Discovering the line of fifths in a large historical corpus

Future Directions of Music Cognition Virtual Conference

Fabian C. Moss, Markus Neuwirth, Martin Rohrmeier 7 March 2021

Digital and Cognitive Musicology Lab École Polytechnique Fédérale de Lausanne fabian.moss@epfl.ch – @fabianmoss



Corpus Studies - Music Theory - Music Analysis

A possible corpus study scenario:

- 1. analyze a piece or some salient aspect
- compare it to other pieces and wonder whether observation holds more generally
- 3. gather/create suitable corpus to study question on a large scale
- 4. operationalize question for quantitative analysis
- 5. test against data
- 6. usually: new (general) observations
- 7. do they apply to individual pieces?
- 8. go back to 1.

Different realizations of hexachords



Figure 1: Josquin Desprez, Missa sine nomine, Agnus Dei II (1514).

Different realizations of hexachords



Figure 1: Josquin Desprez, Missa sine nomine, Agnus Dei II (1514).



Figure 2: Chick Corea, Children's Song No. 2 (1984), mm. 1–4.

The Line of Fifths



Figure 3: The Line of Fifths. Color mapping emphasizes flat- and sharpwards directions.

The Line of Fifths



Figure 3: The Line of Fifths. Color mapping emphasizes flat- and sharpwards directions.

The LOF contains many common scales as contiguous subsegments, e.g.

- diatonic
- pentatonic
- hexachord (Desprez, Corea)

• ...

Research question

If the line of fifths (and its subsegments) are central for tonal organization in many pieces, we should be able to see its relevance in a corpus study.

Research question

If the line of fifths (and its subsegments) are central for tonal organization in many pieces, we should be able to see its relevance in a corpus study.

Operationalization

- \cdot vector-space model: a piece is represented by the distribution of its notes
- representation of notes as tonal pitch-classes (e.g. F#, Gb, A##, ...)
- $\cdot\,$ important: no inherent ordering of pitch classes
- restriction fo Fbb to B## (35 dimensions)

Research question

If the line of fifths (and its subsegments) are central for tonal organization in many pieces, we should be able to see its relevance in a corpus study.

Operationalization

- \cdot vector-space model: a piece is represented by the distribution of its notes
- representation of notes as tonal pitch-classes (e.g. F#, Gb, A##, ...)
- important: no inherent ordering of pitch classes
- restriction fo Fbb to B## (35 dimensions)

Expectation: similar pieces (in terms of PC distribution) should lie close together in 35-D space.

The corpus

The corpus was gathered from various scholarly resources and surely includes certain biases (instrumentation, canonic composers...)



Figure 4: Convenience sample of 2012 pieces by 75 composers, 1361–1942 (Moss, Neuwirth, and Rohrmeier, 2020).

Method

Principal Components Analysis (PCA)

- We use PCA for dimensionality reduction.
- Principal components are dimensions that maximize variance in the data.
- Inspecting principle components thus allows to draw conclusions about which factors affect the data distribution most.

Method

Principal Components Analysis (PCA)

- We use PCA for dimensionality reduction.
- Principal components are dimensions that maximize variance in the data.
- Inspecting principle components thus allows to draw conclusions about which factors affect the data distribution most.

Procedure

- $\cdot\,$ corpus 'lives' in 35-D space, all dimensions are assumed to be independent
- \cdot we inspect the first two principal components, explaining 64% of variance
- \cdot each data point is colored according to its most frequent pitch class



Figure 5: PCA, global view.



Figure 5: PCA, global view.



Figure 6: PCA, first two components: direction (flatwards/sharpwards; top) and distance to center D (bottom).



Figure 7: Average tonal pitch-class distribution in the corpus.

Observations:

- almost normally distributed
- centered on D (mid-point on line of fifths)



Figure 7: Average tonal pitch-class distribution in the corpus.

Observations:

- almost normally distributed
- centered on D (mid-point on line of fifths)

This is a global statistic. What does it look like under a historical perspective?



Figure 8: Relative pitch-class frequencies over time (moving average).

- local variations possibly due to specific sample (corpus)
- amount pitch-classes with one or several accidentals increases
- entropy remains relatively stable
- relative frequencies appear to be correlated

 high correlations around the central axis



- high correlations around the central axis
- highest for P5/P4



- high correlations around the central axis
- highest for P5/P4
- lowest correlations for A4/A1 (blue)



- high correlations around the central axis
- highest for P5/P4
- lowest correlations for A4/A1 (blue)
- only for central region of line of fifths



- high correlations around the central axis
- highest for P5/P4
- lowest correlations for A4/A1 (blue)
- only for central region of line of fifths
- Pieces further away from the LOF center are generally more chromatic!





Figure 9: PC co-evolution in the 18th (left) and 19th (right) centuries.

Summary

• unsupervised recovery of the line of fifths from a historical corpus using dimensionality reduction

Summary

- unsupervised recovery of the line of fifths from a historical corpus using dimensionality reduction
- LOF central organization principle for pitch-class distributions

Summary

- unsupervised recovery of the line of fifths from a historical corpus using dimensionality reduction
- LOF central organization principle for pitch-class distributions
- stark historical contrasts between epochs

Summary

- unsupervised recovery of the line of fifths from a historical corpus using dimensionality reduction
- LOF central organization principle for pitch-class distributions
- stark historical contrasts between epochs

Summary

- unsupervised recovery of the line of fifths from a historical corpus using dimensionality reduction
- LOF central organization principle for pitch-class distributions
- stark historical contrasts between epochs

Future work

• expanding the corpus (also beyond canon)

Summary

- unsupervised recovery of the line of fifths from a historical corpus using dimensionality reduction
- LOF central organization principle for pitch-class distributions
- stark historical contrasts between epochs

Future work

- expanding the corpus (also beyond canon)
- taking other parameters into account (harmony, meter, form...)

Summary

- unsupervised recovery of the line of fifths from a historical corpus using dimensionality reduction
- LOF central organization principle for pitch-class distributions
- stark historical contrasts between epochs

Future work

- expanding the corpus (also beyond canon)
- taking other parameters into account (harmony, meter, form...)
- $\cdot\,$ zooming on from the global view

Thank you for your attention!

Slides available at fabian-moss.de

References

Moss, F. C., Neuwirth, M., & Rohrmeier, M. (2020). Tonal Pitch-Class Counts Corpus (TP3C) (Version v1.0.0). Zenodo. https://doi.org/10.5281/zenodo.3600088
Temperley, D. (2000). The line of fifths. *Music Analysis*, 19(3), 289–319. https://doi.org/10.2307/854457