GAMuTh: Guide to Advanced Music Theory

Release 0.0.1

Fabian C. Moss

Jan 28, 2022

CONTENTS

Warning: The content on these pages is very much under construction! Since this is ongoing work, I can give no guarantee for completeness or accuracy. Feel free to contact me with your questions and suggestions!

Welcome!

This is not a pedagogical resource for basic music theory concepts but an in-depth introduction into the structures of Western music, built axiomatically from tones and their relations. The logo, a maxima, the longest note value in medieval music, symbolically reflects this level of difficulty.

The content presented on these pages is inspired by a number of great books, e.g. Aldwell *et al.* [1], Lewin [15], Straus [38], Cadwallader and Gagné [4], and Müller [26]. What is new and unique about the approach taken here is that we take a computational perspective and implement all introduced concepts. This does not only provide us with sharp and unequivocal definitions, but also allows us to scale music theory up from the analysis of individual bars, sections, or pieces to that of entire repertoires and corpora!

I recently also discovered Music for Geeks and Nerds by Pedro Kroger which looks very interesting. The Python project musthe also seems to pursue a similar goal.

Note: Some content is only available in the online HTML version and not in the PDF, e.g. scores rendered with Lilypond. Please visit the **website** to see the full version.

Note: TODO: add link to website

Quickstart

Warning: These instructions may not work yet.

Installation of the *gamuth* Python library is as easy as it can be. Just type the following in your terminal:

```
pip install gamuth
```

Then, in a Python script or Jupyter notebook, import the library or individual classes:

```
>>> from gamuth import Tone, Interval

>>> t = Tone(octave=0, fifth=0, third=0) # C
>>> s = Tone(octave=0, fifth=0, third=1) # E

>>> i = Interval(t, s)

>>> print(i.specific interval) # directed interval in semitones (major third)
4
```

See the documentation *API - gamuth* for examples how to use the library, its classes and their methods.

CHAPTER

ONE

FUNDAMENTALS

The theory presented in here can be described as a *tonal theory* in the sense that its most fundamental objects are *tones*, discrete musical entities that have a certain location in tonal space. A tonal space is then a metrical space describing all possible tone locations, and the metric is given by an *interval function* between the tones. Note that by this definition, there are as many different tonal spaces as there are interval functions.

While many aspects and examples will be taken from Western (classical) music, the theory is in principle not restricted to this tradition but extends well to virtually all musical cultures where a tone is a meaningful concept.

Perhaps the most simple description of music is *sound organized in time* (attributed to Edgar Varèse, see also [45]). Later we will see that this description falls short of encompassing many central aspects to music but it provides a good starting point for our considerations. Taking this definition for granted means that we can conceptualize music within a two-dimensional framework, where the axes represent sound and time, respectively (see Fig. 1.1). This is also the

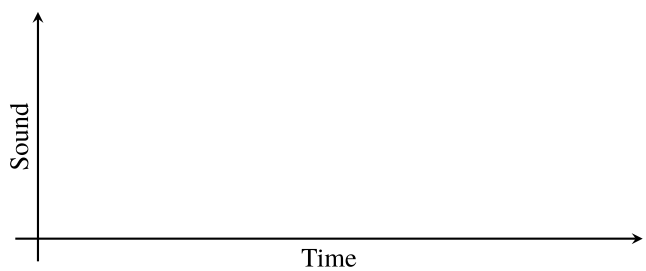

Fig. 1.1: Two-dimensional depiction of music: music as sound organized in time.

way music is usually displayed in *Digital Audio Workstations* (DAW) that feature a master window where music blocks can be arranged along a fixed timeline. Producing music in these environments thus quite literally consists of stacking blocks on top of one another.

1.1 Tones

Let's start with a mental exercise: imagine a tone. Contemplate for a while what this means. Does this tone have a pitch? A duration? A velocity (volume)?

- Riemann (1916). *Ideen zu einer Lehre von den Tonvorstellungen*:

"The ultimate elements of the tonal imagination are single tones." (Wason and Marvin [43], p. 92).

Bearing that in mind, let's create (or *instantiate*) a tone. To do so, we need to conceptualize ("vorstellen" in Riemann's terminology) a *tone location* ("Tonort", Mazzola [17], p. 241). There are many different ways to do this. In fact, the way we specify the location of a tone defines the tonal space in which it is situated. The figure below is an adaptation from Lewin [15]. The fact that music operates with discrete pitches has also been argued to be crucial for its evolution [41].

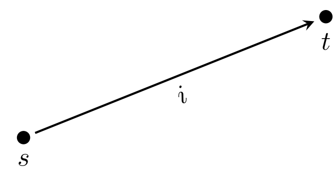

Fig. 1.2: Two abstract tones s and t, and the interval i between them.

1.1.1 Frequencies

Each tone corresponds to some *fundamental frequency* f in Hertz (Hz), oscillations per second.

- Overtone series
- frequency ratios
- logarithm: multiplication => addition

1.1.2 Euler Space

One option is to locate a tone *t* as a point $p = (o, q, t)$ in Euler Space, defined by a number of octaves o, fifths q, and thirds *t*. We will use the `gamuth.Tone` class for this

```
from gamuth import Tone

t = Tone(o=0, q=0, t=0)
```

From this representation we can derive a variety of others, corrsponding to transformations of tonal space.

1.1.3 Octave equivalence

Octave equivalance considers all tones to be equivalent that are separated by one or multiple octaves, e.g C1, C2, C4, C10 etc. More precisely, all tones whose fundamental frequencies are related by multiples of 2 are octave equivalent.

1.1.4 Tonnetz

The *Tonnetz* does not contain octaves and thus corresponds to a projection

$$\pi : (o, q, t) \mapsto (q, t).$$

1.2 Pitch classes

A very common object in music theory is that of a *pitch class*. Pitch classes are equivalence classes of tones that incorporate some kind of invariance. The two most common equivalences are *octave equivalence* and *enharmonic equivalence*.

1.2.1 Enharmonic equivalence

If, in addition to octave equivalence, one further assumes enharmonic equivalence, all tones separated by 12 fifths on the line of fifths are considered to be equivalent, e.g. A♯ and B♭, F♯ and G♭, G♯, and A♭ etc.

The notion of a pitch class usually entails both octave and enharmonic equivalence. Consequently, there are twelve pitch classes. If not mentioned otherwise, we adopt this convention here. The twelve pitch classes are usually referred to by their most simple representatives, i.e.

C, C♯, D, E♭, F, F♯, G, A♭, A, B♭, B,

but it is more appropriate to use *integer notation* in which each pitch class is represented by an integer $k \in \mathbb{Z}_{12}$.

$$\mathbb{Z}_{12} = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\},$$

and usually one sets $0 \equiv$ C. This allows to use *modular arithmetic* do calculations with pitch classes.

1.2.2 Other invariances

OPTIC

1.2.3 Tuning / Temperament

1.3 Intervals

- Pitch intervals
- Ordered pitch-class intervals (-> rather directed)
- Unordered pitch-class intervals
- Interval classes
- Interval-class content
- Interval-class vector

1.3.1 GISs

Transformations between representations of tones are actually *transformations of tonal space*.

[Diagram of relations between different representations.]

CHAPTER

TWO

SCALES AND MODES

A vast majority of musical cultures globally have some notion or concept of organization systems for discrete tones [3]. Those can be called *scales* or *modes*. While this general description holds, the musical reality is incredibly diverse and seems to defy uniform analytical approaches [18].

In this chapter, we are going to learn some tonal characteristics and organization systems from a range of musical cultures from different continents. While this will already hint at the global variation, it is by no means exhaustive. Nor are the descriptions below meant to entirely represent those cultures. Many factors, especially contextual sociopolitical ones, are out of the scope of this book, and we restrict the descriptions to what we can model using the approch established in the earlier chapter: a formal description of tonal relations.

Powers *et al.* [31]

2.1 Indian classical music

2.2 Turkish Maqam

Makam Dataset [14]

2.3 Arab-Andalusian music

Nuttall *et al.* [29]

2.4 Persian Music

FamourZadeh [9], Sanati [34]

2.5 Western classical music

- Ancient Greek modes [28]
- Ecclesiastic modes [2, 7, 8, 30, 44]
- Major and minor [13]
- Modes of limited transposition
- Georgian liturgic chorales (quartal harmonies) [33]

2.5.1 The diatonic scale

Music in the Western tradition fundamentally builds on so-called *diatonic* scales, an arrangement of seven tones that are named with latin letters from A to G. "Diatonic" can be roughly translated into "through all tones". Within this scale, no tone is privileged, so the diatonic scale can be appropriately represented by a circle with seven points on it. Mathemacally, this structure is equivalent to $\mathbb{Z}_7$.

[tikz figure here]

Now, if we want to determine the relative relations between the tones, it is necessary to assign a reference tone that is commonly called the *tonic*, or *finalis* in older music.

For example, if the tone D is the tonic, we can determine all other scale degrees as distance to this tone. Scale degrees are commonly notated with arabic numbers with a caret:

$$\begin{aligned} \mathrm{D} &: \hat{1} \\ \mathrm{E} &: \hat{2} \\ \mathrm{F} &: \hat{3} \\ \mathrm{G} &: \hat{4} \\ \mathrm{A} &: \hat{5} \\ \mathrm{B} &: \hat{6} \\ \mathrm{C} &: \hat{7} \end{aligned}$$

Taking these seven notes in scalar order, they can be converted to their *fifth order* via

$$\phi : t \mapsto 4t \mod 7$$

because the octave is divided into 7 steps and there and a fifths consists of 4 steps. Under this view the diatonic scale is a subsegment of the *line of fifths* [21, 39]

Exercise 2.1

How does the fifth order relate to the Euler Space / Tonnetz mentioned earlier?

Solution to Exercise 2.1

This is the solution

2.5.2 Modes

scale plus order plus hierarchy (but order already defined above?)

2.5.3 Keys

2.6 Jazz

2.7 Other scales

- chromatic
- hexatonic
- octatonic
- whole tone
- Messiaen

Before we move on to another important musical dimension, time, we have to consider one of the most famous musical scale systems (at least among music academics): Balinese Pelog and Slendro.

2.8 Balinese Pelog and Slendro

CHAPTER

THREE

PITCH-CLASS SET THEORY

Content adapted from Forte [10], Straus [38]. In this chapter, "pitch class" always entails octave *and* enharmonic equicalenve.

3.1 Pitch classes

3.1.1 Octave equivalence

Octave equivalence maps all pitch classes with the same label/name to one class, e.g. all C's on the piano get mapped to pitch class C, all Ab's get mapped to pitch class Ab, and all G#'s get mapped to G#.

Octave equivalence has perceptual correlates (cite paper) and is frequently employed in composition. For instance, the double bass sounds an octave lower as a cello but they are notated identically, or a fugue subject enters in different registers. The octave is the most fundamental interval with a frequency ratio of 2:1. Octave equivalence transforms the Euler space to the Tonnetz $\mathbb{Z} \times \mathbb{Z}$.

[diagram]

3.1.2 Enharmonic equivalence

Enharmonic equivalence describes the identification of pitches like F#4 and G#4. Enharmonic equivalence transforms the Euler Space to a tube, namely $\mathbb{Z}_{12} \times \mathbb{Z}$ (the chromatic circle and the octave line).

[diagram]

Taking octave equivalence and enharmonic equivalence together, we arrive at the torus that contains all twelve enharmonic (and octave-related) pitch classes, and which can be represented by numbers $k \in \mathbb{Z}_{12}$

[diagram]

Both equivalences are indepentend of one another, i.e. we have the following commuting diagram between tonal spaces:

[diagram]

Assuming enharmonic equivalence has a number of implications:

- we cannot distinguish between diatonic and chromatic semitones
- we cannot distinguish between augmented fourths and diminished fifths
- we cannot distinguish between dominant seventh chords and augmented sixth chords
- and more

Note: Describe relation to equal temperament.

For the remainder of this chapter, pitch classes designate entities for which both equivalences are assumed. It is thus appropriate to represent them on a circle. Since each pitch class can have infinitely many spellings, going the reverse direction is a difficult inference problem, for which a number of algorithms have been proposed (see advanced chapter on *Pitch Spelling*).

3.2 Intervals

Generally, intervals describe the distance between two points (see Chapter *Intervals*). But depending on which representation for tones we chose (or, equivalently, which equivalences we assume) the types of the intervals changes as well.

Note: Add paragraph about what types are.

3.2.1 Pitch Intervals

Pitch intervals describe the distance between two pitches in semitones. That is, we do assume enharmonic equivalence but not octave equivalence. Consequently, we can visualize pitch intervals in

Ordered pitch-class intervals

Unordered pitch-class intervals

Interval class

Interval class content

Interval class vector

3.3 Pitch-class sets

Normal form

Transposition

Inversion I

Index number

Inversion II

Set class

Prime form

Segmentation and analysis

3.4 Relationships

Common tones under transposition

Transpositional symmetry

Common tones under inversion

Inversional symmetry

Z-relation

Complement relation

Subset and superset relations

Transpositional combination

Contour relations

Composing out

Voice-leading

Atonal pitch space

3.5 Advanced concepts

Tonality

Centricity

Inversional axis

The diatonic collection

The octatonic collection

The whole-tone collection

The hexatonic collection

Collectional interaction

Interval cycles

Triadic post-tonality

3.6 Twelve-tone theory

Twelve-tone series

Basic operations

Subset structure

Invariants

CHAPTER

FOUR

TIME

- beats
- seconds
- onsets

4.1 Notes

(Tones + Duration) blablabla…

4.2 Rhythm

(Duration patterns)

4.3 Meter

(Hierarchy)

4.4 Musical time vs. performance time

CHAPTER

FIVE

HARMONY

- major-minor tonality
- chords: - root on (altered) diatonic scale degree - stack of thirds - chromatic alterations

First note based (Tone-based)

Then arrive at annotation standard (simple version) and do simple regex filtering (find all Vs…)

Moss *et al.* [23], Neuwirth *et al.* [27]

5.1 Transformational theory

We have already talked about GISs in Section 1.3.1

CHAPTER

SIX

PITCH-CLASS BASED MUSIC ANALYSIS

Musical pieces are made up of notes. This almost trivial truth is the starting point for this endeavour: to uncover hidden patterns in compositions dispersed through the centuries from the late middle ages to modern times.

First, we will develop an understanding which notes exist and what their mutual relations are. We will then proceed to use this information to discover aspects of the structural relations between them and how the constitute musical spaces (generalized interval systems).

We will see that musical notes in compositions are used far from randomly and that the order encoded in musical pieces reveals acquaintances between composers that sometimes transcends separation in time.

Finally, we will see how the usage of musical notes changes over the course of history. Observing large-scale changes in compositional style we will be able to distinguish historical trends.

…

The study is based on the music as it is notated. Moreover, we assume octave equivalence. These two assumptions–one given by the representation of notes in the data, the other based on theory and previous literature–leads to the Spiral Array representation of tonal space.

Tonality is constituted by the ways composers navigate pitch space in their compositions. This is, of course, hierarchal. Each piece constitutes a unique instance of tonality which nonetheless shares many properties with "similar" compositions (similar in style).

But what is pitch space?

6.1 Frequency space

A frequency f is measured in Hertz (1/sec). We can represent the set of fundamental frequencies as $F = \{K \cdot 2^o \cdot 3^q \cdot 5^t \mid o, q, t \in \mathbb{Q}\}$ for a fixed *Kammerton* K, and $o, q, t \in \mathbb{Q}$. K has been standardized to $K = 440$ Hz in 1939. Hence, each fundamental frequency can be represented as $f = 440 \text{ Hz} \cdot 2^o \cdot 3^q \cdot 5^t$. Here, 2, 3, and 5 are arbitrary pairwise coprime integers (See Mazzola, 1985, p. 26) but this choice relates to the numerators of the fundamental frequency ratios of the music theoretical basic intervals of two frequencies f and g in just intonation:

- octave: $f : g = 2 : 1$
- fifth: $f : g = 3 : 2$
- major third: $f : g = 5 : 4$

The fact that o, q, t are $\in \mathbb{Q}$ allows for unique solution of the equations **:raw-latex:`\citep[A1.1.2]{Mazzola1985,Mazzola2018}`**.

6.2 Pitch space

Pitch (or *pitch height*) is the perceptual correlate to a fundamental frequency and is measured on a linear scale. It can be calculated as $p(f) = 69.0 + 12 \cdot \log\_2(f/440\,\text{Hz})$, and 12 is the size of the octave in the unit of p. The formula thus gives the distance from the reference pitch (69=A4) as a proportion ($\log_2(f/g)$) of the octave (12). In this representation middle C (C4) gets the pitch number 60.

$$
\begin{aligned}
p(f) &= c_1 + c_2 \cdot \log_2\left(\frac{440\text{ Hz} \cdot 2^o \cdot 3^q \cdot 5^t}{440\text{ Hz}}\right) \\
&= c_1 + c_2 \cdot \log_2\left(2^o \cdot 3^q \cdot 5^t\right) \\
&= c_1 + c_2 \cdot (o \cdot \log_2(2) + q \cdot \log_2(3) + t \cdot \log_2(5))...
\end{aligned}
$$

Equivalently,

$$p(f) - c_1 = c_2 \cdot (o \cdot \log_2(2) + q \cdot \log_2(3) + t \cdot \log_2(5))$$

Accordingly, each frequency f or equivalently each tone t can be conceived as a point $x = (p, s, r) \in \mathbb{Q}^3$. And vice versa, each point is associated with its fundamental frequency $f(x) = f(p, s, r) = 440\text{Hz} \cdot 2^p \cdot 3^s \cdot 5^r$. In this representation, $p = o \cdot \log_2(2) + q \cdot \log_2(3) + t \cdot \log_2(5)$ is a linear combination from the basis vectors over $\mathbb{Q}^3 = \mathbb{T}^3$ with coefficients o, q, t, and $\mathbb{T}^3$ is a module over the ring $\mathbb{Z}$ of integers. Another way to state this is that octave, fifth and major third are the basis vectors of tonal space.

6.3 Transformations of tonal spaces.

For a fixed (arbitrary) reference tone (e.g. chamber tone A4=440Hz), and for integers $o, q, t \in \mathbb{Z}$ we can represent all tones in just intonation. The integer lattice $\mathbb{Z}^3 \subseteq \mathbb{Q}^3$ corresponds to Euler's representation **:raw-latex:`\citet{Euler1739}`** and we will denote it with $\mathbb{T}^3$, the tonal space incorporating the three dimensions of the octave o, the pure fifth q and the pure major third t.

Usually, descriptions of musical space do not consider octaves and octave-related tones are subsumed into equivalence classes (chroma classes or pitch classes); $\pi_O : \mathbb{T}^3 \to \mathbb{T}^2$; $(o, q, t) \mapsto (q, t)$

Notes are abstract symbolic representations of tones. They can be modeled as pairs $n = (p, d) \in \mathbb{T}^3 \times \mathbb{R}$ where p encodes *pitch* and d *duration*. The pitch dimension can be represented in one of the tonal pitch spaces (TPS).

6.4 Generalized Interval Systems

- (GIS) maps a TPS representation to an appropriate mathematical space
- A *Generalized Interval System* (GIS) is an ordered triple $(S, (G, \circ), \text{int})$, where S is a musical space, $(G, \circ)$ is a group, and int is an interval function that maps $S \times S \to G$. Common instances for G are $(\mathbb{Z}, +)$ (suitable for the line of fifth) or $(\mathbb{Z}_{12}, +)$ (suitable for the circle of fifths).
- …. many names: Pitch Space, Tonal Space, Tonal Pitch Space, Music Space, Musical Space….

These models of tonal space (line of fifths, circle of fifths, tonnetz, torus) can serve as support for probability distributions. These in turn describe the generative process for tonal pieces.

For the scope of this dissertation, the bag-of-notes model is assumed. Meaning, that the grammar G responsible for the sequential arrangement of notes is factored out.

Problem: "Theorizing in the wrong space" (Wiggins, 2012: Music, Mind, and Mathematics)

Pitch space encompasses the pitches and their mutual relations, the intervals. Certain assumptions about pitches transform pitch space.

There are numerous theoretical models of pitch space.

- Euler space
- Tonnetz
- Line of fifths
- Circle of fifths
- Spiral Array
- MIDI

6.5 Models of pitch space

Pitches can be expressed as $2^x 3^y 5^z$ for $x, y, z \in \mathbb{Q}$ **:raw-latex:`\citep{LonguetHiggins1962a,LonguetHiggins1962b,Mazzola2018}`**. "Fundamental theorem of harmony" **:raw-latex:`\citep{LonguetHiggins1965}`**. Pitches thus form a 3-dimensional space, also called the Euler space that incorporates just intonation (pure integer ratios of frequencies). Distances between pitches in this space are called intervals. Music theorists consider a number of equivalence relations that transform the space. The most common equivalence relation is octave equivalence that identifies all pitches that are related by frequency ratios of 2, effectively projecting the 3-dimensional Euler space to the plane given by $3^y 5^z$. This plane is commonly called the Tonnetz and has numerous historical precursors in the 19th century.

Since the Tonnetz expresses just intonation, one can distinguish, for instance, between the just third E above C and the Pythagorean third E' that lies four fifths. The difference between the just and the Pythagorean third is called the syntonic comma,

$$\begin{aligned} 4/5 : (2/3)^4 &= 4/5 : 16/81 \\ &= 4/5 \cdot 81/16 \\ &= 81/80 \\ \approx 1.0125. \end{aligned}$$

Identifying just and Pythoagorean thirds wraps the Tonnetz to a cylinder, also called the Spiral Array **:raw-latex:`\citep{Chew2000, Chew2014}`**.

On this cylinder, the line of fifths wraps around in such a way that every fourth fifth coincides with a third. This also means that all points on this cylinder lie on this line of fifths. The pitches in this space are sometimes called tonal pitch classes **:raw-latex:`\citep[TPCs][]{Temperley2000,Temperley2001}`**. The line of fifths is sufficient to capture all TPCs but the 2-dimensional surface of the cylinder emphasizes triadic relations. Moreover, a segment of six fifths contains all notes of a major or a (natural) minor scale and hence all pitches and intervals in a key. The triads within a key form Moebius strip **:raw-latex:`\citep{Mazzola1985,Noll2016}`**. Closing this segment to a circle involves only diatonic fifths, but one of them is diminshed (B–F in C major). This pitch set can be mapped to $\mathbb{Z}_7$.

[FIGURE: Diatonic chord sequence C-a-F-d-bo-G-e-C]

[FIGURE: Moebius strip embedded in $\mathbb{Z}_7$]

Finally, another importent equivalence relation is that of enharmonic equivalence. Enharmonic equivalence identifies octaves and augmented sevenths,

$$(1/2)^7 : (2/3)^1 2 \approx 1.0136.$$

This equivalence relations transforms the cylindrical pitch space to a torus, and the line wrapped around the cylinder to a circle, the circle of fifths. The tonal pitch classes are transformed into neutral pitch classes, or simply pitch classes.

The pitch classes on the circle of fifths can be reorderd to the chromatic circle by

$$p \mapsto 7p \mod 12,$$

resulting in the order on which the keys within one octave occur in the piano. Both the chromatic circle and the circle of fifths can be identified with $\mathbb{Z}_{12}$.

Especiall, review **:raw-latex:`\citep[115-136]{Temperley2001}`**, line of fifths, center of gravity and describe it in the language of distributions.

- distinction tonal/spelled vs. (neutral) pcs
- explain mapping to $\mathbb{Z}_{12}$
- sharp tpcs are mapped to positive numbers, flat tpcs to negative numbers
- already by this definition, the white-note diatonic is more sharp than flat (and not balanced!)

6.5.1 The bag-of-notes model

The bag-of-notes model conceives pieces simply counts the occurrences of notes without taking into account the order in which they appear in the piece.

It is in this sense a much more general model than the theoretically motivated ones that we have seen in the previous section. This model does not make any specific assumptions about the relations between notes other than that their respective frequency is relevant.

In the terminology of probability theory, relative note frequencies derived from note counts under the bag-of-words model correspond to multinomial distributions.

Take for example the first movement of Alkan's *Concerto for Solo Piano*, op. 39, No. 8, in G$\sharp$ minor. Figure [fig:Alkan_39-8_freqs] shows the note counts, weighted by duration (CHECK) and ranked by frequency.

One can see an almost linear relation between frequency and rank.

Only towards the end of the movement the key signature changes from five sharps (G sharp minor) to four flats Ab major but even in the sharp parts of the movement the notated score changes to flats where convenient.

Since musical pieces can have very different lengths–some pieces last only a few minutes while others may last more than an hour–it is useful to normalize the note counts and to derive the relative frequencies.

Interpret as distributions, show pitchplots… compare to Figure [fig:tonal_spaces].

These and many more conceivable transformations of tonal space do not serve the goal of merely reflecting abstract algebraic or geometric relations. It is important to emphasize that these transformations reflect rather practical assumptions for performance and instrument construction (such as dealing with the syntonic comma for keyboard instuments) or compositional decisions (such as enharmonic equivalence).

6.6 Dataset

MusicXML files from diverse sources… - musescore.com
- ELVIS
- Humdrum
- DCML transcriptions
- CPDL
- other websites…

In this chapter we look at the historical development of tonality. Although the dataset contains ca. 2000 pieces, there are unfortunately huge gaps in the timeline as can be seen in Figure [fig:piece_dist]. Attributing one year to a piece is not easy, in particular for older pieces. If available, we use the year of composition, otherwise the year of publication. Where both dates were unavailable, the middle year of the composer's life was chosen to represent the piece. Following this procedure leads to only 157 years for which we have pieces in the whole range of 582 years from 1361 to 1942.

If the hypothesis is true that tonality is constituted by the pitch usage in pieces and that certain compositional assumptions transform pitch space, then it should be possible to discover aspects of these assumptions and the structure of pitch space by analyzing the usage of pitches in musical compositions.

Moreover, comparing different sets of pieces, e.g. from different time periods or composers, should reveal historical and stylistic differences.

CHAPTER

SEVEN

TONAL PITCH-CLASS DISTRIBUTIONS

The tonal pitch-class distribution of a musical piece is the relative frequency of each tonal pitch class in that piece. Each piece can thus be represented as a V-dimensional vector, where V is the number of different pitch classes in the corpus, and that sums to one. In this view, pieces are points in a V-dimensional vector space.

In this space, pieces that have similar tonal pitch-class distributions will be close together whereas pieces with very different tonal pitch-class distributions will be more distant.

If all pieces are transposed to the same root, clusters in this space correspond to different types of distributions that can be interpreted as modes (take root out). This fact has been used in **:raw-latex:`\citep{Harasim2019}`** and also shown that there are historical developments.

If one does not transpose pieces, pieces that have similar root and mode (and, accordingly, similar distributions) should cluster together. Since V is usually quite large, it is difficult to visualize these clusters. One can use methods for dimensionality reduction to represent the data in lower-dimensional spaces (2D or 3D) in order to visualize them while at the same time maintaining characteristic properties of the original space.

One of the most popular and classic methods is **principal component analysis** **:raw-latex:`\citep{Bishop2006}`**, that can be used to project the data onto a two-dimensional plane while keeping as much of the variance in the data as possible. A more recent method for dimensionality is called **:math:`t`-distributed stochastic neighbor embedding** **:raw-latex:`\citep[$t$-SNE;][]{VanDerMaaten2008}`**. PCA is better to get a global understanding of the structure of the space and t-SNE is better in illustrating local relationships. Figure [fig:tsne_pca] shows the data reduced to the Euclidean plane by both methods.

The reduction using t-SNE (top panel) shows that there are many clusters that are relatively homogenuous with respect to their coloring. The PCA reduction on the bottom panel of Figure [fig:tsne_pca] also shows that pieces with similar coloring are close together but additionally shows that the colors are ordered along the line of fifths. This means that pieces in keys that are close on the line of fifths have similar tonal pitch-class distributions. Another advantage of PCA is that the axes, called the principal components, have clear interpretations. They reflect how much the data varies in this direction. Applying this interpretation to the right panel of Figure [fig:tsne_pca], one can see that the first principal component ("PC1") roughly represents the "distance to C" or "diatonic" pieces (white or very light colors) of more chromatic ones (darker shades). This distinction accounts for 55 percent of the variance in the data. The second principal component ("PC2") distinguishes sharp from flat keys (red vs. blue coloring) which is responsible for 21 percent of the data variance.

These two principal components together account for 76 percent of the variance in the data but simplify the space from $V = 35$ dimensions to just two which seems like a good tradeoff.

CHAPTER

EIGHT

HISTORICAL USAGE OF TONAL PITCH CLASSES

Apart from counting the number of tonal pitch classes in an individual musical piece, comparing these distributions between pieces and across historical time is interesting. In the last section we compared a small number of pieces manually. This section attempts at quantifying these intuitions and gaining a picture of the larger view.

The question is, how does the usage of tonal pitch classes change over time? Can we infer something about tonality from this change? An immediate caveat that comes to mind is that pieces often feature very different sets of notes because they are, for instance, in a different mode (both in the pre-tonal as well in the tonal sense), or key. It is therefor a standard preprocessing step in computational musicology to normalize pitch class distributions by transposing every piece to the same key in order to make them commensurate. For the same reasone, the chord symbols in the datasets analyzed in part [part:meso] where encoded with relative Roman numerals and not their absolute chord names. But in order to perform this normalization step, one needs to know the key of a piece. **(Well, not really: Harasim et al. 2019)** Moreover, the concept of "key" does not mean the same thing for all musical styles. Bach's B-minor Mass and Liszt's B-minor Sonata share the same nominal key but differ greatly with respect to their pitch-class distributions. Since the underlying tonality has changed, the derivative concept of key has changed, too. And just identifying B as the most common note in both pieces as indicative for the key (**check if that is the case**) is not a solution either because this procedure would also identify Renaissance locrian pieces as having the same key without even having touched the problem of how to infer the mode.

We come back to this issue in later chapters (**WHERE?**). Maybe it is appropriate to inspect the absolute pitch distributions of pieces before delving into the issue of relative pitch classes. This is what this chapter is about.

8.1 Modeling tonal pitch-class evolution

Tonal pitch classes on the line of fifths can be mapped to integers $k \in \mathbb{Z}$. An interval $I = [a, b] \subseteq \mathbb{Z}$ is called a **line-of-fifths segment** and its length is $n = |b - a|, a < b$. The distribution of tonal pitch classes at time t (in a piece or in historical time) is modeled as a draw from a Dirichlet distribution:

$$X^{(t)} \sim \mathrm{Dir}(\alpha), \alpha \in \mathbb{R}^n.$$
$$Importantly, in this model, the dimensions of : math : `X^t`have no$$

inherent order. This means that the model knows nothing about the line of fifths anymore. The ordering of pcs along this line is just for convenience. The probability of the pitch class $k = i - a$ at time t is given by the $k + a$th component of the vector $X^{(t)}$, $p(k|t) = X^t_{k+a}$ The diachronic change of these distributions forms a process

$$\mathbf{X} = (X^{(1)}, \ldots, X^{(t)}, \ldots, X^{(T)}) \in \mathbb{R}^{n \times T},$$
$$such that : math : `\sum_i X_i^{(t)} = 1, \forall t`.$$

8.2 Variability in tonal pitch-class usage

We count the occurrence of tonal pitch classes in all pieces and trace the change between them across the historical timeline. Based on theoretical reasoning **:raw-latex:`\citep{Temperley2000,Gardonyi2002}`**, we have already seen in section [sec:bagofwords] that it seems to be the case that sorting pitch classes along the line of fifths reveals structural connections between the pitch classes. For that reason we plot the pitch classes along this axis and also use colors to encode this relation.

As Figure [fig:piece_dist] has shown, the dataset is not uniformly distributed over time. On one hand, there are some large gaps between periods, whereas on the other hand some years contain many pieces at the same time.

For years without data, we take the assumption that “nothing changes” and keep the values from the last where were data was available. For the years with many pieces, we add up the pitch class counts, so that they all contribute to the calculation.

…

A **rolling mean**, also called a moving average, is calculated over the whole historical range. It is common that sliding windows are centered. But because it makes more sense for historical data to only consider previous events because future events have no impact, the result of the sliding window takes into account all t previous years.

For a value x_t in year t, and window size s, the rolling mean $\bar{x}$ is defined as

$$\bar{x} = \frac{1}{s}\sum_{i=0}^{s-1} x_{t-i}.$$

This definition allows a scalable perspective on historical developments. Adjusting the windows size allows to all historical periods in the range of the historical frame under consideration. For instance, setting $s = 50$ will lead to a curve that at any point represents the average value of the last 50 years, if years are the unit of time.

This is done for the tonal pitch-class distributions of aggregated pieces and is shown in Figure [fig:evolution_tpcs]. It is a complex plot and we will discuss each part at a time.

The legend above the two subplots show the mapping of tonal pitch classes to colors. Since tpcs are isomorphic to $\mathcal{Z}$, as mentioned above, it is possible to map flat tpcs to negative numbers, shown as graded blue colors, and to map sharp tpcs to positive numbers, shown as graded red colors. The tpc C is mapped to zero which corresponds to the color white in this plot.

The plot immediately below the legend shows the smoothed distribution of tonal pitch classes over time, sorted by the associated colors. The two dashed curves demarcate the white-note diatonic tonal pitch classes F to B. It is important to note here that in the bag-of-notes model tonal pitch classes are expressed as multinomial distributions. This means that there is no inherent order to the pitch classes–there is no structure in the bag. The coloring and sorting is done on theoretical grounds, but we will soon see that this ordering makes also sense for the data at hand.

The dark line throughout this plot shows the normalized entropy of the pitch-class distributions at any point in time. This line is smoothed by the same procedure as the individual per-year pitch-class distributions and is thus an adequate measure for the randomness of these distributions for a given year. Taking into account a 50-year window shows that randomness slightly increases over time with some wiggles along the way. The value of this line is independent of the number of tonal pitch classes in a given year, since it is normalized by its maximal value which is given by $\log(n)$ where n is the number of non-zero tonal pitch classes in that year.

The red line in the bottom plot in Figure [fig:evolution_tpcs] shows the ratio of non-zero “sharp” tpcs (G, D, A, …) to non-zero “flat” notes (F, B♭, E♭, etc.), defined as $q = s/f$, where s is the number of sharp tpcs (not unique but the actual number), and f is the number of flat notes. If $f = 0$, the ratio q is not defined. Since the analysis is based on the moving average, as well, a piece with no flats (which implies also F) is excluded. Since the window size is considerably large, there is no sliding window that contains only pieces with non-flat notes so that q is always defined as can be seen by the smoothness of the red line. As can be seen, this curve shows considerable variation. In both subplots, saddle points correspond to regions where no data is available so no interpretation should be given for these areas.

Bootstrap sample CIs!

If a musical piece exclusively contains the seven diatonic tpcs, and if they are furthermore uniformly distributed in this piece, the sharp-to-flat ratio is $q = |\{G, D, A, E, B\}/|F| = 5$. Which is exactly what we see in the beginning of our timeline.

The reverse statement that diatonic notes are uniformly distributed if the ratio is $q = 5$ is not necessarily true. In fact, there are non-diatonic notes present at the beginning of the timeline, namely B♭ in the flat direction, and F♯ and C♯ in the sharp direction. A uniform ratio would be then $q = |\{G, D, A, E, B, F\sharp, C\sharp\}/|B\flat, F| = 3.5$ So we can rule out uniformity, also because the entropy (the black line in the upper plot) is not maximal. The question is, whether the non-randomness in these distribution tells us something about tonality and its historical development. We come back to this question later.

The smoothed trends in both subplots show that sharpward tpcs are generally much more common if not only because all diatonic pitches are already sharps except F. More precisely, sharp notes occur roughly five times more often than flat notes until the last quarter of the fifteenth century. This might be due to the fact that almost only diatonic notes are being used, with relatively constant but low B♭s **(transposed modes!)**. On the sharp side of the spectrum, F♯s occur rarely, as do C♯s which lends itself to the interpretation that these notes do reflect the **musica ficta**. Other accidentals occur vanishingly seldom.

Around 1460 there is a decline in q that stabilizes around 1530 where the sharps occur only three times as often as flats. This is due to an increased use of flat notes F and B♭. Somewhat surprisingly, F♯ and higher sharps are absent in this period. But for modal music that is a logical consequence. If the transposed modes are used more often, sharp notes are less likely to occur.

In the second half of the 16th century, E♭ appears for the first time in the corpus in a substantial and stable way. But also F♯ comes back so it is counterbalanced and the ratio stays roughly the same.

Towards the end of the 16th century, we see a dramatic increase in the sharp-flat ratio that continues until the middle of the 17th century and reaches a more than 7-fold peak. This is due to the disappearance of almost all flats below B♭, while the sharps C♯, G♯, and D♯ become even stronger (and never vanish again). In this period, music seems to shift to the sharp side. While modal music featured the basid diatonic modes plus downward transposition to the flat side by one, here we see more and more accidentals.

…going into dominant regions means going sharpwards.

But this peak lasts only shortly. Around 1700 the sharp-flat ratio has fallen back to its earlier point around 2.5. But although the ratio is the same, the tpc usage is quite different. Now many more sharps and flats are employed than ever before. More importantly, this peak marks the beginning of the Baroque period. The first Baroque composer in the corpus is Corelli (also the most frequent one). There are a lot of pieces from him at the end of the 17th century.

A surge of flats around 1800 brings the ratio down to its lowest point since ca. 1530 and remains relatively stable throughout the 19th century. There is a slight rise and decay over the course of this century. Both sharps and flats increase in this time but more so do the flats.

In the early 20th century there is the third lowest point where flats dominate sharps ("renaissance of the Renaissance"? Vaughan Williams, Finzi, …)

CHAPTER

NINE

TONAL PITCH-CLASS COEVOLUTION

9.1 Modeling tonal pitch-class coevolution

The change/evolution of each pitch class $k = i - a$ is given by the changes in $\mathbf{X}_i = (X_i^1, \ldots, X_i^t, \ldots, X_i^T)^\top \in \mathbb{R}^T$. The pitch-class coevolution matrix is given by

$$\mathbf{\Sigma} = (\mathrm{corr}(\mathbf{X}_i, \mathbf{X}_j))_{ij} \in [-1, 1]^{n \times n}$$

$$and reflects the similarity of the diachronic change of pitch - classes.$$

These upper subplot in Figure [fig:evolution_tpcs] have shown the changes in the usage of each pitch-class over time. The coloring and ordering suggests indeed a coevolution but recall that the ordering was put in manually. The question is whether we can learn something about the structure from the data by analyzing the coevolution of the tpcs which is operationalized as the pairwise correlation (the Pearson correlation coefficient ρ) (maybe use sample coefficient r?) of two pc-evolution vectors p and q:

$$\rho_{p,q} = \frac{\mathrm{cov}(p, q)}{\sigma_p \sigma_q},$$

$$where : math : `\mathrm{cov}(p, q)` is the covariance and : math : `\sigma`$$

the standard deviation. Figure [fig:coevolution_tpcs] shows the pairwise tonal pitch class coevolution values across the entire timeline.

Interesting observations:

1. Three regimes are clearly separated: flats (upper left), diatonics (center), and sharps (lower right)
2. The chromatic regimes are of roughly the same size, (only visible in overall plot; the sharps are slightly larger), i.e. the heatmap has two orthogonal symmetry axes
3. Moreover, the chromatic notes (flats and sharps) are weakly positively correlated
4. F♭♭ (and more extreme flats) does not occur in the entire corpus
5. The weakest correlations are highly interesting as well: The weakest correlation is with the chromatic lower neighbor and the tritone (e.g. A vs. A♭, E♭; E vs E♭, B♭; B vs. B♭, F; F♯ vs. F, C) This is only true for "central" tpcs (white keys diatonic)

We can use this correlation matrix to plot distances between the pitch classes. Restricting the relations to the center of the plot, the diatonic notes plus F♯ and B♭ these distances actually approximate the line of fifths!

9.2 Deciphering pitch-class coevolution

The last section presented how strong the evolution of pitch classes correlates with each other. The heatmap in Figure [fig:coevolution_tpcs] indicated an interesting connection to the ordering of tpcs on the line of fifths. But this ordering was achieved manually, based on theoretical knowledge. How strong is this connection based on the available data?

One way to investigate this is to reduce the high-dimensional space to a smaller one. A common method to achieve this is **principal component analysis** (PCA). PCA analyzes the variance in the data and projects the data to a lower-dimensional space while maximizing the retained variance.

Subsequently, one can inspect the individual principal components individually and interpret the variance within and between them.

The results are very interesting:

1. Roughly, 64% of the variance is explained with the diatonic-chromatic distinction (PC1)
2. About 22% is explained by the sharp-flat distinction (PC2). Note also that C is on the zero-line for PC2 (does this really mean something?).
3. Another 6% of variance is explained by the third principal component. It roughly corresponds to the numbers of accidentals and follows, approximately, a zig-zag pattern for the 5 regions ♭♭, ♭, ”, ♯, and ♯♯.
4. PC4 is not easy to interpret, but it still captures a difference between, flat, diatonic, and sharp tpcs. Indeed, it seems that this component captures enharmonic equivalence! The tpcs C, G, D, A, was well as their enharmonic sharp and flat equivalents are all separated from the other notes. The same goes for F and E# (but not Gbb).

It seems that the PCA reduction was not only able to capture meaningful dimensions, but also a meaningful relation between them, namely the hierarchical one depicted in Figure [fig:pca_hierarchy].

The variance explained by each of the components can be interpreted as the weight or importance of these dimensions for the data. The two most important principal components are PC1 and PC2, together contributing approximately 86% of variance to the data. Figure [fig:PCA_2dim] shows how these two dimensions interact. Largely speaking, diatonic and chromatic tpcs can be separated by a vertical line (not exactly), whereas sharpward and flatward tpcs can be separated by a horizontal line, with C, the only tpc that is neither flatwards nor sharpwards, being exactly on the axis. Moreover, the three respectively most extreme tpcs, Fbb–Gbb and Ax–Bx, are located close to the origin of the PCA transformed plot. This means that they do not contribute much to the variance in the data. These are also precisely the ones outside of the enharmonic equivalence shown in PC4.

9.3 TPC coevolution per historical period

This “global view” can be broken down to compare how the tpc correlations change over time. The next figure shows the correlations for 50-year periods

1. 1500-1550: Two clusters emerge
2. 1550-1600: Clear separation between recta and ficta.
3. 1600-1650: Dahlhaus situates the origin of harmonic tonality in the early 17th century (Untersuchungen, p. 14), namely (following Fétis) in Monteverdi’s Cruda Amarilli SV 94, mm. 9-19, 24-30. Without diminishing Monteverdi’s influence we can see here that the first half of the 17 century was indeed a time of change, at least with respect to the conjunct usage of tones. But note also that the most prominent composer in that epoch in the dataset is Gesualdo who is well-known for his unusual harmonies.
4. 1650-1700: Confusion
5. 1650-1800 The separation between flat, diatonic, and sharp tpcs stabilizes. This is the closest to the overall picture above (although not as centered). The closest distribution to the overall distribution (check!) is the one in the late 18th century. It coincides with the common-practice period. Since we see that tpc behavior is different

before and after, the CPT should not be taken as a synonym to tonal music. This affects large portions of empirical research of tonality presupposing two modes with clear and stable patterns. Review also Harasim et al. (2019)

6. 1800-1900: Strong correlation between all accidentals vanishes. The diagonal line is very clear. In this time, all pitch classes exhibit the greatest independence historically speaking.
7. 1900-…: Looks like a mix of CPT and Extended

Dahlhaus situates the origin of harmonic tonality in the early 17th century (Untersuchungen, p. 14), namely (following Fétis) in Monteverdi's *Cruda Amarilli* SV 94, mm. 9-19, 24-30. Without diminishing Monteverdi's influence we can see here that the first half of the 17 century was indeed a time of change, at least with respect to the conjunct usage of tones. But note also that the most prominent composer in that epoch in the dataset is Gesualdo who is well-known for his unusual harmonies.

Turn argument around: Use inter-pc correlations to show importance of fifth structure! What about thirds?

CHAPTER TEN

DIATONICISM – CHROMATICISM – ENHARMONICISM

"When we think about harmony, we automatically think about chords. In fact, we are so fixated on chords that we sometimes forget they tell only part of the story" **:raw-latex:`\citep[154]{Tymoczko2011}`**

The development of tonality can also be described as a change in two dimensions: key-distance and separatedness (tonal closure/unity).

- Baroque: Keys are relatively close to each other but changes occur frequently, tonicizations are commonplace
- Classic: Keys are relatively close to each other and key sections are larger and relatively homogenuous
- early Romantic: Keys are further apart and key sections are larger and relatively homogenuous
- late Romantic: Keys are further apart but changes occur frequently

Here a tabular overview of this hypothesis:

	small	large	
	small	Baroque	Late Romantic
	large	Classic	Early Romantic

Table: Stages of Tonality.

10.1 Expansion of tonal material

Based on **:raw-latex:`\citep{Gardonyi2002}`**: (see MGG "Diatonik – Chromatik – Enharmonik")

- same diatonic region on LoF: relative keys/scales - although

theoretically, LoF is equivalent to $\mathbb{Z}$, composers use only a relatively small subset of it

- individual intervals can be associated with a regime on the

fifth-width space: m2 (5Q) is diatonic, wheras A1 (7Q) is chromatic, and A7≈P8 (12Q).e

- compare with "pitch class circulation" **:raw-latex:`\citep[158ff.]{Tymoczko2011}`**
- fifth width measures Diatonicism -> Chromaticism -> Enharmonicism **:raw-latex:`\citep[243]{Gardonyi2002}`**

|image|

- Analyze also the variance of fifth-widths, not only the means!

How can enharmonic exchange (Verwechslung) and enharmonic equivalence (Umdeutung) distinguished? The former implies a reinterpretation of tonal pitch classes, i.e. a transition to a different location in tonal space, whereas the former is only motivated by notational constraints (parsimony) and tonal/diatonic relations remain constant.

For example, in Debussy's *Claire de lune*. It is in D♭ major with a middle segment in C♯ minor which is enharmonically equivalent to D♭ minor but only has four sharps instead of eight flats.

First, enharmonic equivalence should only be invoked to render notation easier, not more difficult. This means, that the number of accidentals has to be reduced by the transformation.

Second, the key in question should be in a direct relation with the preceding and/or consequent key. In the case of the reinterpretation of the German sixth chord as a dominant chord effects a key shift by a semiton, which is far away in tonal space (LoF). In the Debussy example, the keys are only R related after applying the equivalence.

10.2 Expansion of local harmonic content

Fifth width per measure in a piece.

A couple of examples

10.2.1 Over time

The change in fifth widths is differently on a global (piece) and a local (measure) scale. Globally, pieces cross the boundary to chromaticism quite early (which can already happen with ficta), and even to enharmonicism (because modulations to distant keys takes place). At the same time looking on a local harmonic scale we see that chromatic ("dissonant") harmonies are rare on average (mode, mean, median) but are increasing historically (with an interesting wavelike pattern - what does it mean?). Locally, pieces do not cross the enharmonic threshold (on average)

10.3 Enharmonic spectrum

In the extreme case, for each note, a random tonal representative of the neutral pitch class is sampled uniformly. => unpredictable because infinite possibilities.

In practice, only a few representatives are likely candidates: not uniform prior on representatives but concentrated (has *a lot* to do with surrounding notes–context–but this is not possible on the bag-of-notes model).

Anyway, *if* absolute enharmonicism would prevail, the prior on the representatives would be flat (but this does not even happen in 12-tone music, show some examples). In "moderate" enharmonicism, some candidates would be preferred.

I can measure hoe many representatives of a pitch class occur in a piece.

=> **enharmonic pitch-class entropy** is a measure of enharmonicism (works obviously only with spelled pitch classes)

But: Even in a Bach piece (or older), e.g. F♯ and G♭ can co-occur. Because they occur in different contexts (different keys/tonal centers), they are **not** enharmonically equivalent. In the bag-of-notes model we need to factor in the fact of how likely it is to belong to a tonal center:

Which I already can estimate because of the mixture/topic model!

Thus, **enharmonicism** can be operationalized as the pitch-class entropy, weighted by the likelihood to belong to different tonalities/clusters/keys/tone fields.

=> maybe inverse weight, because: higher weights of F# and Gb in a Bach piece should trigger a new mixture component whereas in an extended tonal piece it might just adjust the parameters of existing components (variance)

[By the way, enharmonic distance is 12n (in fifths)]

But maybe also: actually inverse because in tonal music, enharmonic notes are outliers whereas in enharmonic music they get a lot of probability mass.

But then: How to distinguish enharmonically equivalent tonal regions from random enharmonicity?

—> entropy might help

If entropy is low, they should be outliers. If it is high, enharmonicism can be assumed.

Entropy is **highest** when all representatives are equally likely (ideally, 12-tone music).

Thus: the higher the enharmonic pitch-class entropy the higher is enharmonicism.

Hypothesis: EPCE increases over time (and is maximal with 12-tone compositions)

Segmentation can be either achieved in a fixed manner by bars, groups of bars, or segmentation sign posts such as key-signature changes or double bars.

A data-driven segmentation could use the notes themselves.

Then, segment length l would be informative about tonal stability / rate of tonality changes

- given the optimal number of clusters from the mixture model, apply key-scape algorithm. It should give rise to a much clearer segmentation with $K << 24$ components.
- Again, use Information Theory to determine best segmentation. (lowest entropy, per segment?, KL divergence with component distributions?)
- How can these differences be measured? (not always binary)
- Horizontally: Streams (*Auditory Scene Analysis*; **:raw-latex:`\citep{Bregman1990,Huron2016}`** Huron, 2016)
- Vertically: Building blocks / units –> Gestalt laws, Segmentation **:raw-latex:`\citep{Hanninen2012}`**

Laws can govern *primary parameters* which allow for syntactic relations between discrete units (such as pitch, or rhythm), and *secondary parameters* which discribe continous dimensions such as timbre, dynamics, etc.

10.4 Tonal Centers

10.4.1 Number of Tonal Centers

10.4.2 Distance of Tonal Centers

10.4.3 Divergence on the Tonnetz

CHAPTER ELEVEN

DIACHRONIC AND SYNCHRONIC TONAL STYLES

11.1 Clustering

11.2 Principal Component Analysis

11.3 Analysis of Styles

CHAPTER

TWELVE

TOPIC MODELING WITH LATENT DIRICHLET ALLOCATION (LDA)

Note: Rework this chapter based on the pedagogical introduction in [24].

Topic Models: What are corpora, documents, topics? "Distributional hypothesis" (Harris, 1954; Firth, 1957).

In general, topic models describe the generative process of how documents (viewed as bags of words) have been created. A document is defined as a distribution over topics and a topic is defined as a distribution over words. To generate a new document, one first chooses a distribution over topics, and for each word in the document choose a topic from this distribution. The word is then sampled from the distribution over words of this topic.

> A generative model for documents is based on simple probabilistic sampling rules that describe how words in documents might be generated on the basis of latent (random) variables. When fitting a generative model, the goal is to find the best set of latent variables that can explain the observed data (i.e., observed words in documents), assuming that the model actually generated the data. (Steyvers & Griffiths, 2007)

> In probability theory, the multinomial distribution is a generalization of the binomial distribution. For example, it models the probability of counts for rolling a k-sided die n times. [...] When k is 2 and n is 1, the multinomial distribution is the Bernoulli distribution. When k is 2 and n is more than 1, it is the binomial distribution. When n is 1, it is the categorical distribution.
>
> (Wikipedia)

12.1 Background

- LDA in general (short review of relevant papers), numerous extensions of the basic model
- application to music, review model of **:raw-latex:`\citet{Hu2009}`**.

A **corpus** $\mathcal{C}$ is a set of M pieces. For each piece, the **distribution of topics** θ is drawn from a Dirichlet distribution with fixed corpus parameter α.

A collection (multiset) of **notes** $\boldsymbol{u}_n = \{u_{n1}, \ldots, u_{nL}\}$ defines a **segment**. The number of unique notes in a corpus is the **vocabulary size** V. Each segment u_n (e.g. beat, slice, bar, section, ...) is assigned a unique **topic label** z (key, tonality, mode, ...). A **piece** $\mathcal{P} = \{\boldsymbol{u}_1, \ldots, u_N\}$ consists of N segments with associated topic labels. A piece can have at most N topics, if $N \leq V$, otherwise at most V topics. **Topics** β are defined as distributions over notes. Since there are K topics and V distinct notes, β can be represented as a $V \times K$ matrix where β_{ij} encodes the probability of note i in topic j.

12.1.1 Definitions and Assumptions

1. A *note* :math:` u in mathbb Z_{12}`.
2. A *segment* $\mathbf{u}_n = \{u_{n1}, \ldots, u_{nL}\}$. In a bag-of-notes (BoN) model, a segment can also be represented by a 12-dimensional count vector x_n, where x_n^j counts the number of times note j occurs.
3. A *piece* s is a sequence of N segments: $s = \{\mathbf{u}_1, ..., \mathbf{u}_N\}$. Again, in a BoN model a piece can be represented as a sequence of count vectors $X = (x_1, ..., x_N)$.
4. A *corpus* is a collection of M pieces, $\mathcal{S} = \{s_1, ..., s_M\}$.
5. Finally, a *topic* z is a probability distribution over the 12 pitch classes. In their model, a topic models the concept of key and each segment is assumed to have precisely one topic/key. Thus, the sequence of topics in a given piece is modeled as $\mathbf{z} = (z_1, ..., z_N)$.
6. They fix the number of topics to $K = 24$, based on prior music theory knowledge.

12.2 A generative model for a musical piece

Bag of notes model… multinomial distribution… no order/structure among the classes (tpcs)

- Formalization of LDA as a probabilistic graphical model (PGM)
- PGMs are generative models. Toy example to generate pieces.

1. For each piece s_m, $m = 1, ..., M$, draw a K-dimensional topic weight vector θ from a Dirichlet distribution $(\theta \sim \mathrm{Dir}(\alpha))$ to determine which keys are likely to occur:

   $$p(\theta \mid \alpha) = \frac{\Gamma\left(\sum_i \alpha_i\right)}{\prod_i \Gamma\left(\alpha_i\right)} \prod_i \theta^{\alpha_i - 1}.$$

   The corpus-level parameter α determines which topics are likely to co-occur in pieces.

2. For each segment $\mathbf{u}_n$, $n = 1, ...N$, in the piece, choose topic $z_n \in \{1, ..., K\}$ from the multinomial distribution $p(z_n = k \mid \theta) = \theta_k$.
3. For each note u_{nl} in $\approxeq_n$, $l = 1, ..., L$, choose a pitch-class from the multinomial distribution $p(u_{nl} = i \mid z_n = k, \beta) = \beta_{ij}$, where β is a $V \times K$ matrix encoding each topic as a distribution over $V = 12$ pitch classes.

This generative process defines a joint probability distribution over observed and latent random variables for each piece in the corpus:

$$p(\theta, \mathbf{z}, s \mid \alpha, \beta) = p(\theta \mid \alpha) \prod_{n=1}^{N} p(z_n \mid \theta) \prod_{l=1}^{L} p(u_{nl} \mid z_n, \beta).$$

In this model, a piece is a bag-of-segments, and segments are bags-of-notes.

12.2.1 Inference and Learning

The model is fully specified by the corpus-level Dirichlet parameter α and the key-profile matrix β. Under the assumption that they are known, key-profiles for segments or pieces can be inferred by computing the posterior distribution

$$p(\theta, \mathbf{z} \mid \alpha, \beta, s) = \frac{p(\theta, \mathbf{z}, s \mid \alpha, \beta)}{p(s \mid \alpha, \beta)},$$
$$according to Bayes' rule.$$

The denominator in the last equation is called the *marginal distribution* or *likelihood* of a piece. The learning problem for the present setting is to maximize the log-likelihood of all pieces in the corpus ("Which combination of α and β make it most likely that these pieces were generated?"). Thus, we want to maximize

$$\mathcal{L}(\alpha, \beta) = \int d\theta p(\theta \mid \alpha) \prod_{n=1}^{N} \sum_{z_n=1}^{K} p(z_n \mid \theta) \prod_{l=1}^{L} p(u_{nl} \mid z_k, \beta).$$

The simplest learning algorithm for this task is the expectation maximization (EM) algorithm. Since this is not tractable, it has to be approximated. They use variational approximation. I use Gibbs sampling. Gibbs sampling can be understood as a generalization of the EM algorithm. Instead of maximizing at each of its two steps (E and M), Gibbs sampling uses the conditional distributions and samples from them.

- Learn model parameters from corpus given a number K of topics via Gibbs sampling.
- Train-test split not self-evident. Possibilities:
  - Train on whole corpus
  - Lern topics for different periods separately

12.3 Extension of the LDA music model

Describe potential adaptions of the LDA model. In particular the difference to the pitch class representation and the interpretation of topics as tone fields, not keys

- Model notes in $\mathbb{Z}$ instead of $\mathbb{Z}_{12}$ (line of fifths instead of circle of fifths).
- Use different segmentations:
  - Slices (onsets)
  - Beats
  - Bars
  - Key-regions (as defined by accidentals)
  - entire piece
- Allow for more topics. Hypothesis: chromatic passages, hexatonic, octatonic, pentatonic, and variants of several keys will show up as topics.
- Take note-order into account: Griffiths, Steyvers, Blei & Tenenbaum (2005), and Andrews & Vigliocco (2010)
- Dynamic Topic Modeling - Changes of topics over time: Blei & Lafferty (2006)

Other Features / Random Variables

- Length in notes
- Pitch-class distribution in $\mathbb{Z}$ and in $\mathbb{Z}_{12}$
- Number of key changes
- Chromaticism
- Meter / Meter change

General Notes:

- Interpretation: Because the musical vocabulary is quite small when notes are the equivalent of words, it is not sufficient to just look at the most frequent notes in a topic in order to interpret it but rather to inspect the whole distribution over notes.

- Similarity between documents (pieces) and subcorpora –> Clustering

12.4 Topics inferred from the corpus

Figure [fig:topic_dists] shows the note distributions for the $K = 7$ inferred topics.

The relative weights of the topics in the overall corpus can be seen in Figure [fig:topics_weights]. The most common topic is topic 0 ("(transposed) diatonic") and the least common is topic 4 ("far-flats")

Besides calculating the overall importance of topics in the corpus, one can also look at the relative topic weights within individual pieces. Figure [fig:raw_topics] shows this distribution. Analogous to the pitch-class evolution from Figure [], pieces that are assigned to the same year are accumulated (**Describe procedure!**).

It is obvious how the lack of data in earlier periods affects the pattern we see. Nonetheless, it can be seen, that earlier pieces rarely contain the certain topics which only occur later.

12.5 Number of topics reveals hierarchy of tonality

"Vertical": Different values of $K \in [2, 12]$ indicate hierarchical nature of tonality.

1. Compare topic distributions for different values of K
2. Relate topics from different K-stages with each other: coarse to fine, correlations between some topics should increase

All matrices where based on documents. In the classical LDA setting, a corpus is a bag of documents. We are in particular interested in historical developments, so the chronological order is important. Moreover, we do not a piece for each year and for some years we have many pieces. The first step is to re-assign each piece its "display year" (composition, publication, of composer half-life). Then we average all pieces in the same year. We now have at most one topic distribution per year in the corpus.

…

But there are still years for which we do not have data, in particular in the earlier periods. Pragmatically, if we do not have a topic distribution for a given year, we take the one from the previous year. To that end, we create a time index ranging from the earliest to the latest date in the corpus.

We then iterate over all years and use the inferred topic distributions if there is one for that particular year. If not, we use the same as in the year before.

12.6 Historical inferences for each value of K

12.6.1 Sliding windows reveal trends

Figure [fig:raw_topics] takes a very fine-grained view on the evolution of the topic distribution because each year is a single data point. In order to see larger trends, we can zoom out and look at smoothed versions of the same data. We inspect rolling averages with a window size of 30, 50, and a hundred years to see generational, epochal and secular trends.

In these figures we can observe a) the relative topic importances (weights) over time, and b) identify breaking points and local extrema.

Moreover, the entropy of these distributions is informative!

Todo: PCA shows relation between topics and documents in (reduced) note-space

12.7 Topic coevolution

topic correlations motivate hierarchical clustering

12.8 Pitch class – topic coevolution

Overall Figure [fig:tpc_topic_coevolution] almost looks like a block matrix.

Locally, the fifths order is preserved, especially the diatonic, as seen by the row clusters! Regarding the topics, we see two major clusters. The left most one is "chromatic notes" and the right one is "diatonic plus". The diatonic cluster contains the diatonic notes without F, which get clustered with F but without B (note also that when B is included, Bb is most negatively correlated, and, when F is included, F# is most negatively correlated). The tritone is the condition to separate these as well as the chromatic semitone. Then it gets extended by sharps up to G# which makes sense because of the dominant of a minor. The last topic to join this cluster extends it into the flat direction. We have already noted that tonal music is generally sharpwards oriented so it makes sense that the evolution of flat notes is weaker correlated with diatonic notes than sharp ones.

The second cluster…

12.8.1 Beyond the bag-of-notes model: the Hidden Markov Topic Model (HMTM)

Improving the bag-of-notes model with a Hidden Markov Topic Model

12.8.2 Discussion

- Result 1: Historically, ever larger portions of pitch space are

explored

- There is a trend from diatonic > chromatic > enharmonic pieces, but it

is not monotonic. In the 19th century, there are diatonic Lieder (composer) and Alkan, who has the greatest tonality range (diatonic-enharmonic).

CHAPTER

THIRTEEN

SEGMENTATION

13.1 Free segmentation

(pc set analysis)

- Straus [38]
- Hanninen [11], Hanninen [12]

relating segments creates specific graph structure

13.2 Contiguous segmentation

chord labels, Roman numerals, etc.

graph: chain

13.3 Hierarchical segmentation

CFGs, Schenker

graph: tree

13.4 Exhaustive segmentation

Keyscapes Sapp [35], Sapp [36], Pitchscapes Lieck and Rohrmeier [16], Wavescapes Viaccoz *et al.* [42]

See also Müller (form)

CHAPTER

FOURTEEN

ADVANCED APPLICATIONS

14.1 Pitch Spelling

Meredith [19], Meredith and Wiggins [20] Cambouropoulos [5], Chew and Chen [6] Stoddard *et al.* [37] Temperley [40]

14.2 Style (classification)

14.2.1 Feature clustering

(k-means, PCA, …)

14.2.2 Hierarchical clustering

14.3 History

(regression, GPs)

- trends (maybe with a non note-based dataset e.g. metadata)

14.4 Performance

- Spotify API to compare different recordings

14.5 Modeling musical sequences

14.5.1 Hierarchical theories

describe central Schenkerian concepts in terms of tones, intervals, and underlying tonal spaces. E.g., a neighbor note is the next note (upper or lower) in a tonal space that has a notion of neighborhood, e.g. the diatonic or chromatic spaces. But in this generalized sense, a neighbor can be a semitone, a whole tone, a third, or a fifth apart. What the neighbor actually is, depends on the underlying assumed tonal space. Accordingly, the *Bassbrechung* is an upper neighbor on the circle or line of fifths, while the common neighbor note only exists in diatonic spaces.

14.5.2 Formal models for music sequences

See overview in [32]

Regular Expressions

(chord symbols, rhythms)

n-gram models

(melody, rhythms)

Hidden Markov models

(harmony)

Probabilistic Context-Free Grammars

(form; choro) [22, 25]

More advanced models

Neural nets

CHAPTER

FIFTEEN

DEVELOPERS

15.1 API - gamuth

class gamuth.**Interval**(*source*, *target*)

Class for an interval between two tones *s* (source) and *t* (target).

get_euclidean_distance(*precision=2*)

Calculates the Euclidean distance between two tones with coordinates in Euler space.

Parameters **precision** (*int*) – Rounding precision.

Returns The Euclidean distance between two tones *s* (source) and *t* (target).

Return type float

Example

```
>>> s = Tone(0,0,0) # C_0
>>> t = Tone(1,2,1) # D'1
>>> i = Interval(s,t)
>>> i.get_euclidean_distance()
2.45
```

get_generic_interval(*directed=True*, *octaves=True*)

Generic interval (directed) between two tones.

Parameters

- **directed** (*bool*) – Affects whether the returned interval is directed or not.
- **octaves** (*bool*) – returns generic interval class if *False*.

Returns (Directed) generic interval from *s* to *t*.

Return type int

Example

```
>>> db = Tone(0,-1,-1) # Db,0
>>> b = Tone(0,1,1) # B'0
>>> i1 = Interval(db, b) # the interval between Db0 and B1 is an ascending␣
↪thirteenth
>>> i1.generic_interval()
13
```

```
>>> i2 = Interval(b, db) # the interval between B1 and Db0 is a descending␣
↪thirteenth
>>> i2.generic_interval()
-13
```

```
>>> i3 = Interval(b, db) # the interval between B1 and Db0 is a descending␣
↪thirteenth
>>> i3.generic_interval(directed=False)
13
```

get_specific_interval(*directed=True*, *octaves=True*)

Specific interval (directed) between two tones.

Parameters

- **directed** (*bool*) – Affects whether the returned interval is directed or not.
- **octaves** (*bool*) – returns specific interval class if *False*.

Returns (Directed) specific interval from *s* to *t*.

Return type int

Example

```
>>> fs = Tone(0,2,1) # F#'0
>>> db = Tone(0,-1,-1) # Db,0
>>> i1 = Interval(fs, db)
>>> i1.specific_interval()
17
```

```
>>> i1.specific_interval(octaves=False)
5
```

class gamuth.**PitchClass**

Pitch class instance in $\mathbb{Z}_{12}$.

class gamuth.**PitchClassSet**(*st*)

Pitch class sets # For multiple constructors see: https://pythonconquerstheuniverse.wordpress.com/2010/03/17/multiple-constructors-in-a-python-class/

interval_class_vector()

Interval-class vector for given pitch-class set

Returns interval-class vector

Return type list

`invert`(*t=0*)
Invert pitch-class set. If the inversion pc is not specified, it is set to 0 by default.

Parameters `t` (*int*) – inversion pitch class (default: 0)

Returns inverted pitch-class set

Return type set

`normal_form`()
Normal form of pitch-class set

Returns Normal form

Return type set

`transpose`(*t*)
Transposition by *t* semitones.

Parameters `t` (*int*) – number of semitones to transpose up

Returns transposed pitch-class set

Return type set

class `gamuth.Tone`(*octave=None*, *fifth=None*, *third=None*, *name=None*)
Class for tones.

`get_accidentals`()
Gets the accidentals of the tone (flats (♭) or sharps (♯))..

Parameters `None` –

Returns The accidentals of the tone.

Return type str

Example

```
>>> t = Tone(0,7,0) # C sharp
>>> t.get_accidentals()
`#`
```

`get_frequency`(*chamber_tone=440.0*, *precision=2*)
Get the frequency of the tone.

Parameters

- `chamber_tone` (*float*) – The frequency in Hz of the chamber tone. Default: 440.0 (A)
- `precision` (*int*) – Rounding precision.

Returns The frequency of the tone in Hertz (Hz).

Return type float

Example

```
>>> t = Tone(0,0,0)
>>> t.get_frequency(precision=3)
261.626
```

get_label()

Gets the complete label of the tone, consisting of its note name, syntonic position, and octave.

Parameters **None** –

Returns The accidentals of the tone.

Return type str

Example

```
>>> c = Tone(0,0,0)
>>> ab = Tone(0,1,-1)
>>> c.get_label(), ab.get_label()
`C_0` `Ab,1`
```

get_midi_pitch()

Get the MIDI pitch of the tone.

Parameters **None** –

Returns The MIDI pitch of the tone if it is in MIDI pitch range (0–128)

Return type int

Example

```
>>> t = Tone(0,0,0)
>>> t.get_midi_pitch()
60
```

get_pitch_class(*start=0*, *order='chromatic'*)

Get the pitch-class number on the circle of fifths or the chromatic circle.

Parameters

- **start** (*int*) – Pitch-class number that gets mapped to C (default: 0).
- **order** (*str*) – Return pitch-class number on the chromatic circle (default) or the circle of fifths.

Returns The pitch class of the tone on the circle of fifths or the chromatic circle.

Return type int

Example

```
>>> t = Tone(0,7,0) # C sharp
>>> t.get_pitch_class(order="chromatic")
1
```

```
>>> t = Tone(0,7,0) # C sharp
>>> t.get_pitch_class(order="fifths")
7
```

get_step()

Gets the diatonic letter name (C, D, E, F, G, A, or B) of the tone *without* accidentals.

Parameters **None** –

Returns The diatonic step of the tone.

Return type str

Example

```
>>> t = Tone(0,7,0) # C sharp
>>> t.get_step()
`C`
```

get_syntonic()

Gets the value of the syntonic level in Euler space. Tones on the same syntonic line as central C are marked with _, and those above or below this line with ' or ,, respectively.

Parameters **None** –

Returns The number of thirds above or below the central C.

Return type int

Example

```
>>> e1 = Tone(0,4,0) # Pythagorean major third above C
>>> e2 = Tone(0,0,1) # Just major third above C
>>> e3 = Tone(0,8,-1) # Just major third below G sharp
>>> e1.get_syntonic(), e2.get_syntonic(), e3.get_syntonic()
`'` `_` `,`
```

BIBLIOGRAPHY

[1] Edward Aldwell, Carl Schachter, and Allen Cadwallader. *Harmony and Voice Leading*. Cengage Learning, 4th edition, 2010.

[2] Hernandez Benavides Miguel Angel and Galpin Ixent. Exploring the musicalization of texts in Gregorian Chant using Data Analytics. Technical Report, Universitaria de Bogotá Jorge Tadeo Lozano, 2020.

[3] S. Brown and J. Jordania. Universals in the world's musics. *Psychology of Music*, 41(2):229–248, 2013. doi:10.1177/0305735611425896.

[4] Allen Cadwallader and David Gagné. *Analysis of Tonal Music. A Schenkerian Approach*. Oxford University Press, 1998.

[5] Emilios Cambouropoulos. Pitch Spelling: A Computational Model. *Music Perception: An Interdisciplinary Journal*, 20(4):411–429, June 2003. doi:10.1525/mp.2003.20.4.411.

[6] Elaine Chew and Yun-Ching Chen. Real-Time Pitch Spelling Using the Spiral Array. *Computer Music Journal*, 29(2):61–76, June 2005. doi:10.1162/0148926054094378.

[7] Bas Cornelissen, Willem Zuidema, and John Ashley Burgoyne. Mode classification and natural units in plainchant. In *Proceedings of the 21st International Society for Music Information Retrieval Conference*, 869–875. Montreal, Canada, October 2020. ISMIR. doi:10.5281/zenodo.4245572.

[8] Dimitrios S. Delviniotis. New Method of Byzantine Music (BM) Intervals' Measuring and Its Application in the Fourth Mode. \$A\$ New Approach of the Music Intervals' Definition. *Series Musicologica Balcanica*, 1(2):209–232, February 2021. doi:10.26262/smb.v1i2.7941.

[9] Vedad FamourZadeh. La musique persane, Formalisation algébrique des structures. Technical Report, IRCAM, 2005.

[10] Allen Forte. *The Structure of Atonal Music*. Yale University Press, New Haven and London, 1977.

[11] Dora A Hanninen. Orientations, Criteria, Segments: A General Theory of Segmentation for Music Analysis. *Journal of Music Theory*, 45(2):345–433, 2001.

[12] Dora A. Hanninen. *A Theory of Music Analysis: On Segmentation and Associative Organization*. University of Rochester Press, Rochester, NY, first edition, December 2012.

[13] Daniel Harasim, Fabian C. Moss, Matthias Ramirez, and Martin Rohrmeier. Exploring the foundations of tonality: statistical cognitive modeling of modes in the history of Western classical music. *Humanities and Social Sciences Communications*, 8(1):1–11, January 2021. doi:10.1057/s41599-020-00678-6.

[14] M Kemal Karaosmanoğlu. A Turkish Makam Music Symbolic Database for Music Information Retrieval: SymbTr. In *ISMIR*, 6. 2012.

[15] David Lewin. *Generalized Musical Intervals and Transformations*. Oxford University Press, Oxford, 1987.

[16] Robert Lieck and Martin Rohrmeier. Modelling Hierarchical Key Structure with Pitch Scapes. In *Proceedings of the 21st International Society for Music Information Retrieval Conference (ISMIR 2020)*, 811–818. Montreal, Canada, 2020. doi:10.5281/zenodo.4245558.

[17] Guerino Mazzola. *Geometrie Der Töne: Elemente Der Mathematischen Musiktheorie*. Birkhäuser, Basel, 1985.

[18] Samuel Mehr, Manvir Singh, Dean Knox, Daniel Ketter, Daniel Pickens-Jones, S. Atwood, Christopher Lucas, Nori Jacoby, Alena Egner, Erin J. Hopkins, Rhea M. Howard, Joshua Hartshorne, Mariela Jennings, Jan Simson, Constance Bainbridge, Steven Pinker, Timothy J. O'Donnell, Max Krasnow, and Luke Glowacki. Universality and diversity in human song. *Science*, 366:eaax0868, 2019. doi:10.31234/OSF.IO/EMQ8R.

[19] David Meredith. The ps13 pitch spelling algorithm. *Journal of New Music Research*, 35(2):121–159, June 2006. doi:10.1080/09298210600834961.

[20] David Meredith and Geraint A Wiggins. Comparing Pitch Spelling Algorithms. *Proceedings of the International Symposium on Music Information Retrieval (ISMIR)*, pages 280–287, 2005. doi:10.1007/978-3-540-31807-1_14.

[21] Fabian C Moss, Markus Neuwirth, and Martin Rohrmeier. The line of fifths and the co-evolution of tonal pitch classes. *Journal of Mathematics and Music*, accepted.

[22] Fabian C. Moss. Choro Songbook Corpus. Zenodo, 2020. doi:10.5281/zenodo.3881347.

[23] Fabian C. Moss, Markus Neuwirth, Daniel Harasim, and Martin Rohrmeier. Statistical characteristics of tonal harmony: a corpus study of Beethoven's string quartets. *PLoS ONE*, 14(6):e0217242, 2019. doi:10.1371/journal.pone.0217242.

[24] Fabian C. Moss and Martin Rohrmeier. Discovering Tonal Profiles with Latent Dirichlet Allocation. *Music & Science*, 4:20592043211048827, January 2021. doi:10.1177/20592043211048827.

[25] Fabian C. Moss, Willian Fernandes Souza, and Martin Rohrmeier. Harmony and form in Brazilian Choro: A corpus-driven approach to musical style analysis. *Journal of New Music Research*, 49(5):416–437, 2020. doi:10.1080/09298215.2020.1797109.

[26] Meinhard Müller. *Fundamentals of Music Processing*. Springer, 2015.

[27] Markus Neuwirth, Daniel Harasim, Fabian C. Moss, and Martin Rohrmeier. The Annotated Beethoven Corpus (ABC): A Dataset of Harmonic Analyses of All Beethoven String Quartets. *Frontiers in Digital Humanities*, 5(July):1–5, 2018. doi:10.3389/fdigh.2018.00016.

[28] Edward Nowacki. *Latin and Greek Music Theory*. Eastman Studies in Music. University of Rochester Press, Rochester, 2020.

[29] Thomas Nuttall, Miguel G. Casado, Andres Ferraro, Darrell Conklin, and Rafael Caro Repetto. A computational exploration of melodic patterns in Arab-Andalusian music. *Journal of Mathematics and Music*, 0(0):1–13, May 2021. doi:10.1080/17459737.2021.1917010.

[30] Richard Parncutt. Pitch-class prevalence in plainchant, scale-degree consonance, and the origin of the rising leading tone. *Journal of New Music Research*, 48(5):434–448, October 2019. doi:10.1080/09298215.2019.1642360.

[31] Harold S. Powers, Frans Wiering, James Porter, James Cowdery, Richard Widdess, Ruth Davis, Marc Perlman, Stephen Jones, and Allan Marett. Mode. In *Grove Music Online*. Oxford University Press, Oxford, 2001. doi:10.1093/gmo/9781561592630.article.43718.

[32] Martin Rohrmeier and Thore Graepel. Comparing Feature-Based Models of Harmony. In *Proceedings of the 9th International Symposium on Computer Music Modeling and Retrieval CMMR 2012*, 357–370. London, UK, 2012. Springer.

[33] Sebastian Rosenzweig, Frank Scherbaum, David Shugliashvili, Vlora Arifi-Müller, and Meinard Müller. Erkomaishvili Dataset: A Curated Corpus of Traditional Georgian Vocal Music for Computational Musicology. *Transactions of the International Society for Music Information Retrieval*, 3(1):31–41, 2020. doi:10.5334/tismir.44.

[34] Farshad Sanati. *An Investigation on the Value of Intervals in Persian Music*. Master Thesis, University of Jyväskylä, 2020.

[35] Craig Stuart Sapp. Harmonic Visualizations of Tonal Music. In *The International Computer Music Conference (ICMC)*, 423–430. 2001.

[36] Craig Stuart Sapp. Visual hierarchical key analysis. *Computers in Entertainment*, 3(4):3, 2005. doi:10.1145/1095534.1095544.

[37] Joshua Stoddard, Christopher Raphael, and Paul E Utgoff. Well-tempered spelling: A key-invariant pitch spelling algorithm. In *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR 2004)*. Barcelona, Spain, 2004.

[38] Joseph Nathan Straus. *Introduction to post-tonal theory*. Pearson Prentice Hall, New York, 3rd edition, 2005.

[39] David Temperley. The Line of Fifths. *Music Analysis*, 19(3):289–319, 2000. doi:10.2307/854457.

[40] David Temperley. *The Cognition of Basic Musical Structures*. MIT Press, 2001.

[41] Gary Tomlinson. *A Million Years of Music*. Princeton University Press, 2018.

[42] Cédric Viaccoz, Daniel Harasim, Fabian C. Moss, and Martin Rohrmeier. Wavescapes: A visual hierarchical analysis of tonality using the discrete Fourier transform. *Musicae Scientiae*, pages 10298649211034906, January 2022. doi:10.1177/10298649211034906.

[43] Robert W Wason and Elizabeth West Marvin. Riemann's "Ideen zu Einer 'Lehre von den Tonvorstellungen'": An Annotated Translation. *Journal of Music Theory*, 36(1):69–79, 1992.

[44] Frans Wiering. *The Language of the Modes: Studies in the History of Polyphonic Modality*. Routledge, New York and London, 2001.

[45] Jason Yust. *Organized Time: Rhythm, Tonality, and Form*. Oxford Studies in Music Theory. Oxford University Press, Oxford, New York, 2018.

PYTHON MODULE INDEX

g

INDEX