Vol. 2026-MUS-145, No. 39, pp. 1–6 (2026).
- [24] Hirai, T.: Redesigning a Piano Roll: A Melody Input Interface That Can Play Microtones with an Arbitrary Number of Keys, *Proceedings of 19th Sound and Music Computing Conference (SMC 2022)*, Zenodo, pp. 1–7 (online), DOI: 10.5281/zenodo.6797825 (2022).