Digital Musicology 2022
Tutorials

Assignment 3: Similarity

TA: Gabriele Cecchetti
gabriele.cecchetti@epfl.ch

Deliverables

- Due date: 04.05, 12h
- Deliverables:
  - **Code:** A Jupyter Notebook
  - **Report:** A short report (max. 2 pages) as a Word document or other text file
- Submission:
  - Store your data, code and report in your group’s private GitHub repository (shared with the team members and the TAs). Make sure your notebook is pushed with all output visible, i.e., in a form whereby we do not need to run the code.

The corpus

- Symbolic encoding (.mscz) of 366 chorales by J. S. Bach (1685-1750)
- Files are numbered according to the Riemenschneider catalogue. Inside each file, you also find the corresponding Bach Werke Verzeichnis (BWV) numbering as well as the incipit of the text
- 4-part vocal writing
  - one staff per voice, ordered by register from top to bottom: Soprano, Alto, Tenor, Bass
- “Chorale melody” in the upper voice
- Roughly homorhythmic texture
- Segmentation marked by fermatas

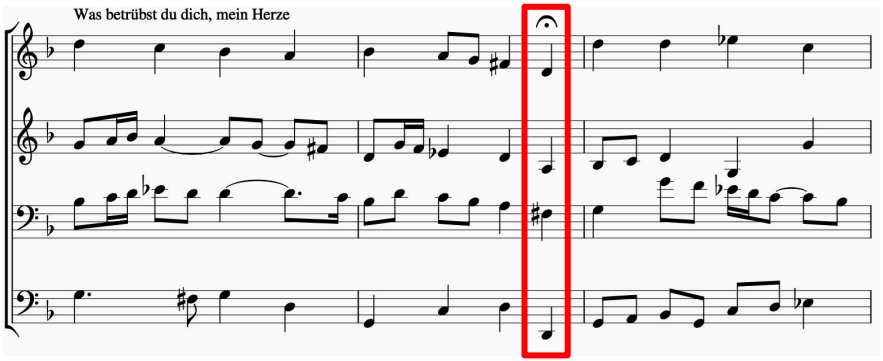
The dataset

- Scores are parsed into a .csv list of notes

- For the purpose of the task, only consider the incipits of the chorales:
  - If the first fermata occurs before the 4th bar, consider until the second fermata
  - If the first fermata occurs in or after the 4th bar, consider until the first fermata

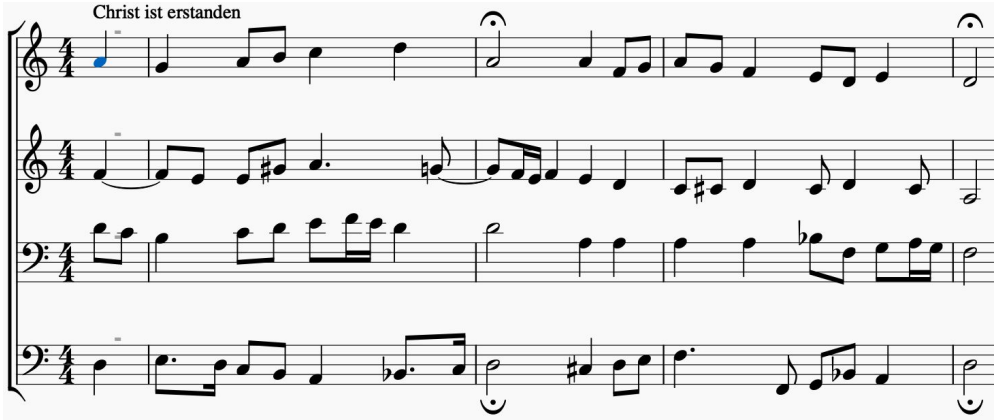
- The entire chorales are still available (e.g., for detecting statistical regularities in the style etc.)

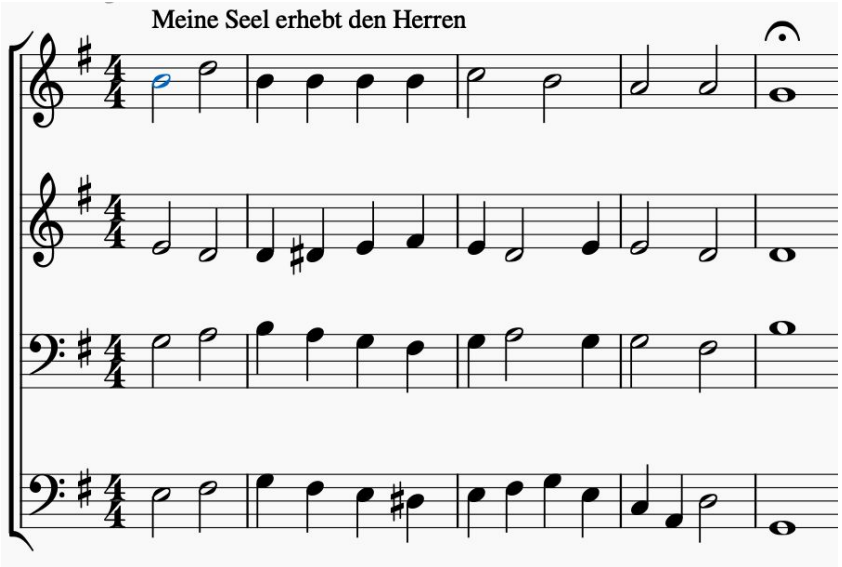
Fermatas



n	piece	mn	mn_onset	timesig	act_dur	staff	voice	duration	nominal_duration	scalar	tied	tpc	midi	gracenote	fermata
34	BachChorales/Chorale237	2	1/4	4/4	1	1	1	1/8	1/8	1	NaN	3	69	NaN	False
35	BachChorales/Chorale237	2	3/8	4/4	1	3	1	1/8	1/8	1	NaN	-2	58	NaN	False
36	BachChorales/Chorale237	2	3/8	4/4	1	1	1	1/8	1/8	1	NaN	1	67	NaN	False
37	BachChorales/Chorale237	2	1/2	4/4	1	4	1	1/4	1/4	1	NaN	2	50	NaN	False
38	BachChorales/Chorale237	2	1/2	4/4	1	3	1	1/4	1/4	1	NaN	3	57	NaN	False
39	BachChorales/Chorale237	2	1/2	4/4	1	2	1	1/4	1/4	1	NaN	2	62	NaN	False
40	BachChorales/Chorale237	2	1/2	4/4	1	1	1	1/4	1/4	1	NaN	6	66	NaN	False
41	BachChorales/Chorale237	2	3/4	4/4	1	4	1	1/4	1/4	1	NaN	2	38	NaN	True
42	BachChorales/Chorale237	2	3/4	4/4	1	3	1	1/4	1/4	1	NaN	6	54	NaN	True
43	BachChorales/Chorale237	2	3/4	4/4	1	2	1	1/4	1/4	1	NaN	3	57	NaN	True
44	BachChorales/Chorale237	2	3/4	4/4	1	1	1	1/4	1/4	1	NaN	2	62	NaN	True

Rhythm





Task

This assignment focuses on quantifying the similarity between musical excerpts: specifically, between the incipits of Bach chorales.

(A) Start with a qualitative exploration of the data: e.g., listen to nn. 221, 300, 332 and try to identify salient similarities and differences

(B) Implement a measure of musical similarity for each of the following parameters:

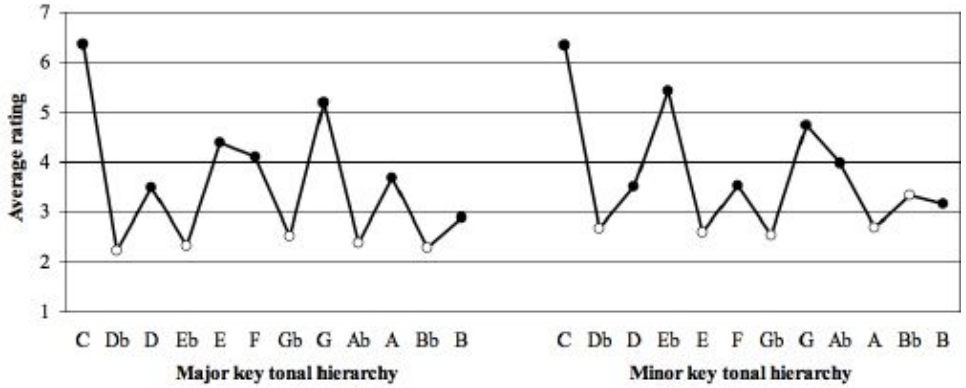
- Melody and melodic contour
- Rhythm
- Harmonic content
- …

The proposed measures should aim at reflecting the similarity between two musical excerpts as it is subjectively experienced.

(C) Evaluate your proposed measures by giving examples of excerpts that are predicted to be similar/dissimilar, and explaining in what ways the measure does or does not match your subjective experience.

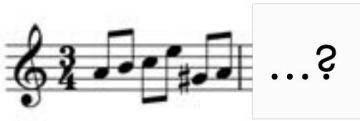
- Which one of these measure captures your hearing the best?
- Reflect on what is missing, and whether/how you could combine these measures to create a “compound” measure that reflects your intuition better

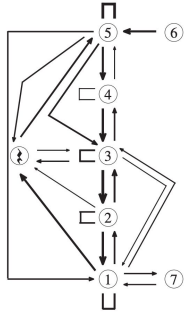
The role of temporality



			1 C.E.G	2 D.G.B	3 C.F.A	4 C.E.A	5 D.F.A	6 D.F.G.B	7 D.F#.A	8 E.G.B	9 C.D.F#.A	10 E.G#.B	11 C.D.F.A	12 C.E.G.A	13 D.F.B	14 C#.E.A	15 C.D.G
1	C.E.G	71	6,37	5,79	4,13	1,68	1,15	0,84	0,48	0,29	0,54	0,15	0,88	0,24	0,37	0,11	0,63
2	D.G.B	61	8,98	2,26	0,63	1,28	0,51	0,51	0,56	0,55	0,25	0,34	.	0,38	0,01	0,04	.
3	C.F.A	52	3,04	1,22	0,70	0,26	0,56	0,50	0,14	0,23	0,11	0,05	0,11	0,06	0,40	0,02	0,07
4	C.E.A	53	0,92	1,38	0,56	0,85	0,63	0,16	0,34	0,55	0,44	0,29	0,43	0,07	0,12	0,05	0,01
5	D.F.A	53	0,99	0,81	0,09	0,50	0,43	0,44	0,03	0,26	0,04	0,26	0,04	.	0,04	0,37	.
6	D.F.G.B	26	2,56	0,09	0,07	0,15	0,04	0,05	.	0,04	.	0,01	0,04	.	0,02	.	.
7	D.F#.A	22	0,11	1,69	.	0,04	.	0,04	0,08	0,04	0,11	0,06	.	.	0,01	0,01	.
8	E.G.B	28	0,35	0,18	0,63	0,42	0,11	0,09	0,04	0,11	0,02	0,03	0,01	0,33	0,09	0,02	.
9	C.D.F#.A	25	0,05	1,04	.	0,06	.	0,01	0,04	0,07	0,01	0,04	.	.	.	.	0,04
10	E.G#.B	18	0,02	.	0,20	1,06	0,03	.	.	0,02	.	0,09	.	0,01	.	0,04	.
11	C.D.F.A	18	0,11	1,12	0,02	0,01	0,04	0,04	.	0,02	.	0,05	.	.	0,01	.	0,04
12	C.E.G.A	19	0,04	0,09	0,08	0,03	0,12	0,04	0,47	.	0,09	.	.	0,01	0,04	.	0,04
13	D.F.B	23	0,80	0,02	0,03	0,04	0,01	0,05	.	0,04	.	0,02	.	.	.	.	.
14	C#.E.A	25	0,08	0,04	0,01	0,12	0,46	.	0,04	0,03	0,03	0,01	.	.	.	0,10	.
15	C.D.G	10	.	0,67	.	0,01	.	0,19	0,01	.	.	.	.	0,01	.	.	0,01

$$p\left(e_i \middle| e_{i-1}\right) = \frac{count\left(e^{i}_{i-1}\right)}{count\left(e_{i-1}\right)}$$



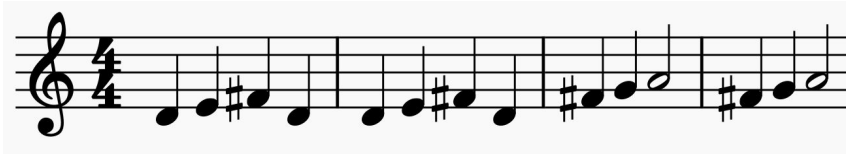


"STATIC"

"DYNAMIC"

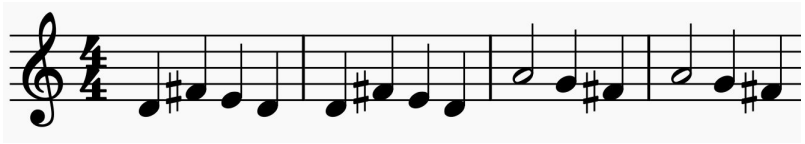
Melody

Transformations and invariances

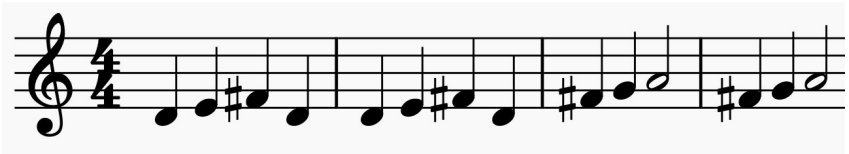


?

=

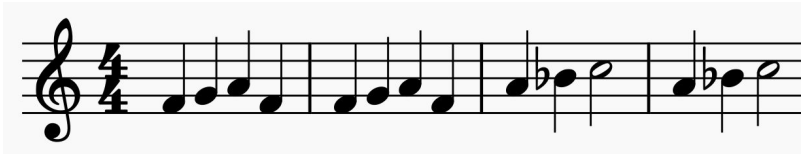


EDIT DISTANCE (note-by-note): 10
EDIT DISTANCE (inversion): 4



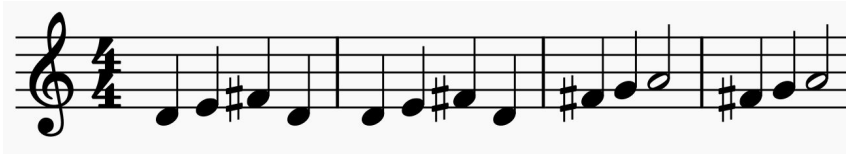
?

=



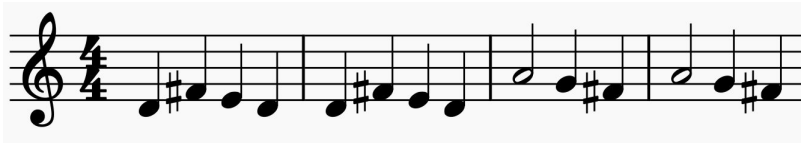
EDIT DISTANCE (note-by-note): 14
EDIT DISTANCE (intervals): 0

Transformations and invariances



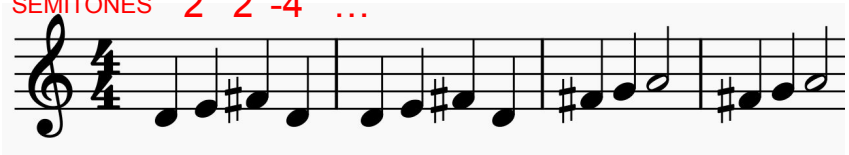
?

=



EDIT DISTANCE (note-by-note): 10
EDIT DISTANCE (inversion): 4

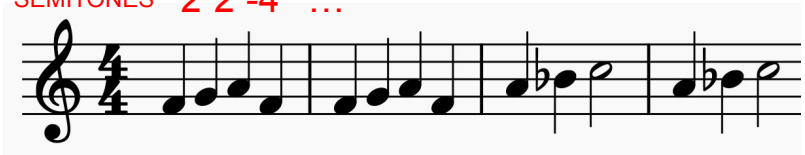
SEMITONES 2 2 -4 …



?

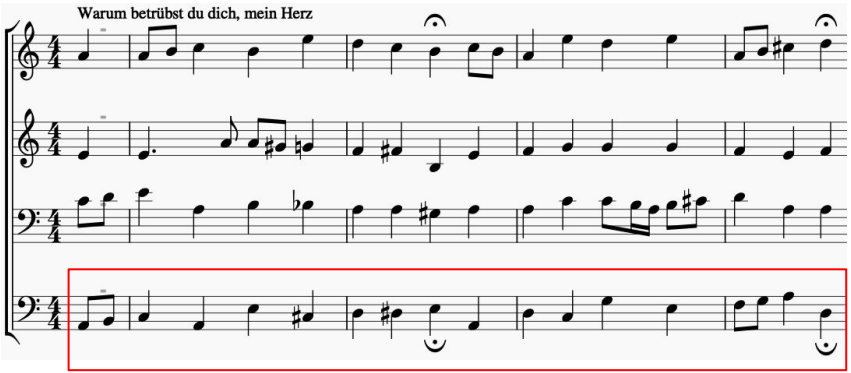
=

SEMITONES 2 2 -4 …

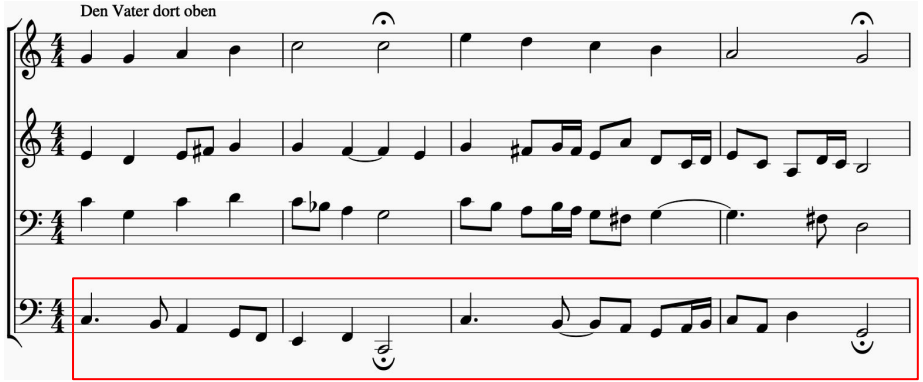


EDIT DISTANCE (note-by-note): 14
EDIT DISTANCE (intervals): 0

Steps vs. jumps

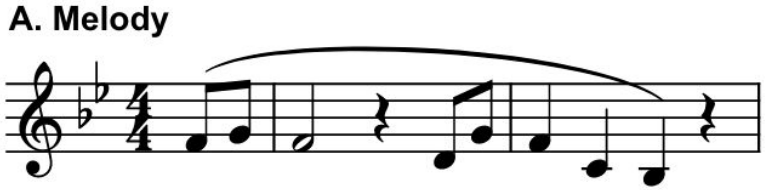


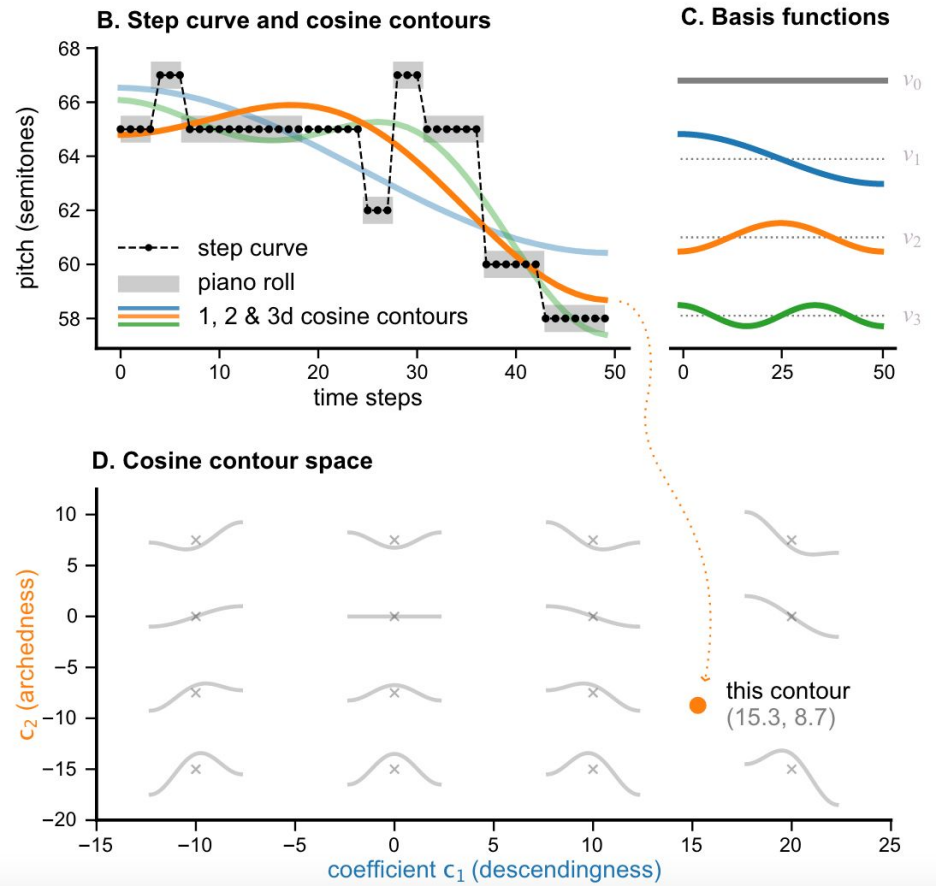
300



239

Contour



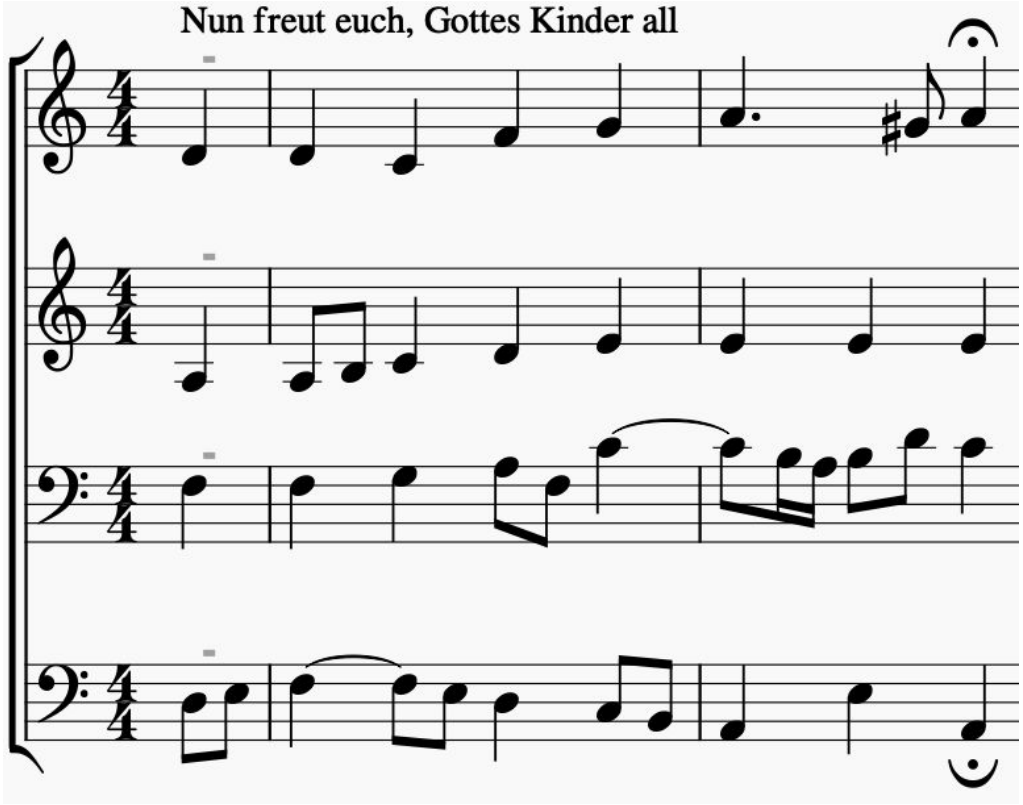


Cornelissen, B., Zuidema, W., & Burgoyne, J. A. (2021).

Rhythm

Rhythm and texture

185



Harmony

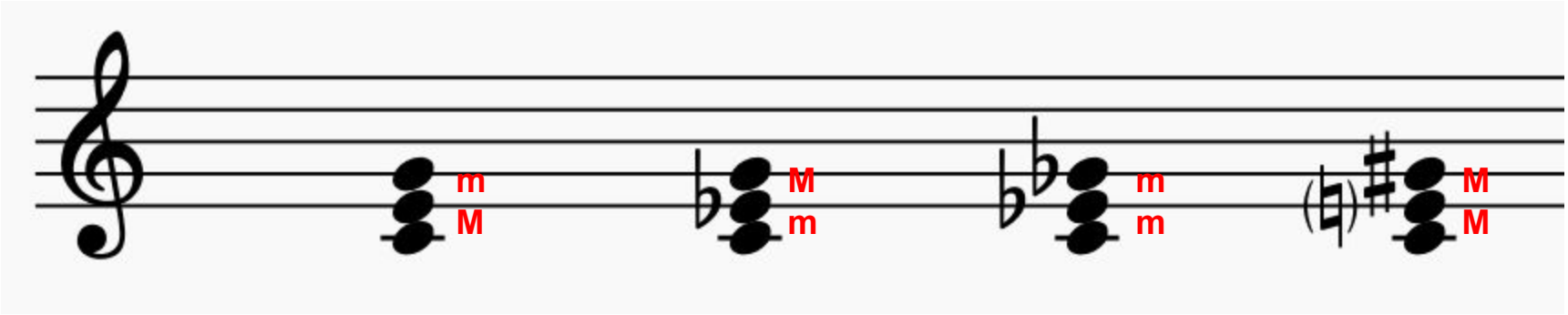
Harmonic content

Chords

- The prototype of a chord in Western tonality is the (root-position) triad:
  - Stacked (major or minor) thirds above a root
  - 4 types (M, m, +, o) x 12 roots = 48 root position triads

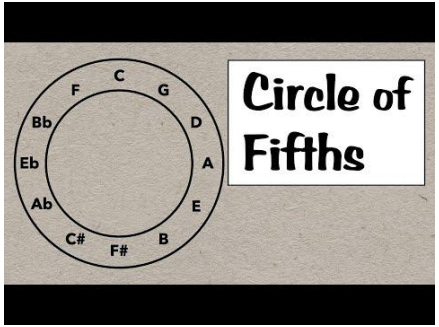
Triads

http://musictheory.pugetsound.edu/mt21c/TriadsIntroduction.html



Major (M) minor (m) diminished (o) Augmented (+)

Chords



- The prototype of a chord in Western tonality is the (root-position) triad:
  - Stacked (major or minor) thirds above a root
  - 4 types (M, m, +, o) x 12 roots = 48 root position triads
  - Music is creative/messy: chords may contain additional (e.g., sevenths) and/or extraneous notes!

- “Distance” between chords:
  - Distance between the roots (in fifths)
  - Number of common tones
  - “Voice-leading” distance: total distance required to match the second chord by moving the notes of the first chord by the minimal amount of semitones
  - …

- Consonance/Dissonance

Consonance/Dissonance

Consonances

Perfect

- Unison and octave
- Fifth (ratio of 3:2)
- Fourth

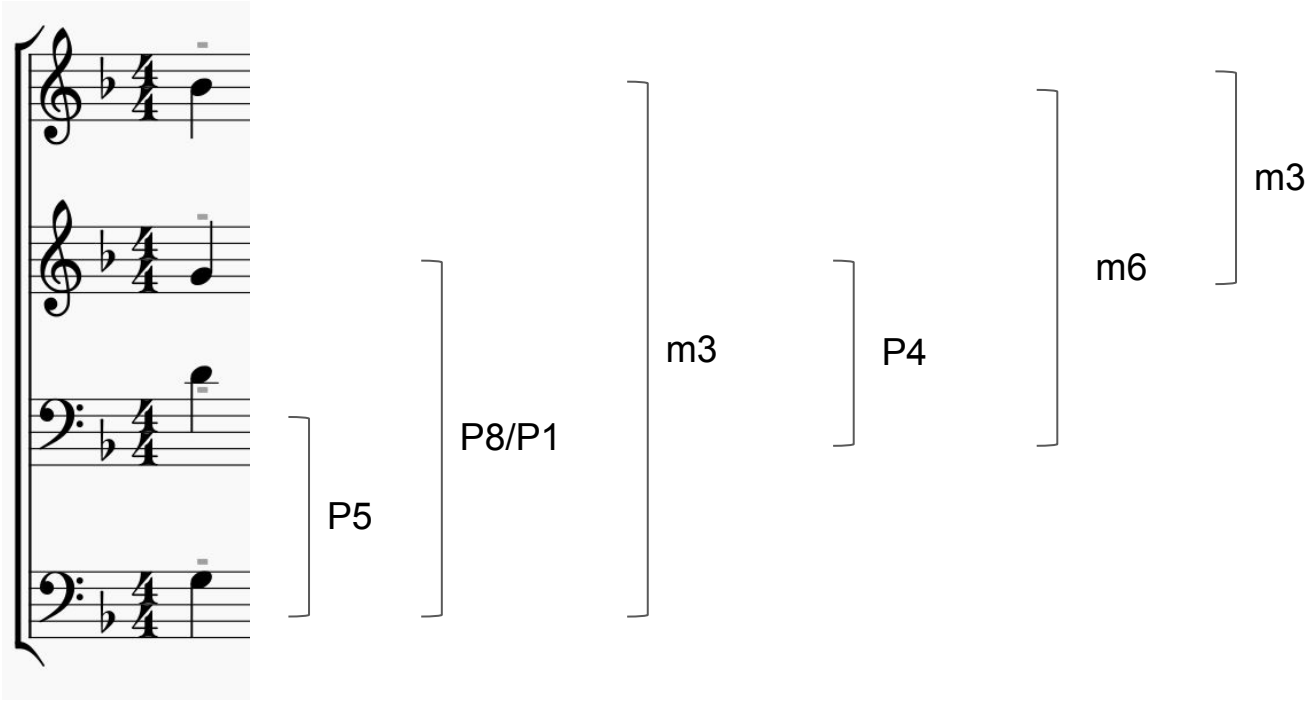
Imperfect

- Third (major and minor)
- Sixth (major and minor)

Dissonances

- Second (major and minor)
- Seventh (major and minor)
- The tritone (ratio of 45:32!)
  - Augmented fourth or diminished fifth
  - *Diabolus in musica* (!)

Chords as interval vectors



Interval vs. interval class

Smallest interval between PC a and PC b (irrespective of direction)

$IC(a,b) = \min\{I(a,b), I(b,a)\}$

P1/P8 → 0

m2/M7 → 1

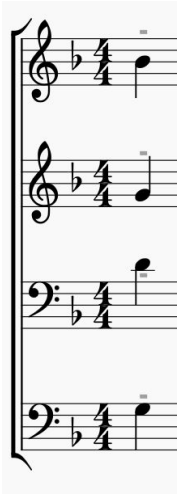
M2/m7 → 2

m3/M6 → 3

M3/m6 → 4

P4/P5 → 5

4+/5° → 6



IC	1	2	3	4	5	6
#	0	0	2	1	2	0

Pitch-class sets

- Chords as sets of pitch-classes
- Reduce redundancy: ignore repeated notes, reduce to “normal form”
- https://musictheory.pugetsound.edu/mt21c/IntervalVector.html