

**Computergestützte Analyse von  
Musikaufnahmen – ein Beispiel für  
interdisziplinäre Forschung**

Dr. Christof Weiß

International Audio Laboratories Erlangen

26.02.2019 | 12.30 Uhr | Raum WE5/02.005

# Christof Weiß



- 2006: Abitur, Max-Reger-Gymnasium Amberg
- 2006-2012: Studium **Physik** Diplom, Universität Würzburg
- 2006-2011: Studium **Komposition**, HfM Würzburg
- 2011-2012: Fortbildungsklasse Komposition

- 2012-2015: **Promotion**  
Fraunhofer Institut für Digitale Medientechnologie, Ilmenau  
gefördert von Stiftung der Deutschen Wirtschaft (sdw)

*Computational Methods for Tonality-Based Style Analysis of Classical Music Audio Recordings*

- Seit 09/2015: AudioLabs Erlangen / freischaffender Komponist
- 2018: KlarText-Preis für Wissenschaftskommunikation

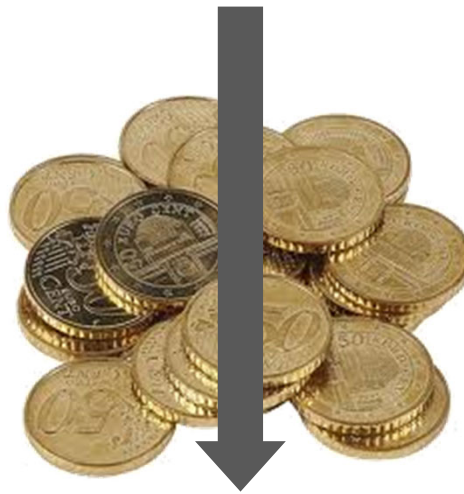


# International Audio Laboratories Erlangen

 **Fraunhofer**  
IIS



**FAU** FRIEDRICH-ALEXANDER  
UNIVERSITÄT  
ERLANGEN-NÜRNBERG



**AUDIO**  
**LABS**

**AUDIO**  
**LABS**

# AudioLabs - FAU

- Prof. Dr. Jürgen Herre  
Audio Coding
- Prof. Dr. Bernd Edler  
Audio Signal Analysis
- Prof. Dr. Meinard Müller  
Semantic Audio Processing
- Prof. Dr. Emanuël Habets  
Spatial Audio Signal Processing
- Prof. Dr. Frank Wefers  
Virtual Reality
- Dr. Stefan Turowski  
Coordinator AudioLabs-FAU



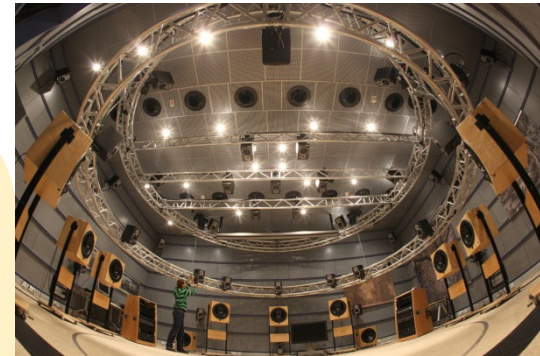
**AUDIO**  
**LABS**

# International Audio Laboratories Erlangen

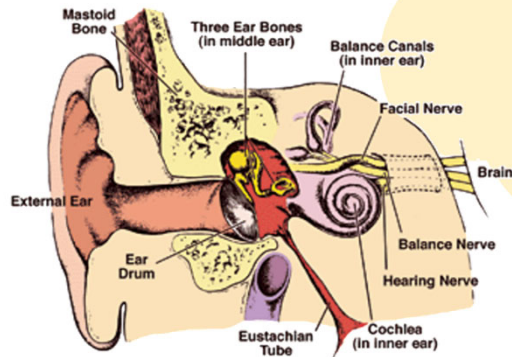
Audio Coding



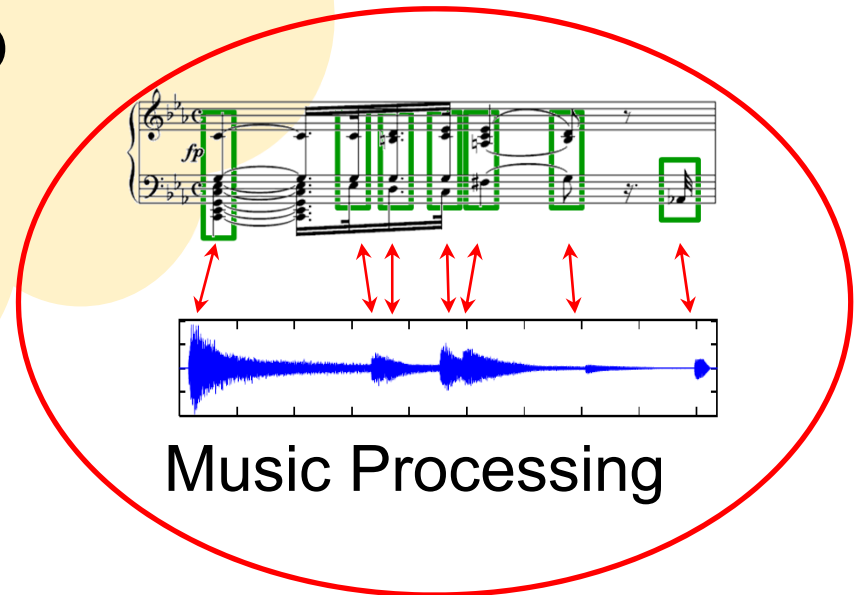
3D Audio



Audio



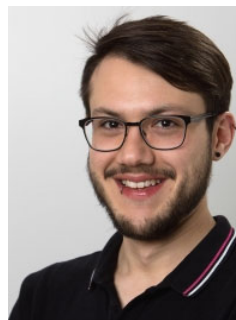
Psychoacoustics



Music Processing

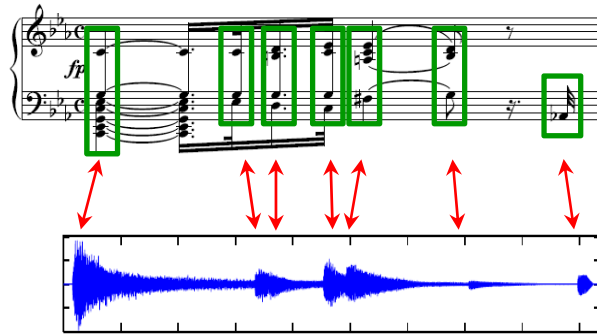
# Group Prof. Meinard Müller

- Patricio López-Serrano
- Frank Zalkow
- Sebastian Rosenzweig
- Christof Weiß

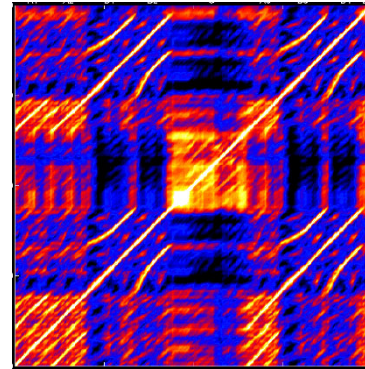


# Music Processing / Music Information Retrieval (MIR)

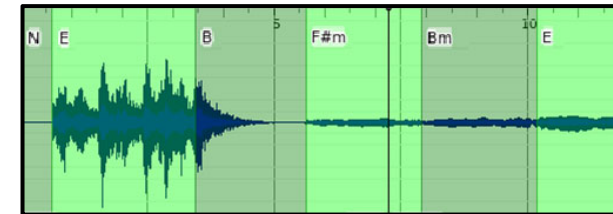
## Music Synchronization



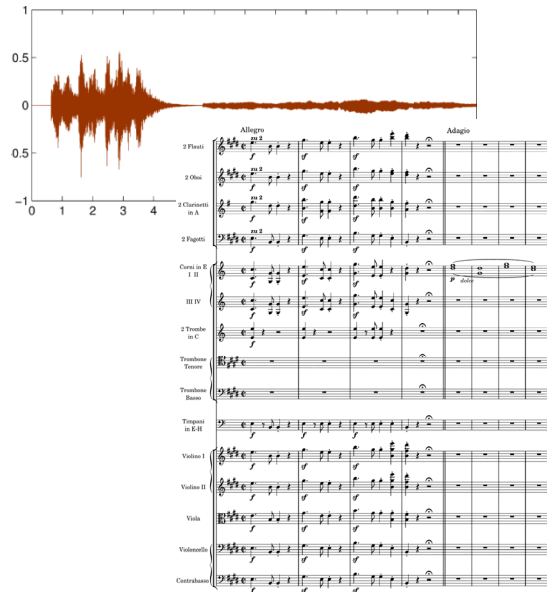
## Structure Analysis



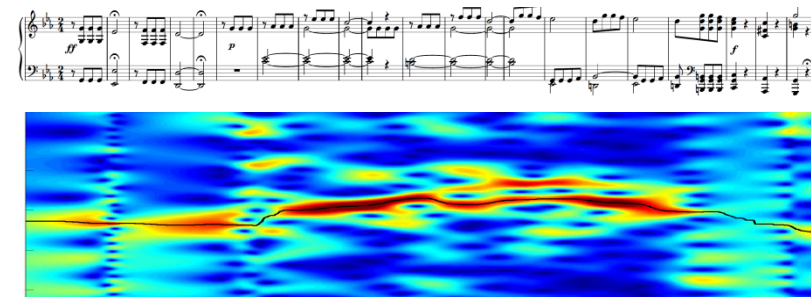
## Harmony Analysis



## Automatic Music Transcription

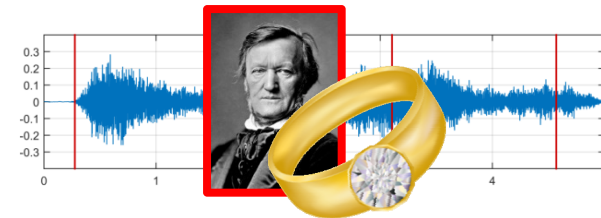


## Tempo Estimation and Beat Tracking

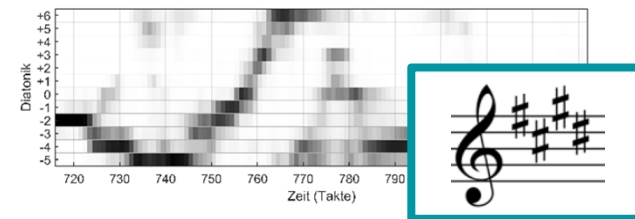


# Outline

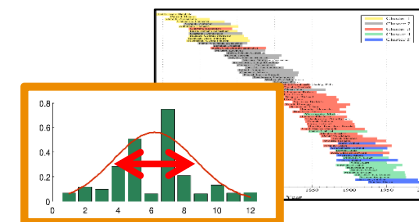
1. Measure Annotations in Wagner's *Ring*



2. Cross-Version Analysis of Harmonic Structures: Local Keys and Chords



3. Machine Learning and Corpus Analyses in Classical Music and Jazz





# Outline

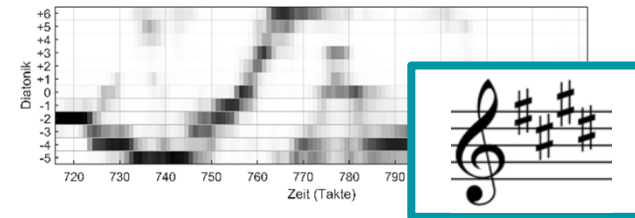
1. Measure Annotations in Wagner's *Ring*

→ *Software example: Sonic Visualizer*



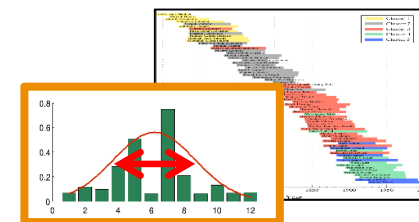
2. Cross-Version Analysis of Harmonic Structures: Local Keys and Chords

→ *Programming example: Python and Jupyter Notebooks*



3. Machine Learning and Corpus Analyses in Classical Music and Jazz

→ *Discussion: Chances and Challenges of Interdisciplinary Research*



# Material

- Workshop website with resources:

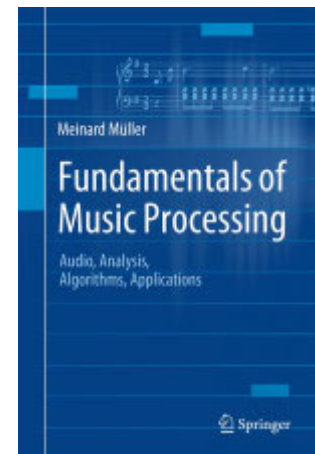
[www.audiolabs-erlangen.de/resources/MIR/2019\\_WorkshopMIR\\_UniBamberg](http://www.audiolabs-erlangen.de/resources/MIR/2019_WorkshopMIR_UniBamberg)

- Book: M. Müller, *Fundamentals of Music Processing*

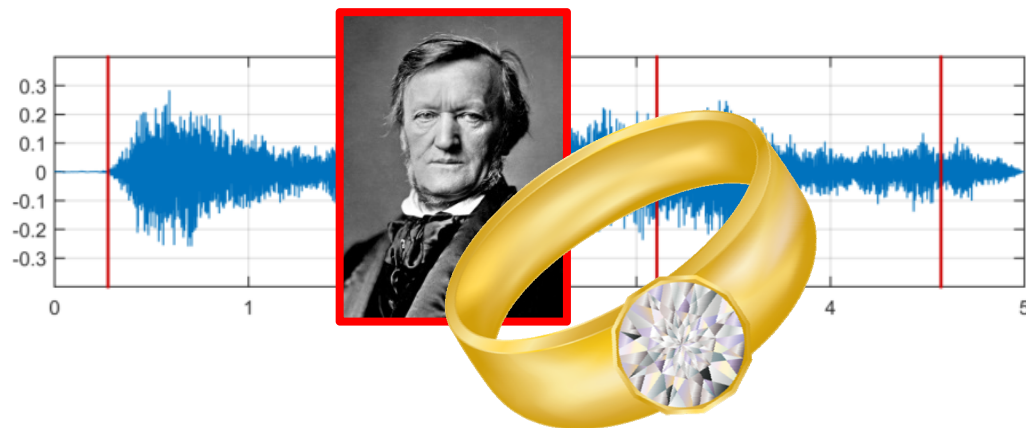
[www.audiolabs-erlangen.de/fau/professor/mueller/bookFMP](http://www.audiolabs-erlangen.de/fau/professor/mueller/bookFMP)

- Jupyter Notebooks in Python:

[www.audiolabs-erlangen.de/FMP](http://www.audiolabs-erlangen.de/FMP)



# 1. A Typical Scenario – Measure Annotations in Wagner's *Ring*



# DFG-funded Project: Computational Analysis of Harmonic Structures

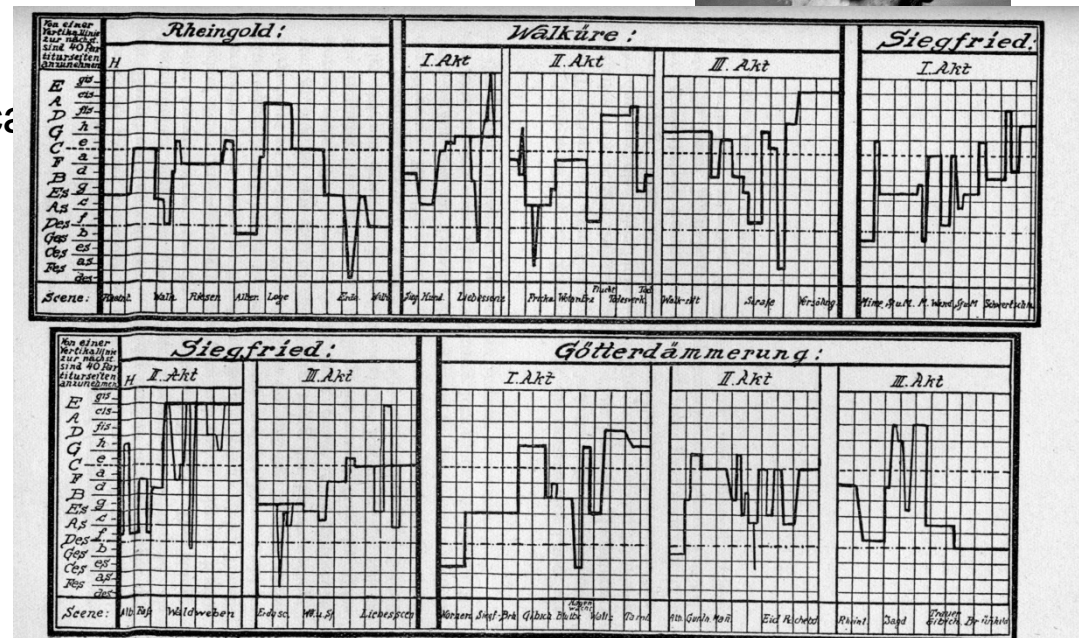


- With Prof. Rainer Kleinertz, Musicology, Uni Saarland



UNIVERSITÄT  
DES  
SAARLANDES

- Richard Wagner, *Der Ring des Nibelungen*
  - Four operas, up to 15 hours of music
  - How is harmony organized at the large scale?
  - Analyses by A. Lorenz 1924
  - Hypothesis of „Poetico-musical“

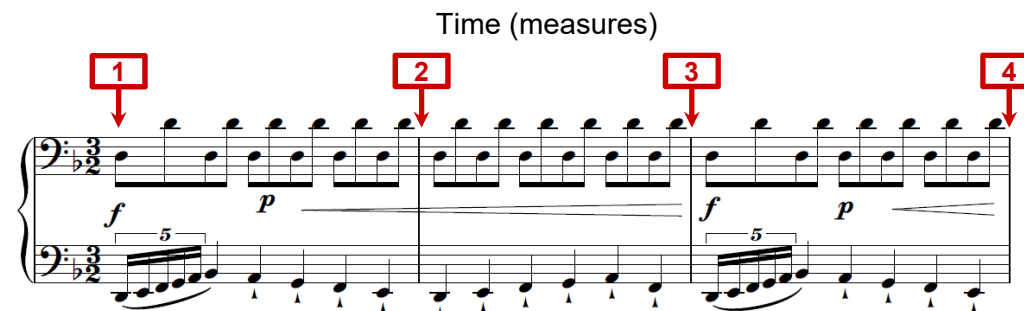


# Music Scenario

- Richard Wagner, *Die Walküre* (opera)
  - Long work (1<sup>st</sup> act: 67 minutes)
  - No interruptions of acts
- Different data types
  - Libretto (**text**)
  - Score / piano reduction (**sheet music**)
  - Recorded performance (**audio**)

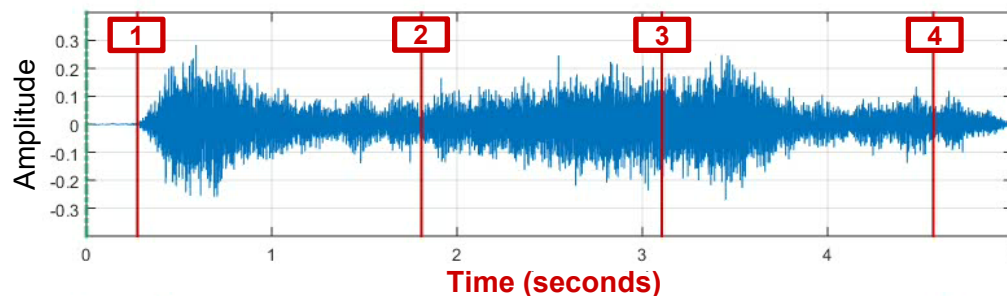
# Music Scenario

Score:

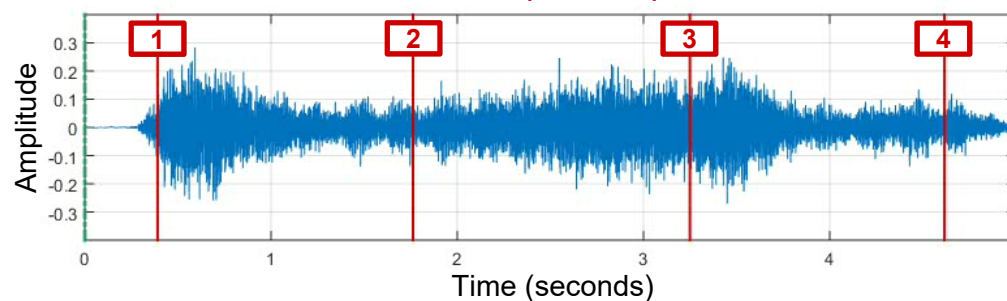


Performance (Karajan 1966):

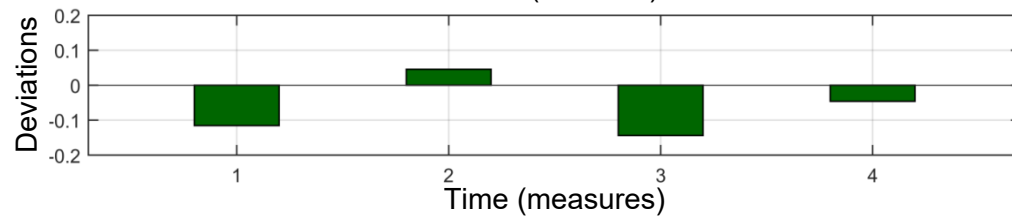
- Annotation 1



- Annotation 2

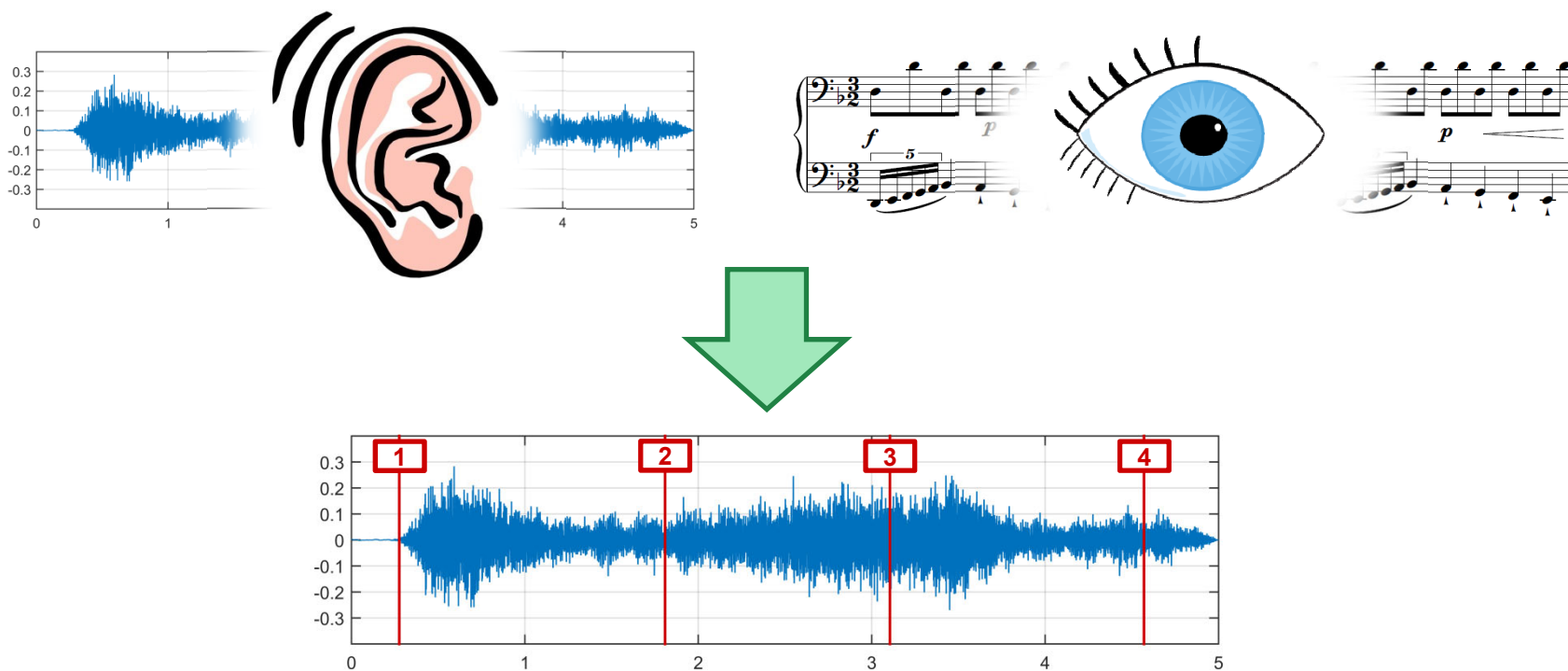


- Deviations



# Manual Measure Annotations

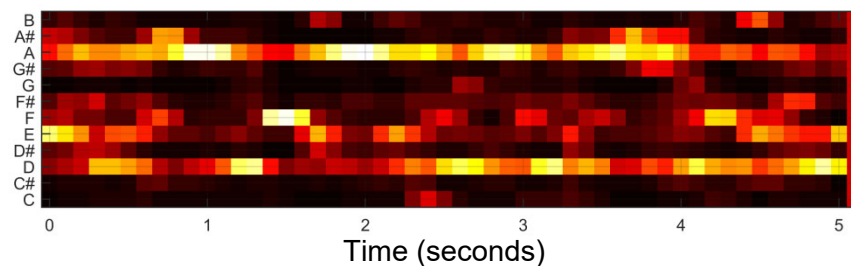
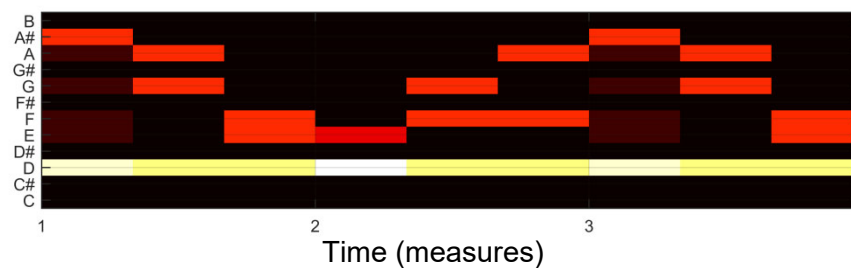
- 5 students with musical background
- Procedure: Listening while reading the vocal score
- Tool: *Sonic Visualiser*



→ *Software example: Measure Annotations in Sonic Visualizer*

# Computed Measure Annotations

- Synchronization (score-to-audio alignment)
- Based on *chroma features*





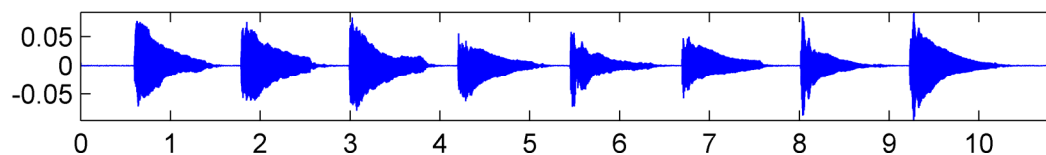
# Excursus: Chroma Features

- Example: C-major scale (piano)

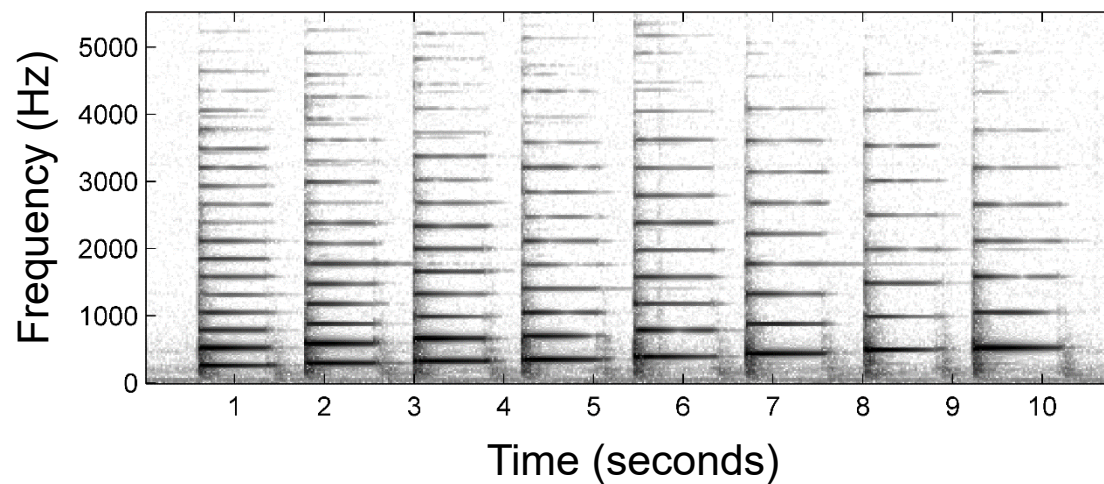
- Score



- Audio – **Waveform**



- Audio - **Spectrogram**



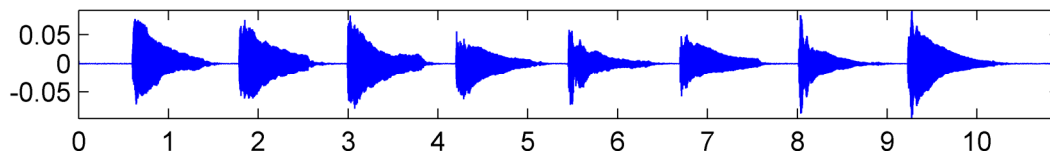
# Excursus: Chroma Features

- Example: C-major scale (piano)

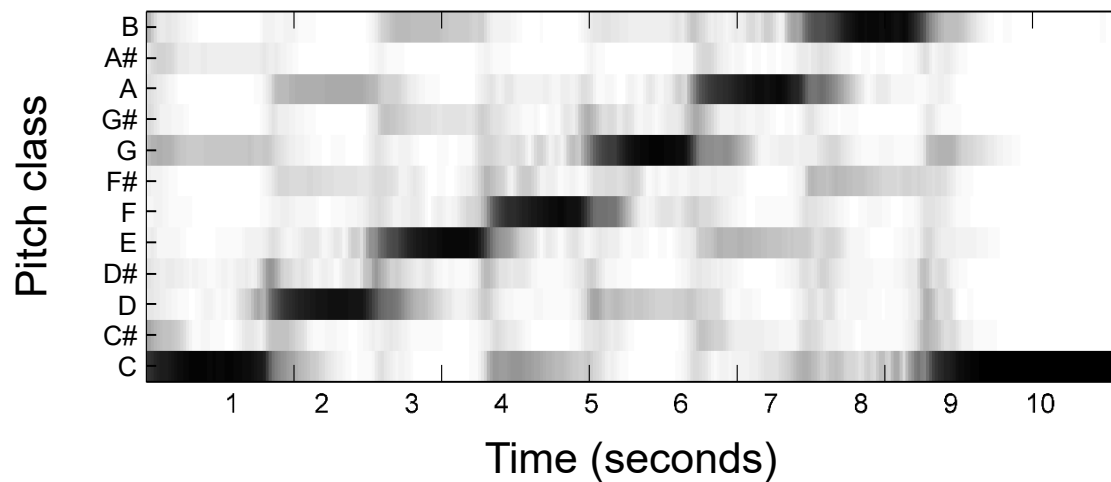
- Score



- Audio – **Waveform**



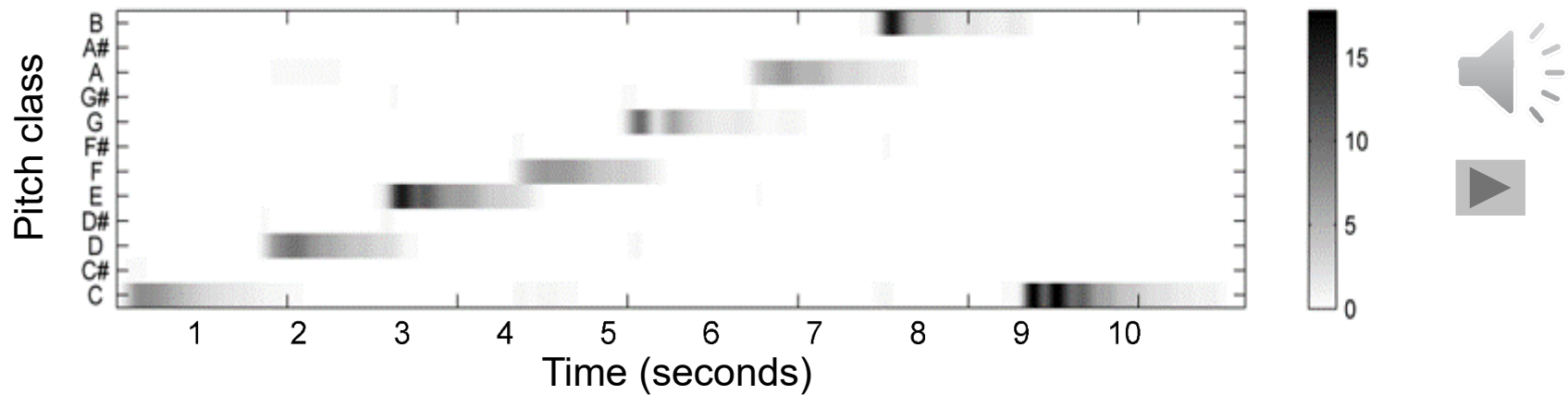
- Audio - **Chromagram**



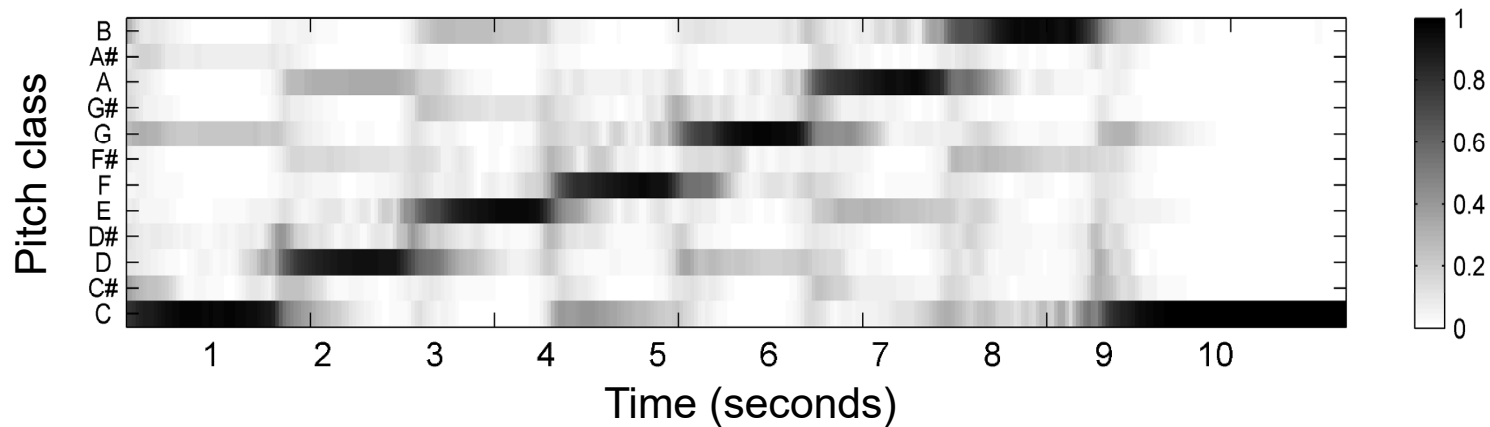
# Chroma Features

- Example: C-major scale (piano)

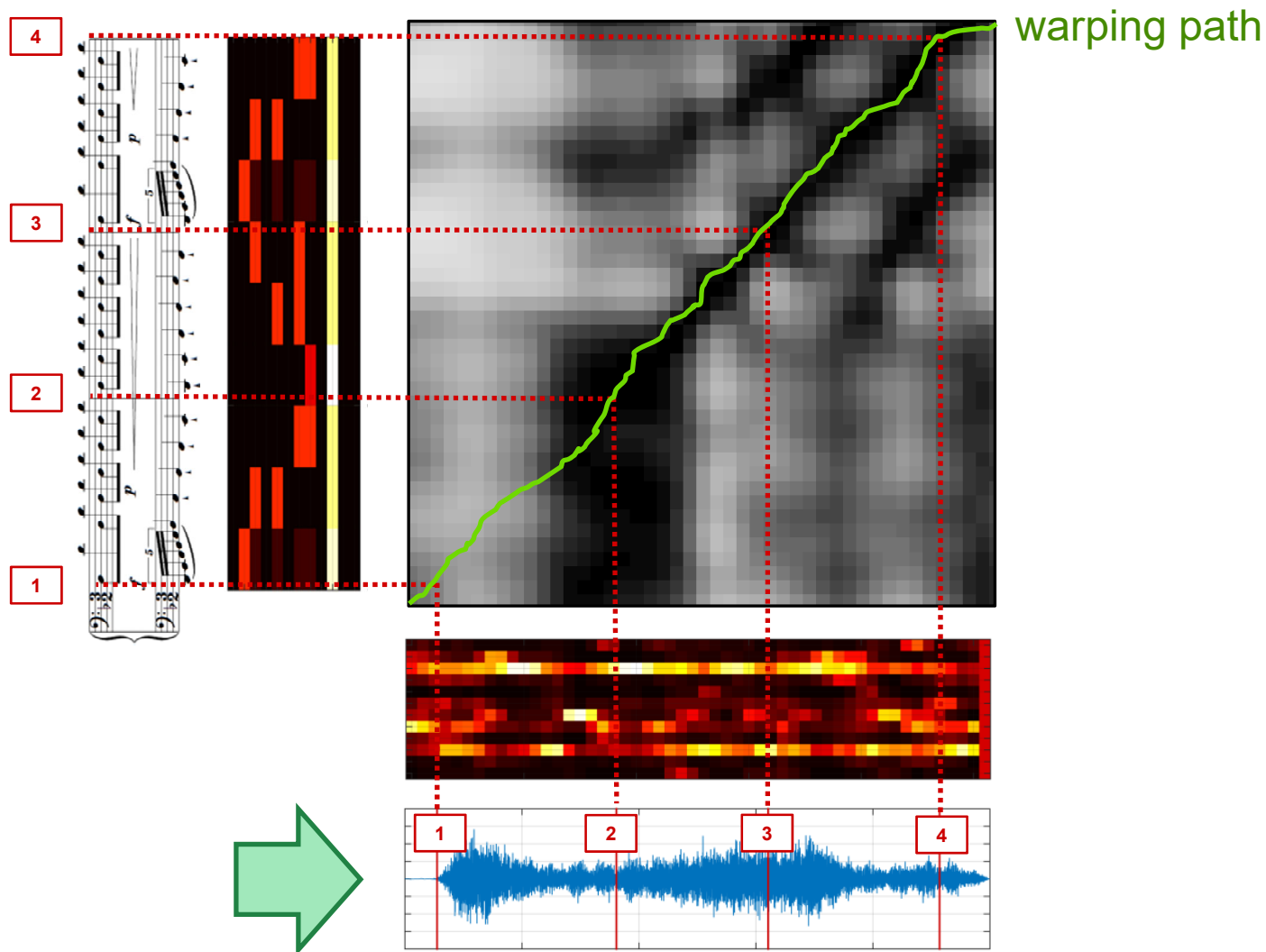
- Audio – **Chromagram**



- Audio – **Chromagram (normalized)**



# Computed Measure Annotations

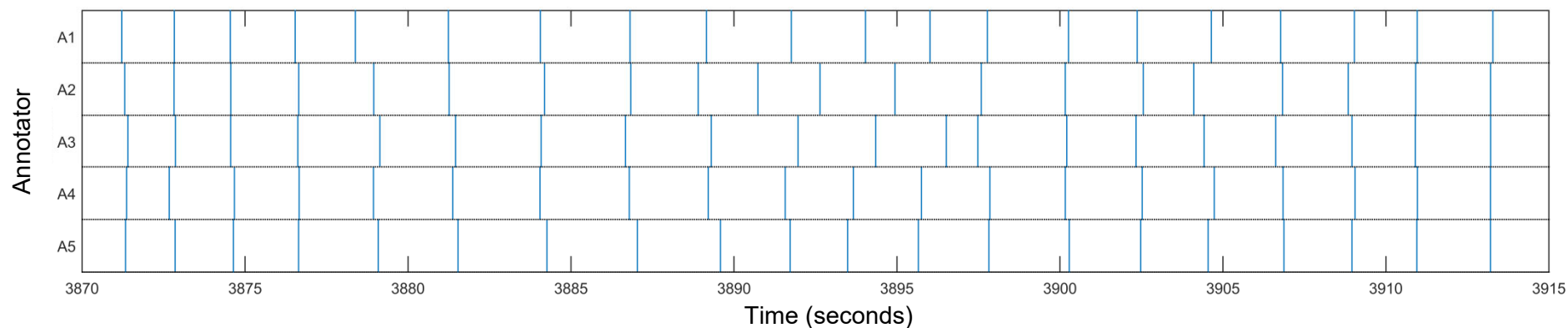


# Analysis of Manual Annotations

- Compare 5 different annotators
- Questions:
  - Accuracy?
  - Typical errors?
  - Systematic offsets?
- Example passage (Karajan 1966)

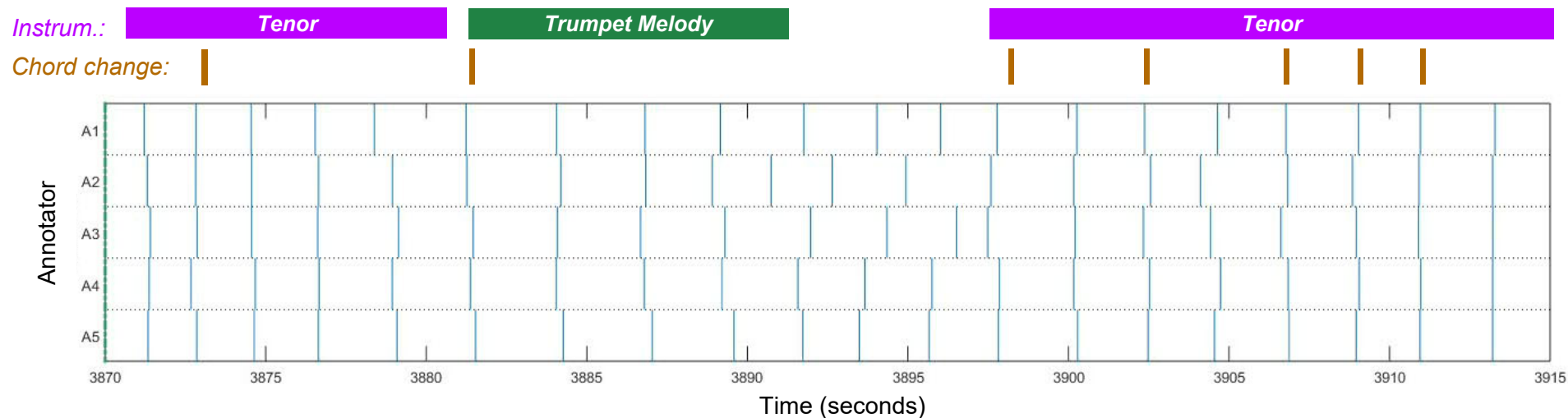
Instrum.: Tenor

Chord change: |

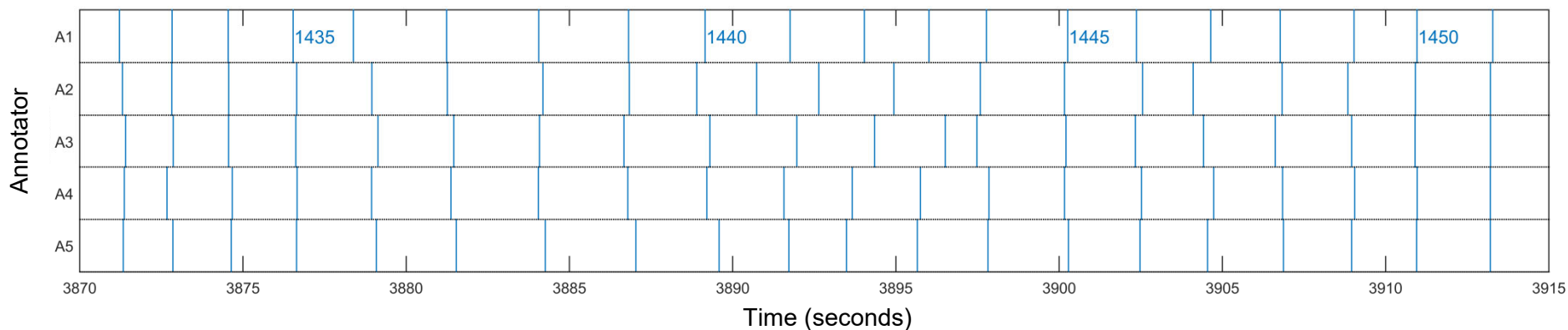
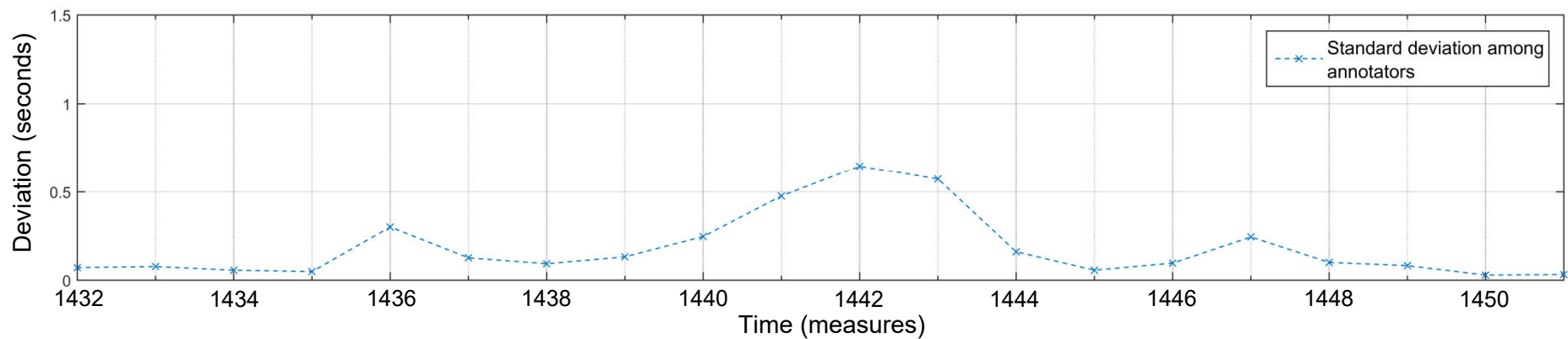


# Analysis of Manual Annotations

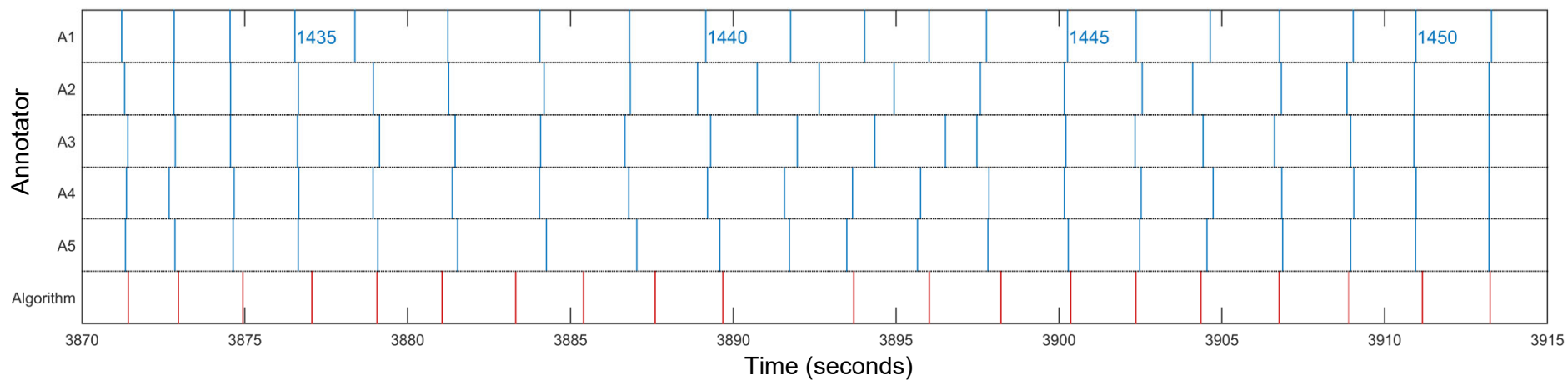
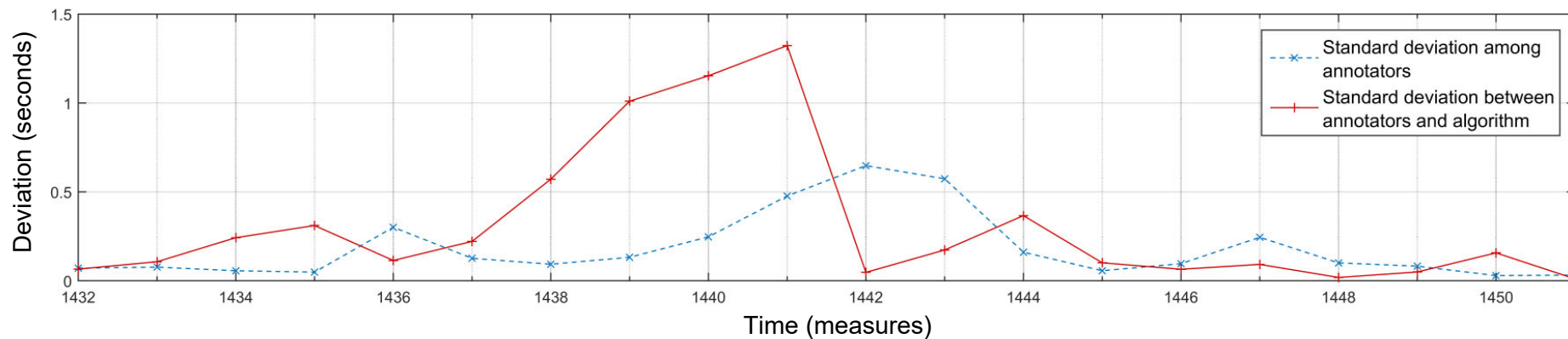
- Compare 5 different annotators
- Questions:
  - Accuracy?
  - Typical errors?
  - Systematic offsets?
- Example passage (Karajan 1966)



# Analysis of Manual Annotations

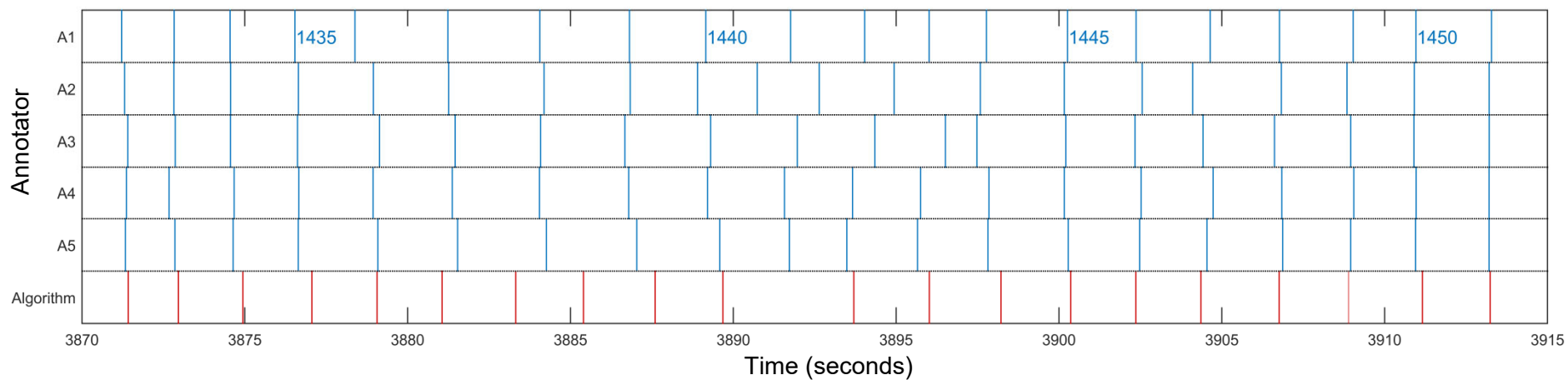
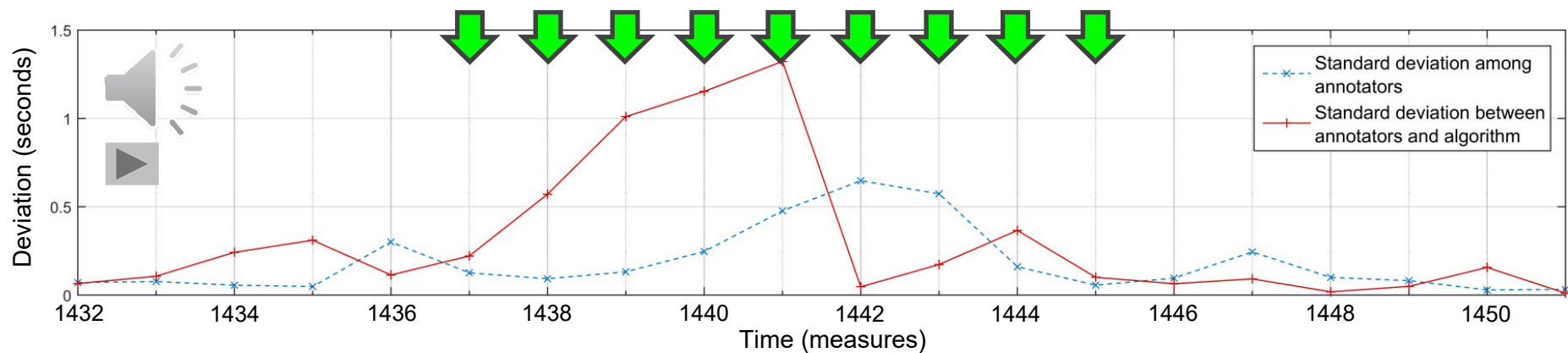


# Analysis of Computed Annotations



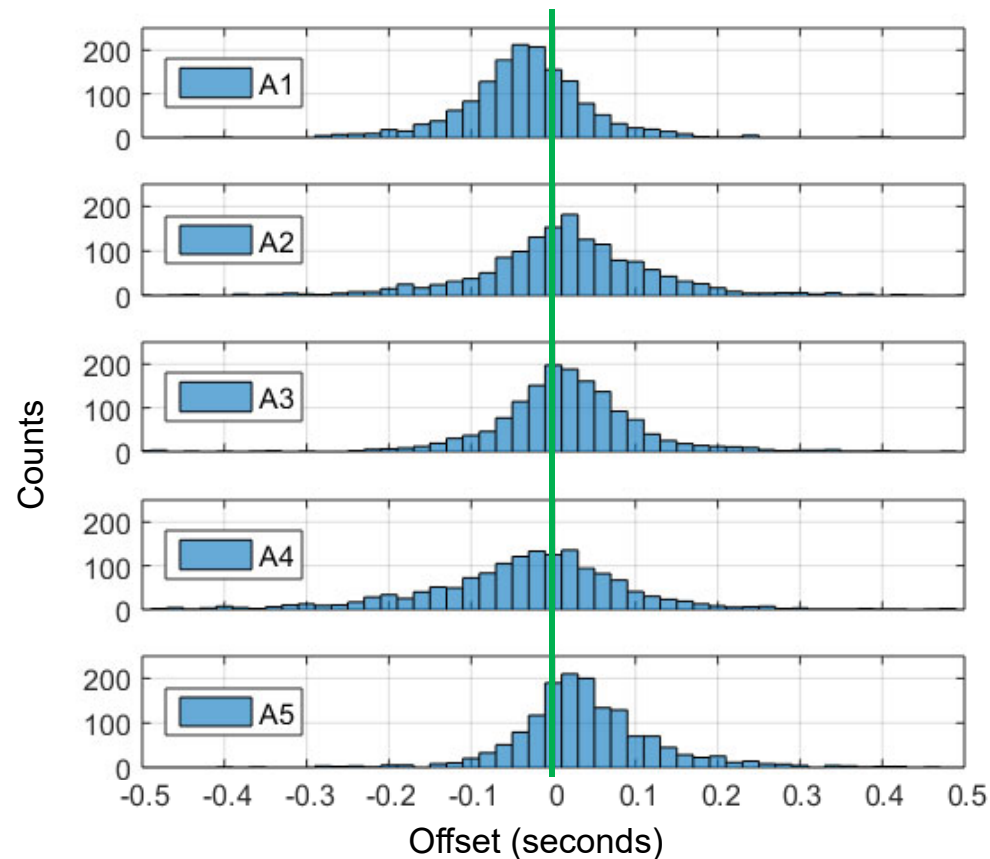


# Analysis of Computed Annotations



# Analysis of Manual Annotations

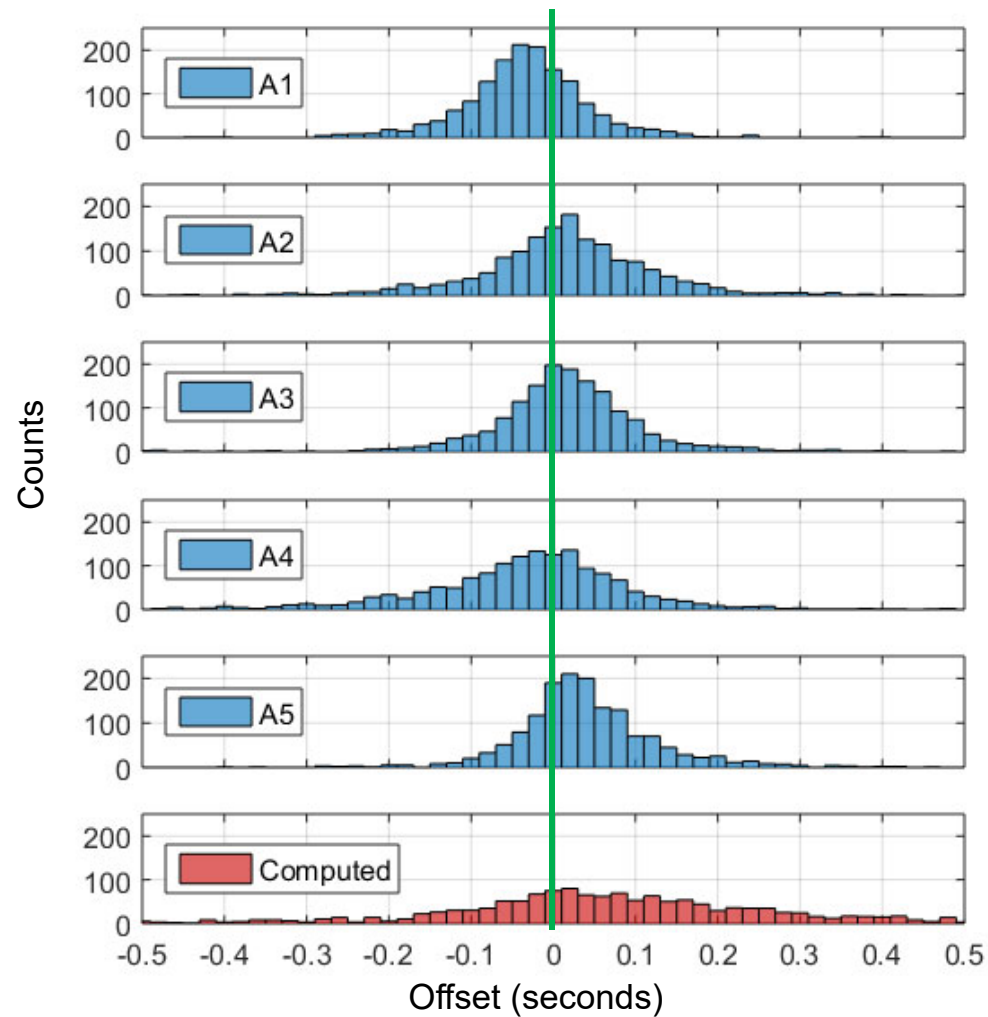
Dataset: Full act (67 minutes, 1523 measures)



0 s = mean position  
of A1 ... A5

# Analysis of Computed Annotations

Dataset: Full act (67 minutes, 1523 measures)

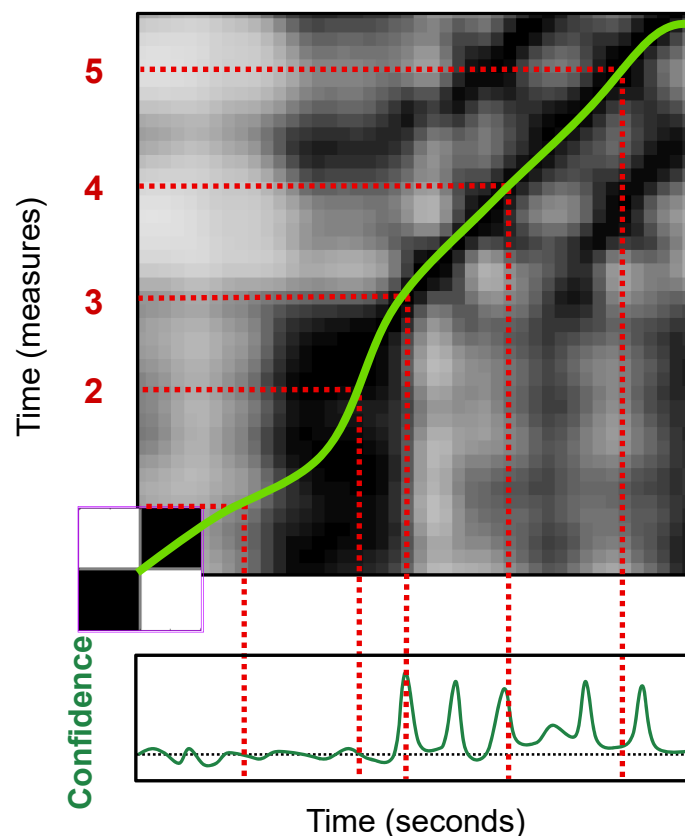


0 s = mean position  
of A1 ... A5

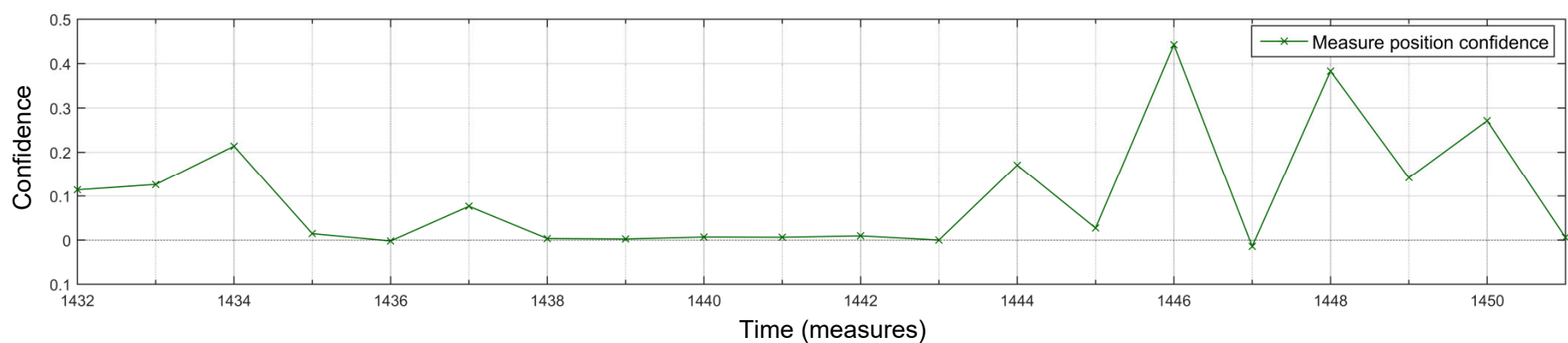
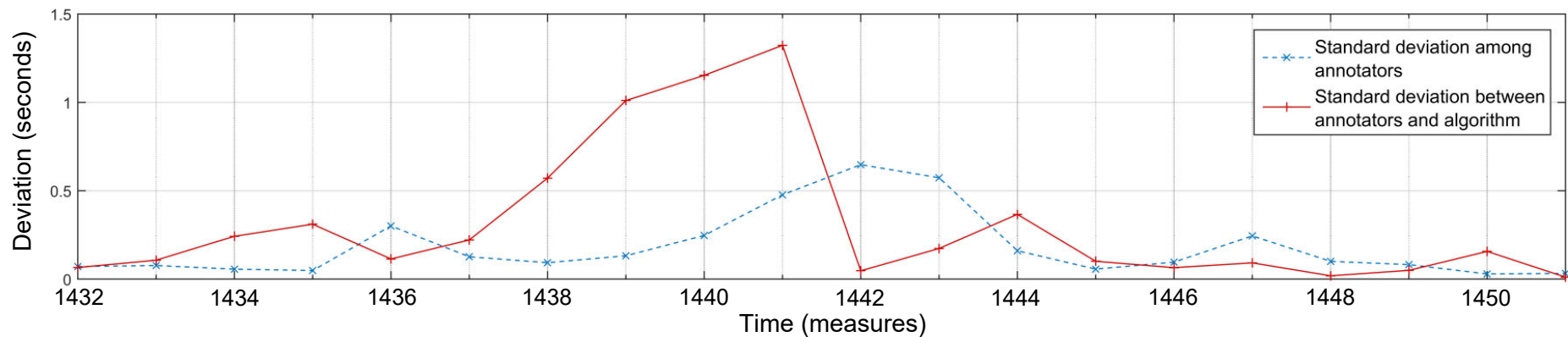
# Confidences for Computed Annotations

- Use information from **similarity matrix**
- Shift checkerboard kernel **along the warping path**

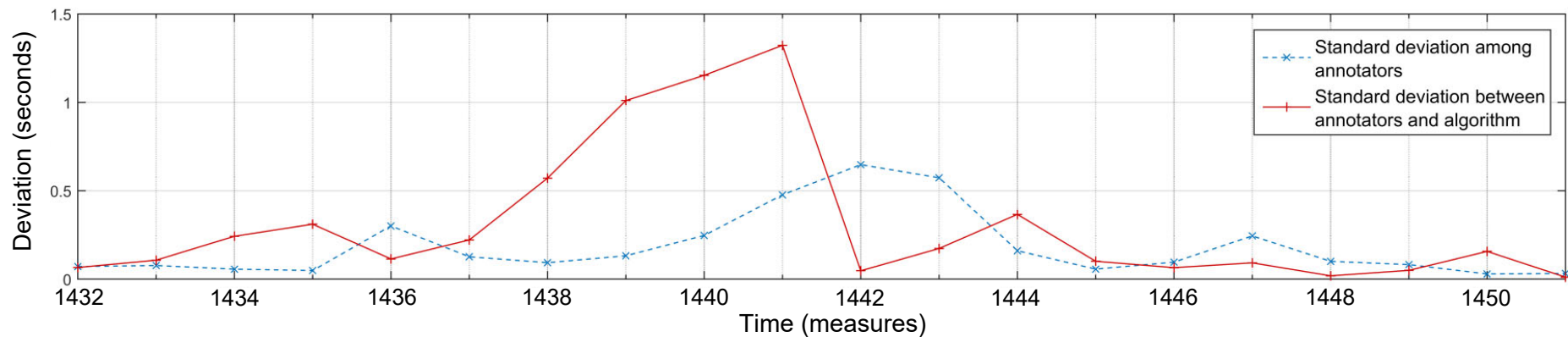
*J. Foote (2000):*  
Shift checkerboard kernel  
**along the main diagonal**  
**of a self-similarity matrix**



# Confidences for Computed Annotations

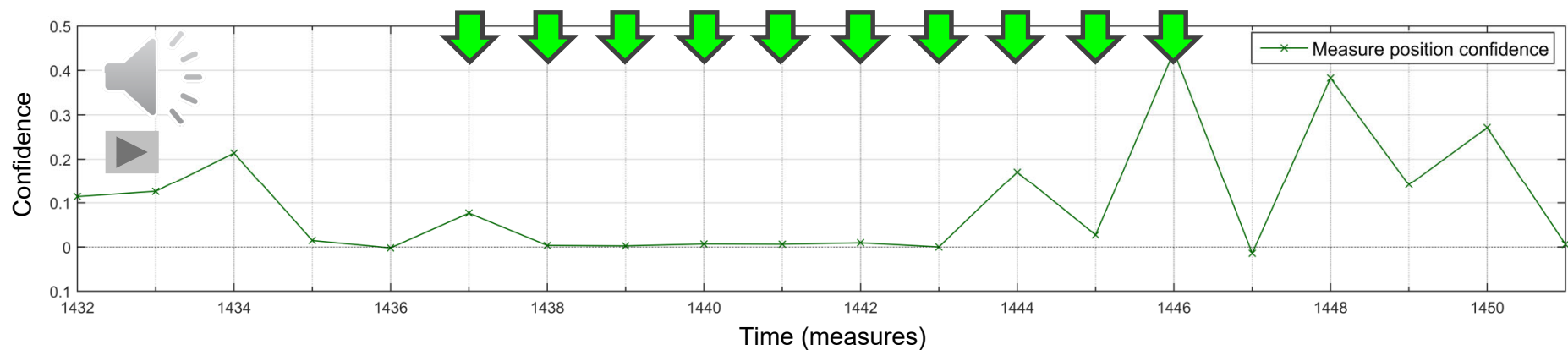


# Confidences for Computed Annotations



Instrum.: Tenor Trumpet Melody Tenor

Chord change: | | | | | | | | | |



# Towards Confidence-based Measure Annotations

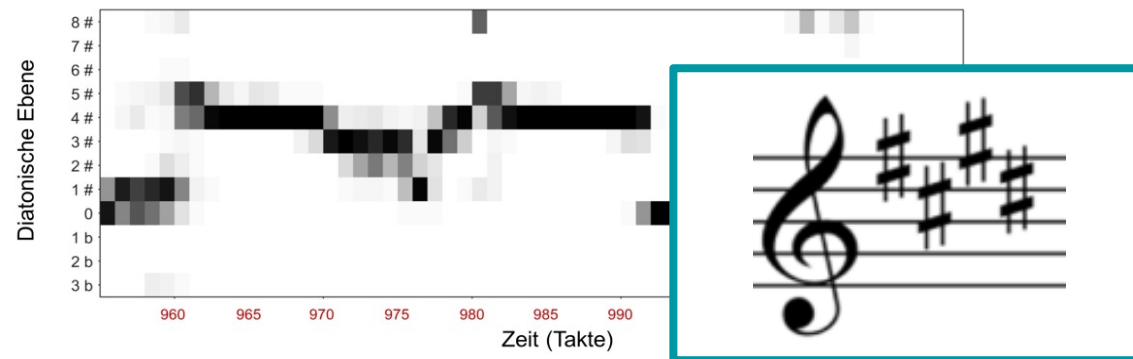
High confidence implies high reliability (quantitative evaluation, full act)

## Future work:

- Reliable measures → anchor points
- → improve computed annotations
- Other types of features

C. Weiß, V. Arifi-Müller, T. Prätzlich, R. Kleinertz, M. Müller  
“Analyzing Measure Annotations for Western Classical Music Recordings”  
In: *Proceedings of the 17th International Society for Music Information Retrieval Conference (ISMIR)*, New York, USA 2016.

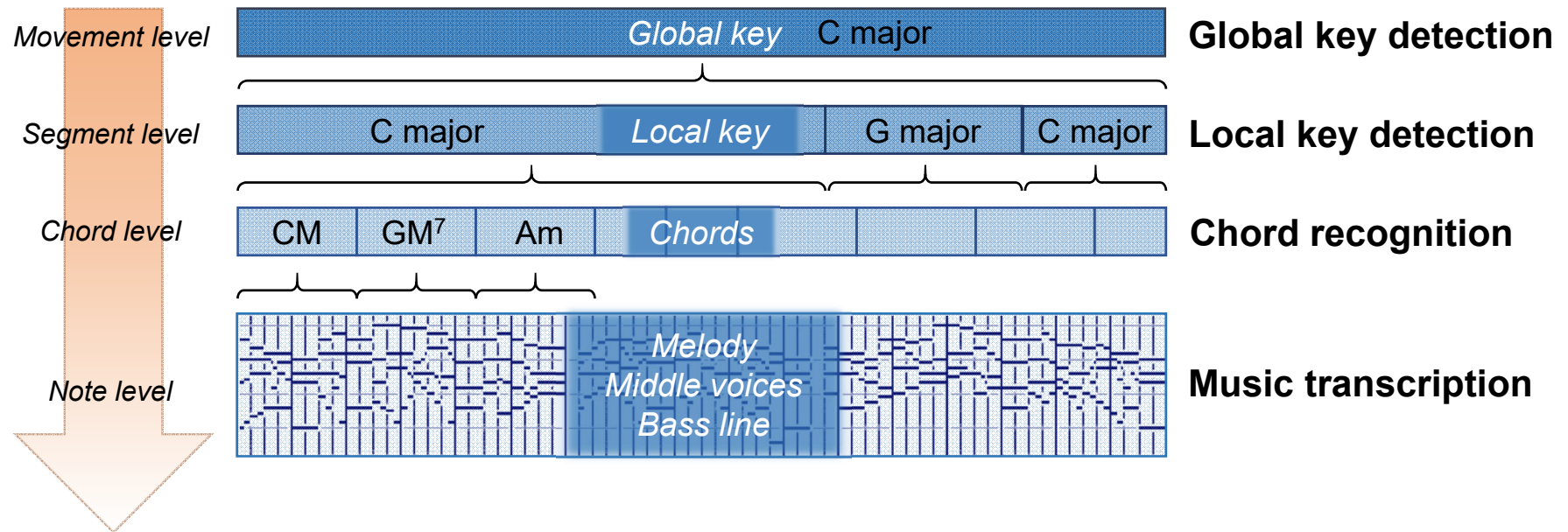
## 2. Cross-Version Analysis of Harmonic Structures: Local Keys and Chords





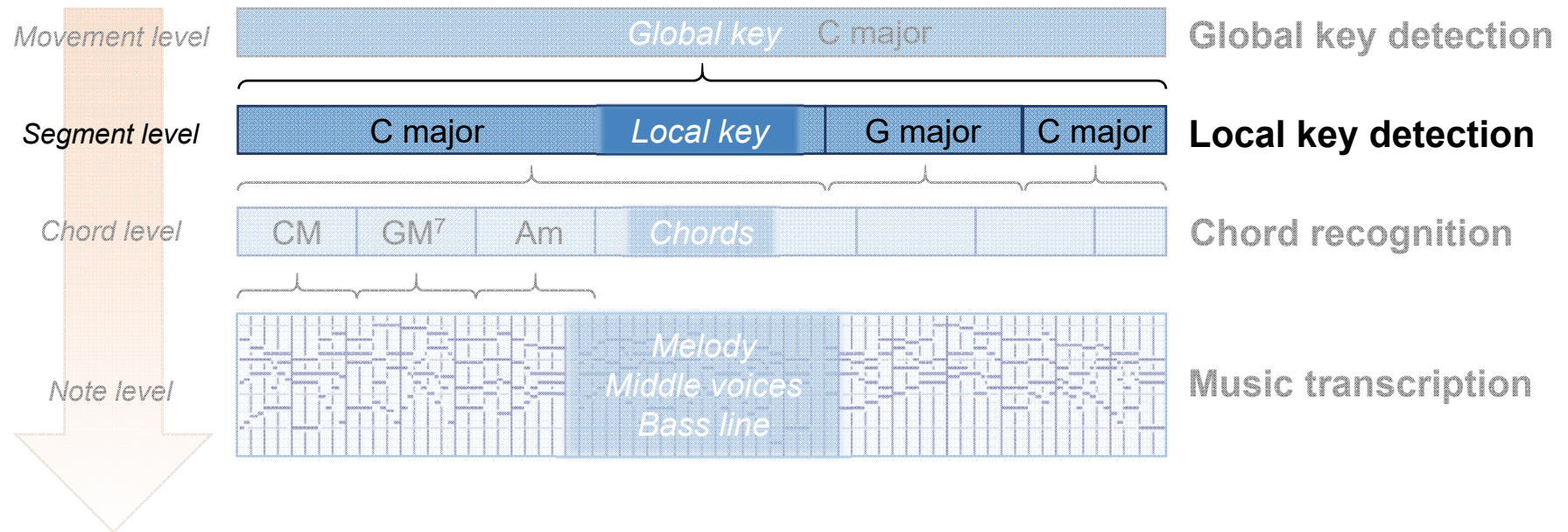
# Motivation

- **Harmony analysis** of music:
  - Different concepts
  - Concepts relate to different **temporal granularity**



# Motivation

- **Harmony analysis** of music:
  - Different concepts
  - Concepts relate to different **temporal granularity**



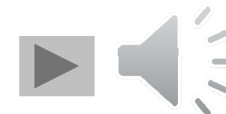
# Musicological Foundations

- Method: estimate **diatonic scales** – 7 fifth-related pitches
- Relationship of diatonic scales:
  - Fifth-neighbouring scales share 6 of 7 notes
  - Ordering of scales according to the **circle of fifths**:



# Visualization of Diatonic Scales

- Example: J.S. Bach, Choral "Durch Dein Gefängnis" (*Johannespassion*)
- **Score** – Piano reduction



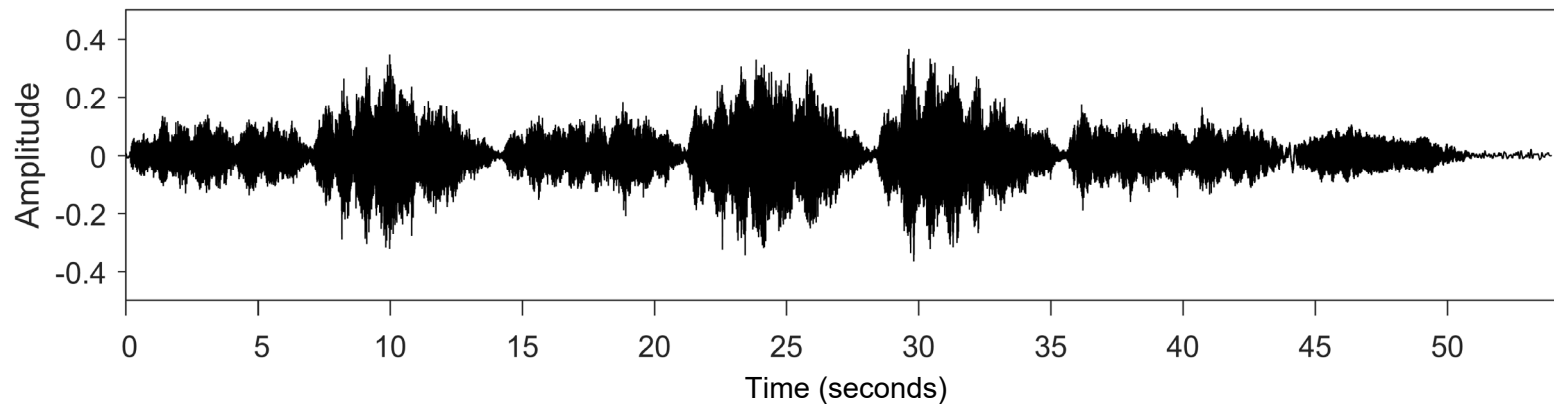
Durch dein Ge-fäng-nis, Got - tes Sohn, muß uns die Frei - heit kom - men;  
Dein Ker-ker ist der Gna - den - thron, die Frei-statt al - ler From - men;

9  
Denn gingst du nicht die Knecht schaft ein, müßt uns - re Knecht-schaft e - wig sein.

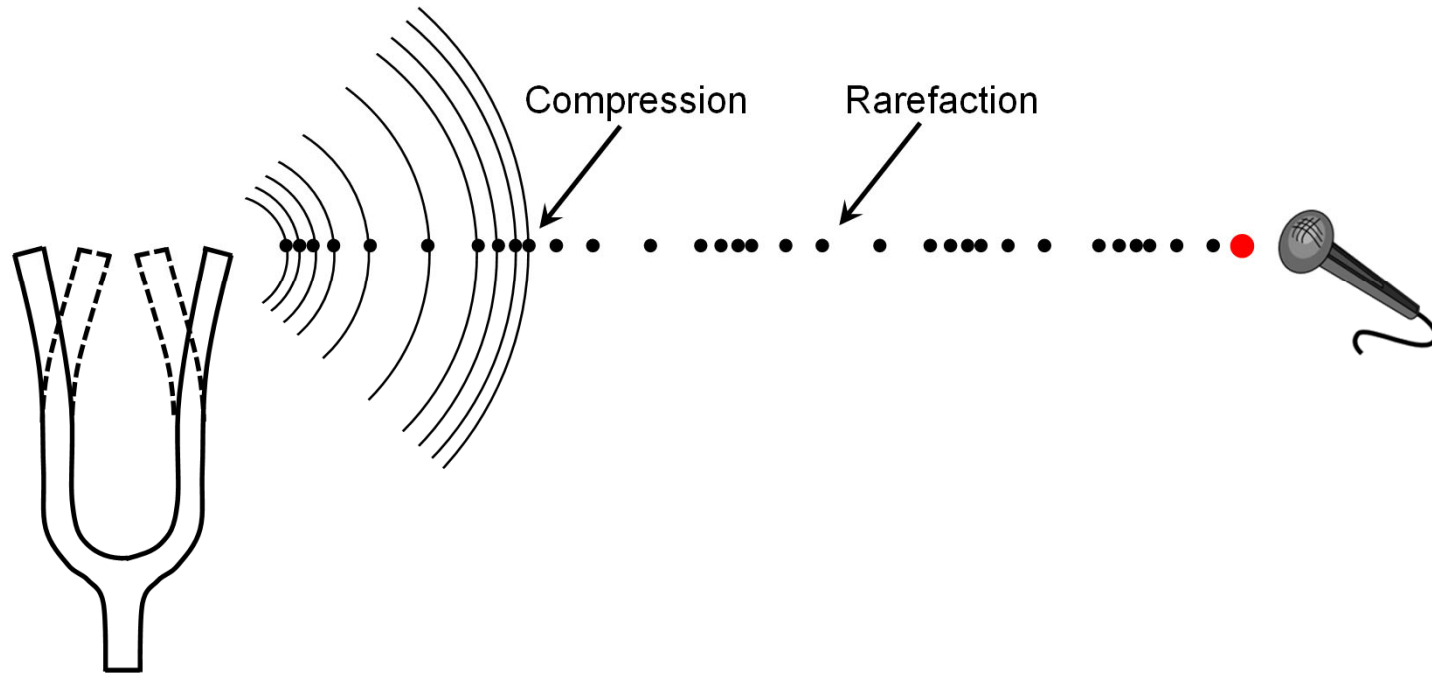
The image shows a piano reduction of a choral piece by J.S. Bach. It consists of two systems of music. Each system has a treble and bass staff. The first system contains the first two lines of the score, with lyrics in German. The second system starts at measure 9 and contains the next two lines of the score, also with lyrics. The music is in G major (one sharp) and 4/4 time. The piano part features a steady accompaniment with some melodic lines in the right hand.

# Visualization of Diatonic Scales

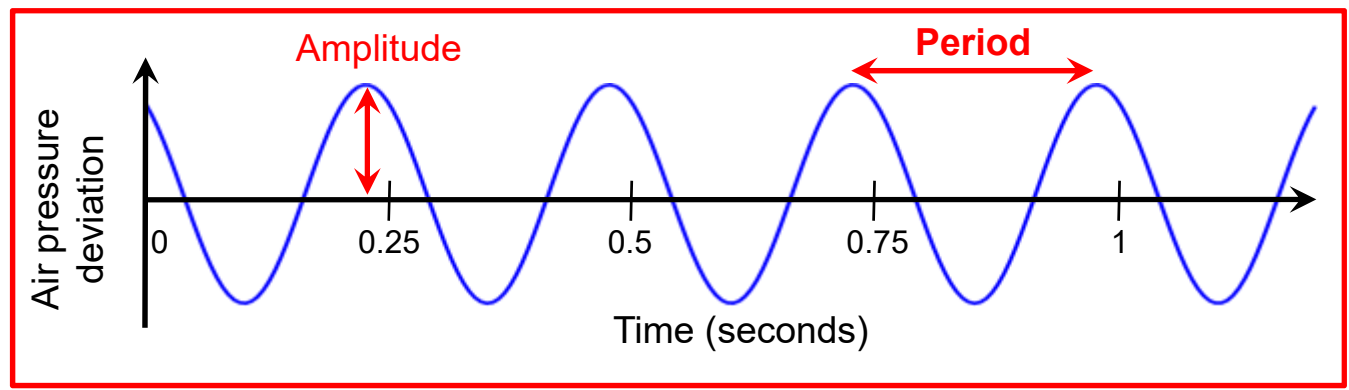
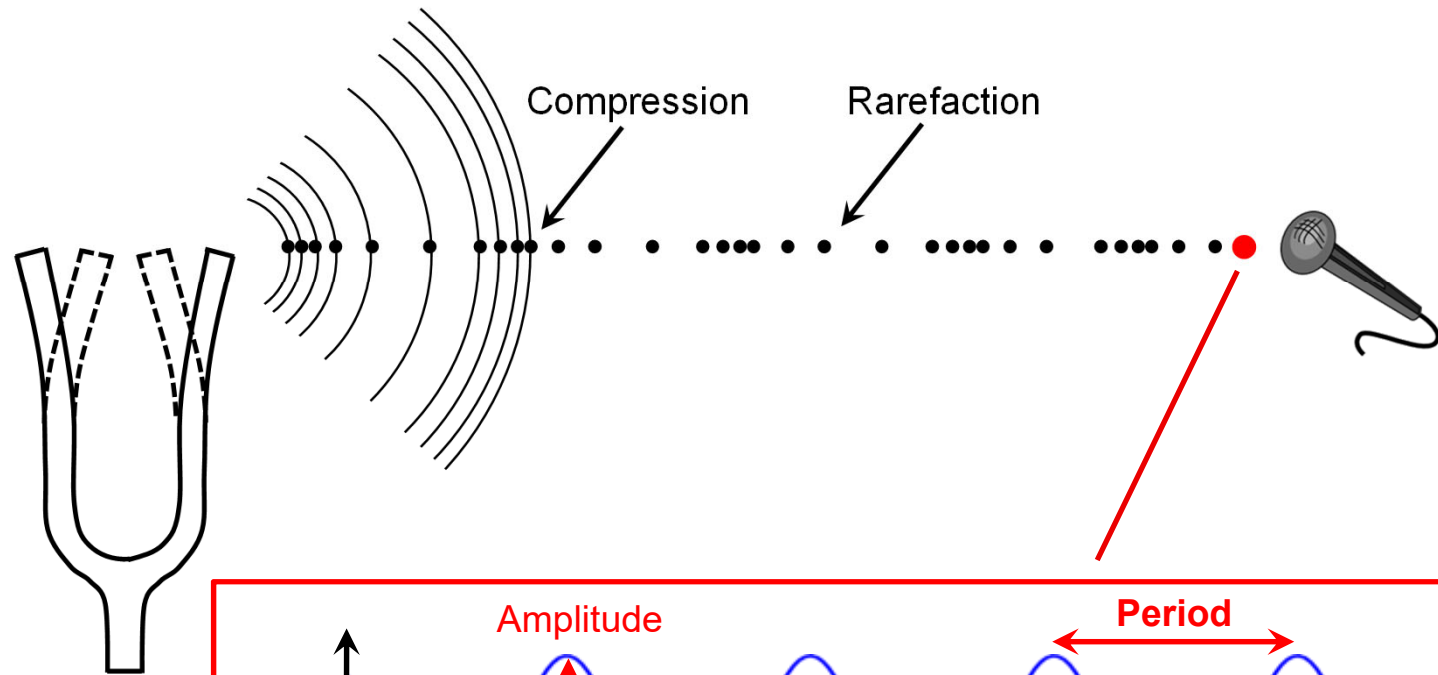
- Example: J.S. Bach, Choral "Durch Dein Gefängnis" (*Johannespassion*)
- **Audio** – Waveform (Scholars Baroque Ensemble, Naxos 1994)



# Excursus: Waveform



# Excursus: Waveform



Frequency:  $1/\text{Period}$

Unit: Hertz (Oscillations per second)

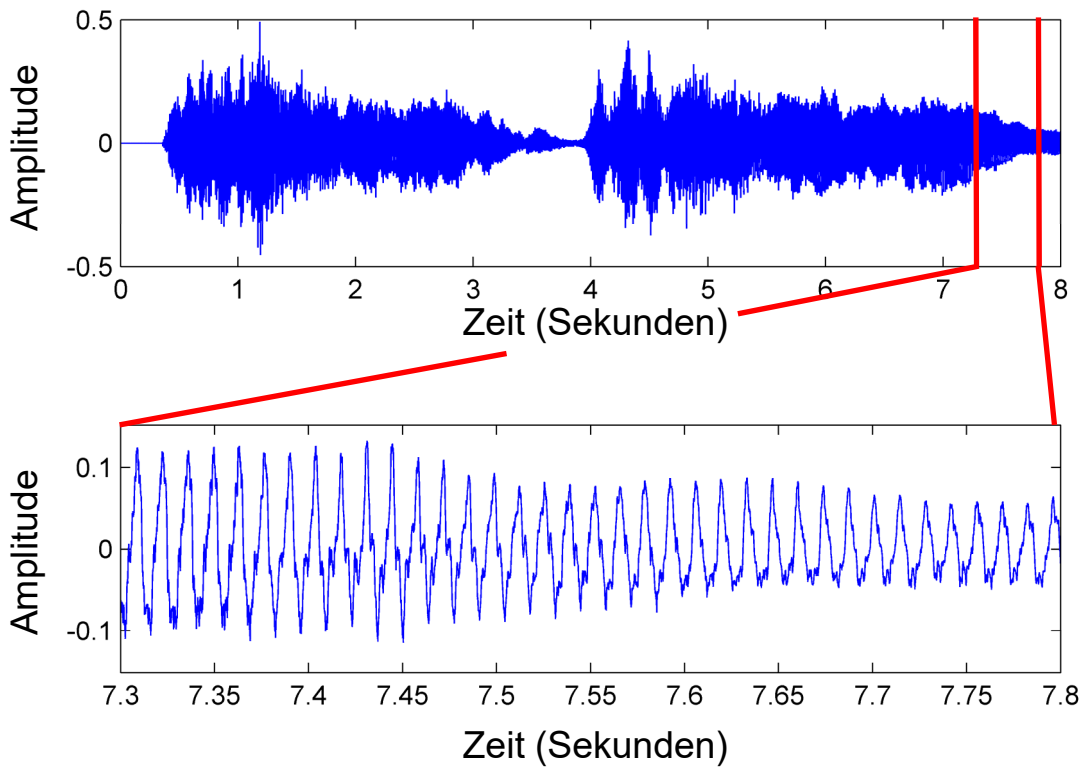
# Excursus: Waveform



Allegro con brio ( $\text{♩} = 108$ )

*ff*

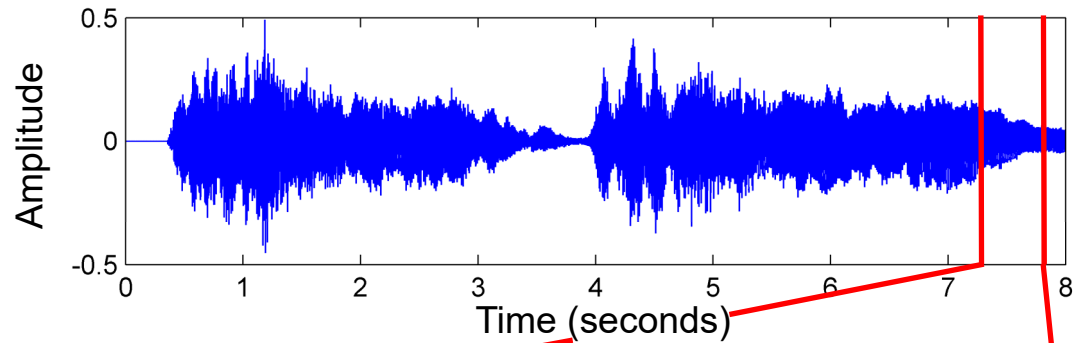
*ped.* \*



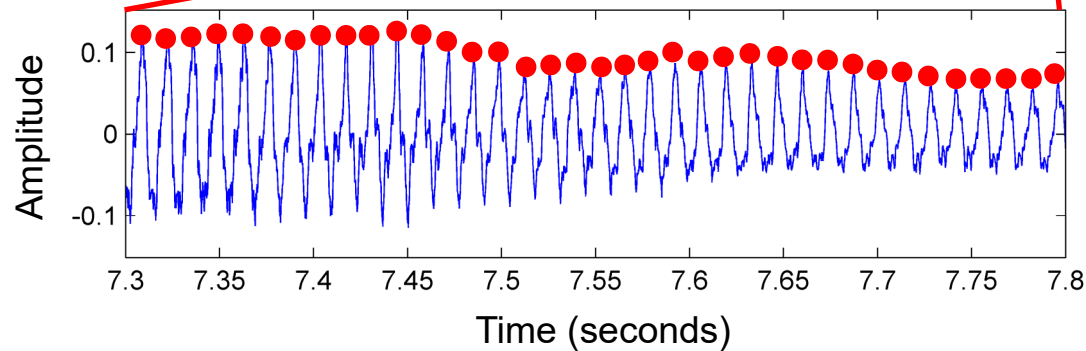
D2 (73.4 Hz)



# Excursus: Waveform



D2 (73.4 Hz)



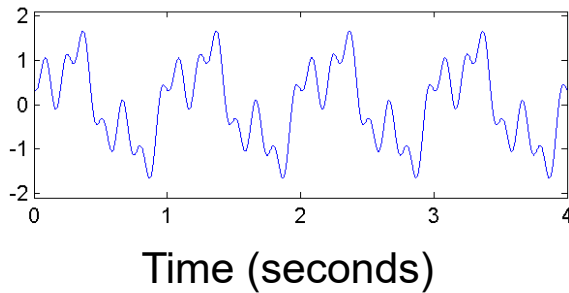
37 Periods  
in 500 ms

# Excursus: Fourier Transform

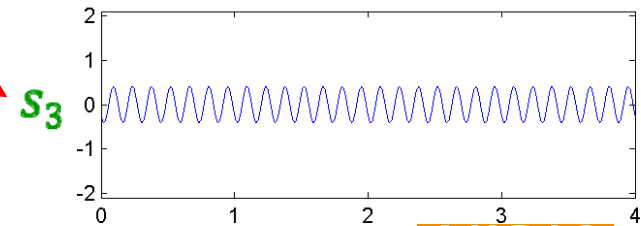
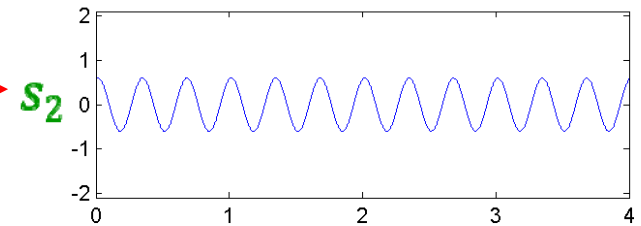
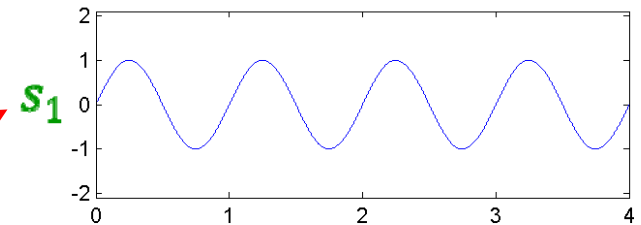
Idea: **Decompose** a given **signal** into a superposition of **sinusoids** (elementary signals).

$$f = s_1 + s_2 + s_3$$

Signal  $f$



Sinusoids



# Excursus: Fourier Transform

Each **sinusoid** has a physical meaning and can be described by three parameters:

$$s(A, \omega, \varphi)(t) = A \cdot \sin(2\pi(\omega t - \varphi))$$

$\omega$  = frequency

$A$  = amplitude

$\varphi$  = phase

## Interpretation:

The amplitude  $A$  reflects the intensity at which the sinusoidal of frequency  $\omega$  appears in  $f$ .

The phase  $\varphi$  reflects how the sinusoidal has to be shifted to best correlate with  $f$ .

$$A_1 = 1$$

$$\omega_1 = 1$$

$$\varphi_1 = 0$$

$$A_2 = 0.6$$

$$\omega_2 = 3$$

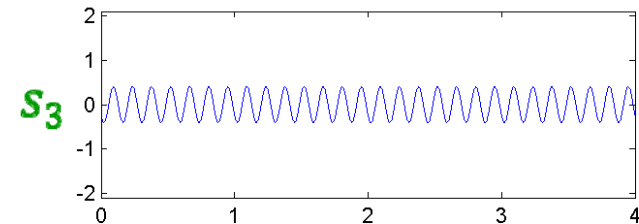
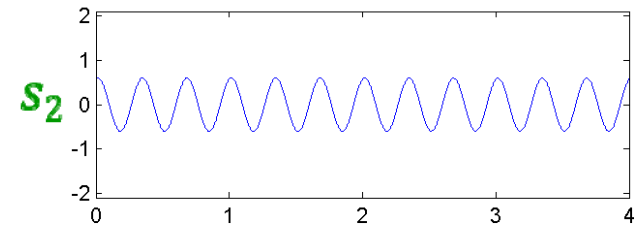
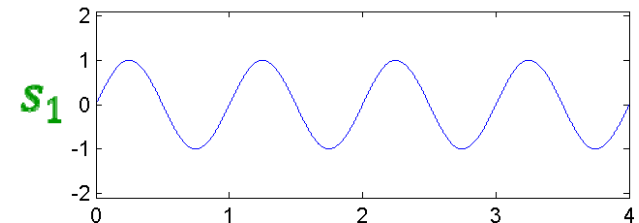
$$\varphi_2 = -0.2$$

$$A_3 = 0.4$$

$$\omega_3 = 7$$

$$\varphi_3 = 0.4$$

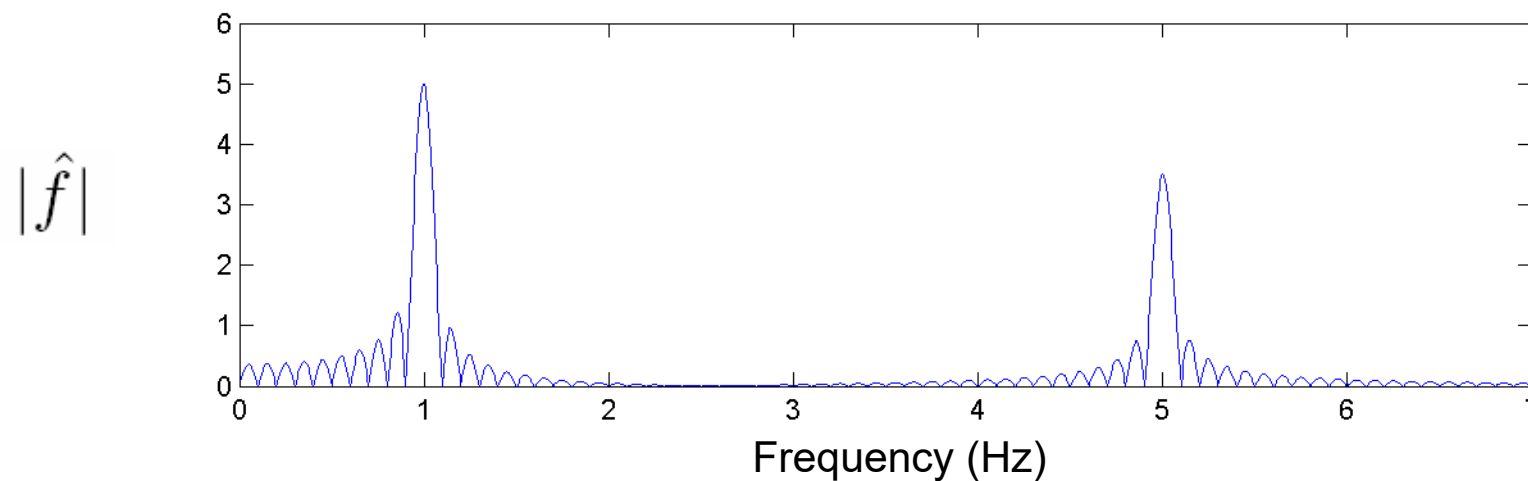
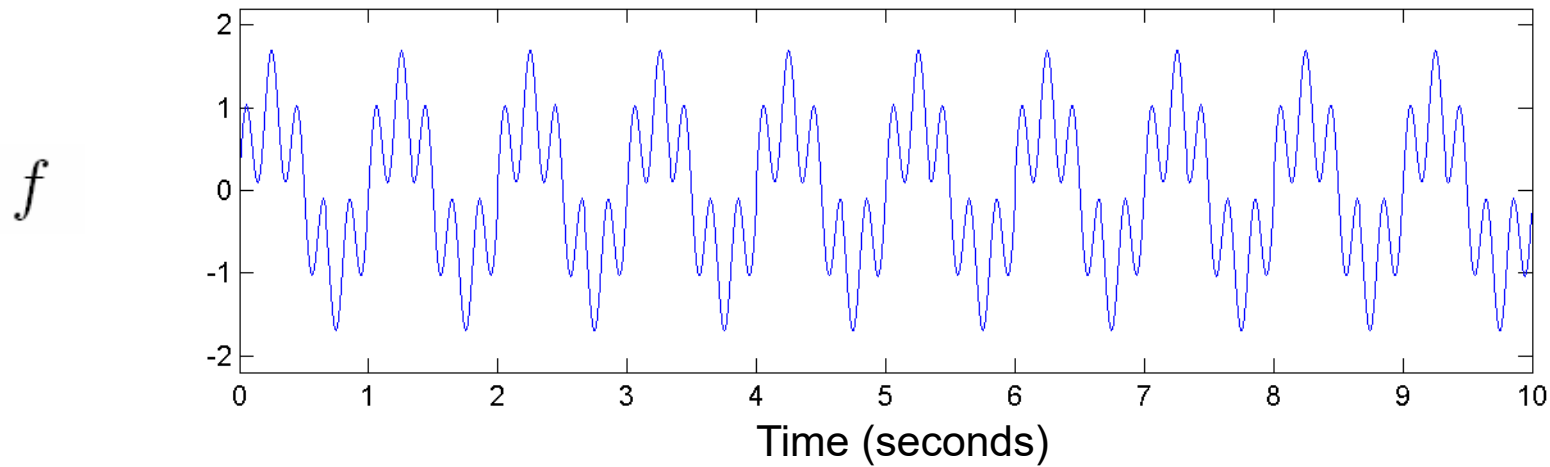
## Sinusoids



Time (seconds)

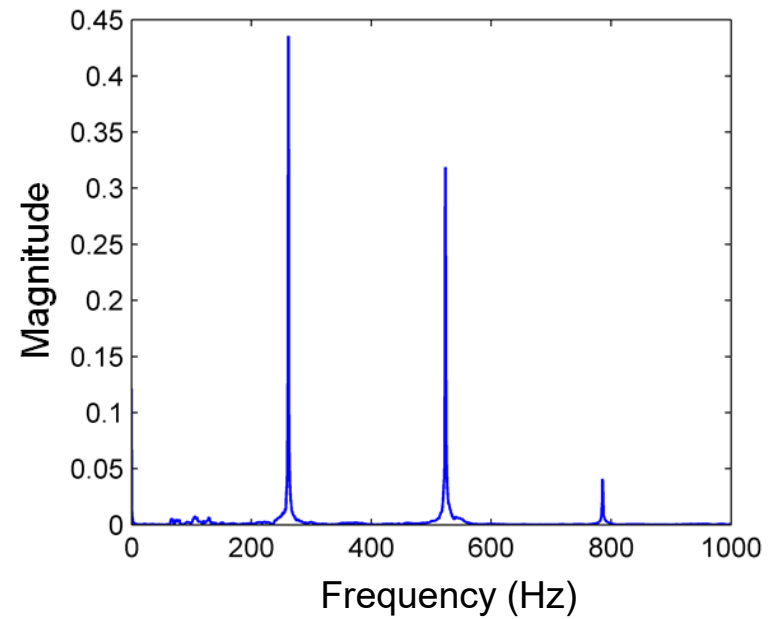
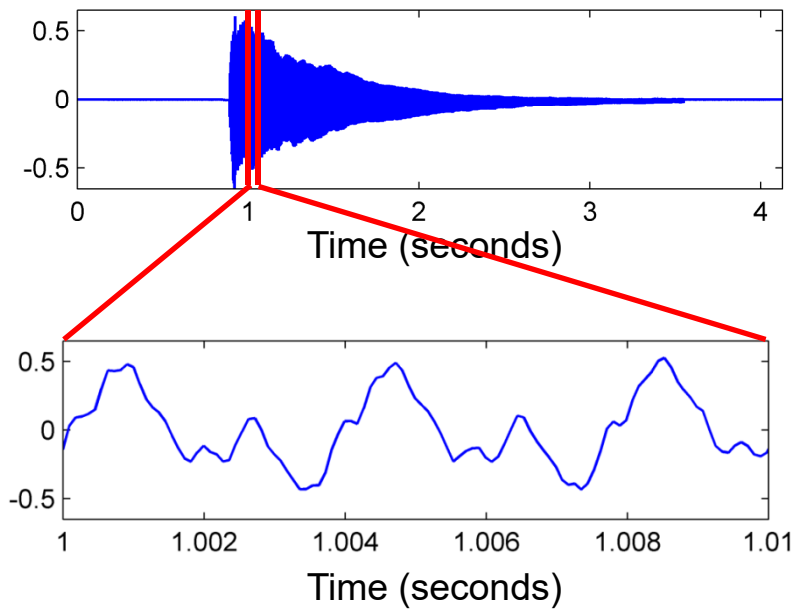
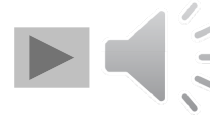
# Excursus: Fourier Transform

Example: Superposition of two sinusoids



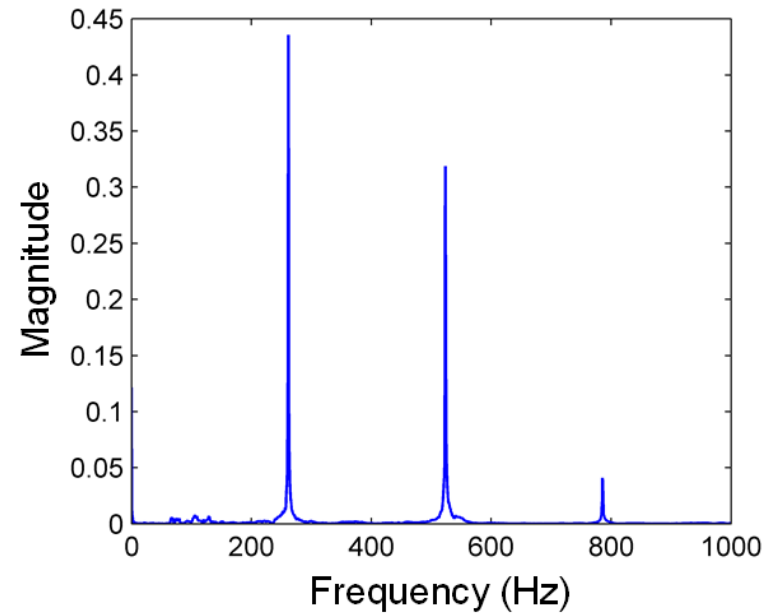
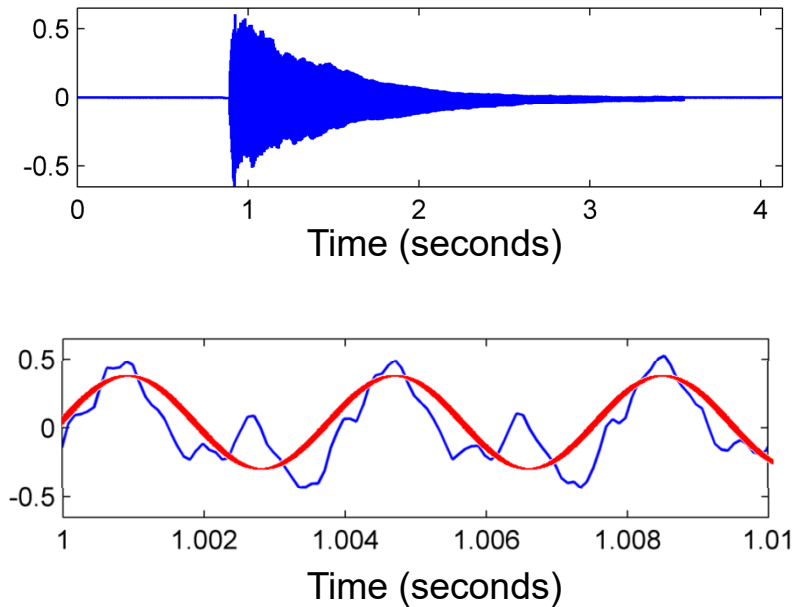
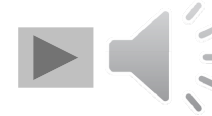
# Excursus: Fourier Transform

Example: Piano tone (C4, 261.6 Hz)



# Excursus: Fourier Transform

Example: Piano tone (C4, 261.6 Hz)



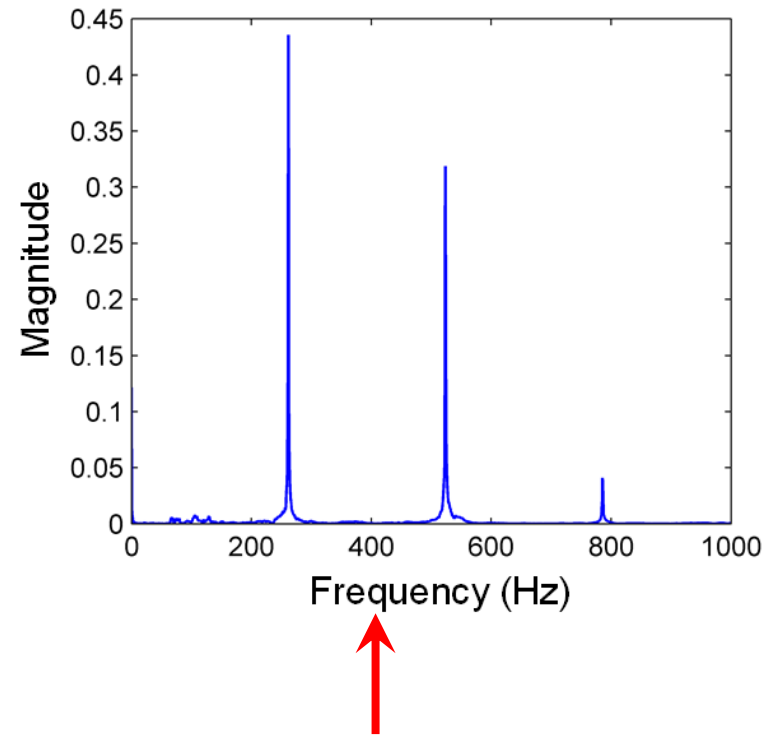
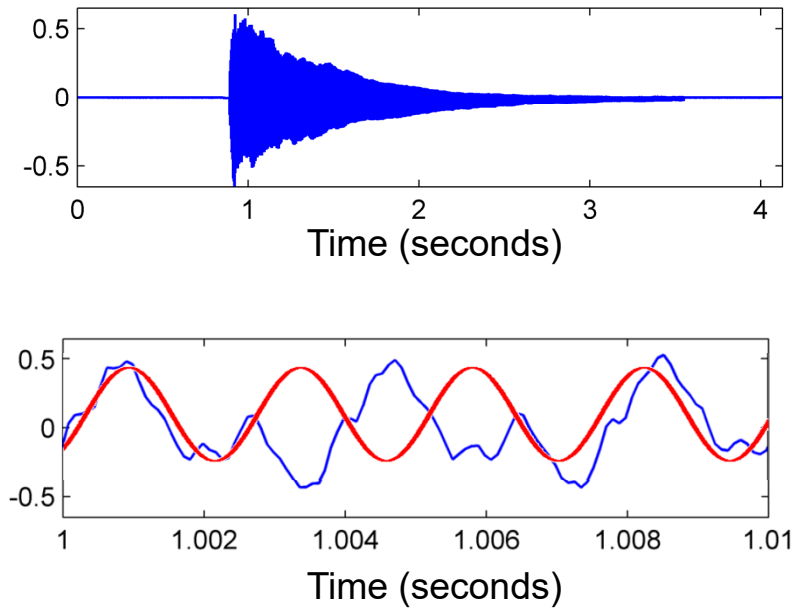
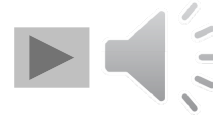
Analysis using sinusoid with **262 Hz**

→ high correlation

→ large Fourier coefficient

# Excursus: Fourier Transform

Example: Piano tone (C4, 261.6 Hz)



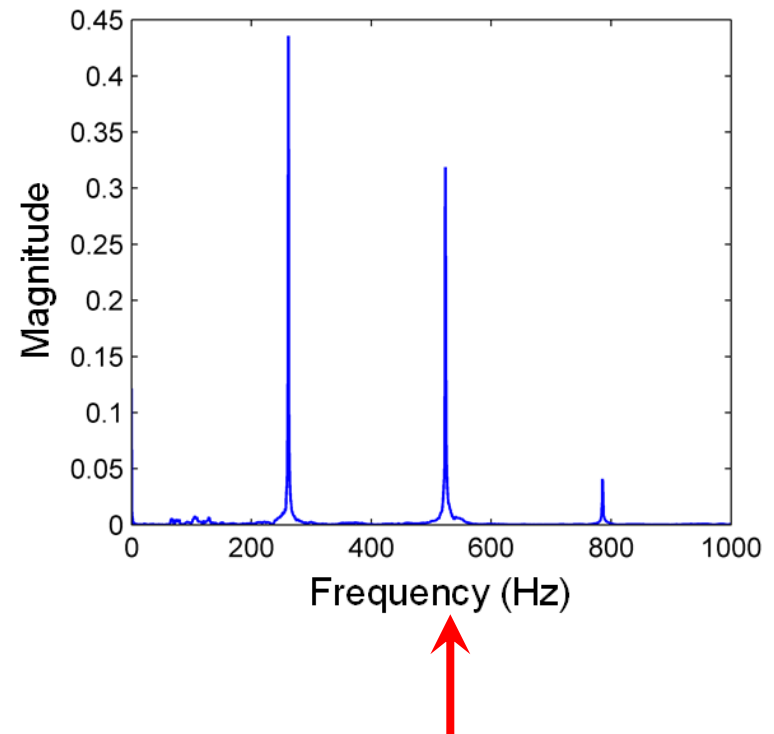
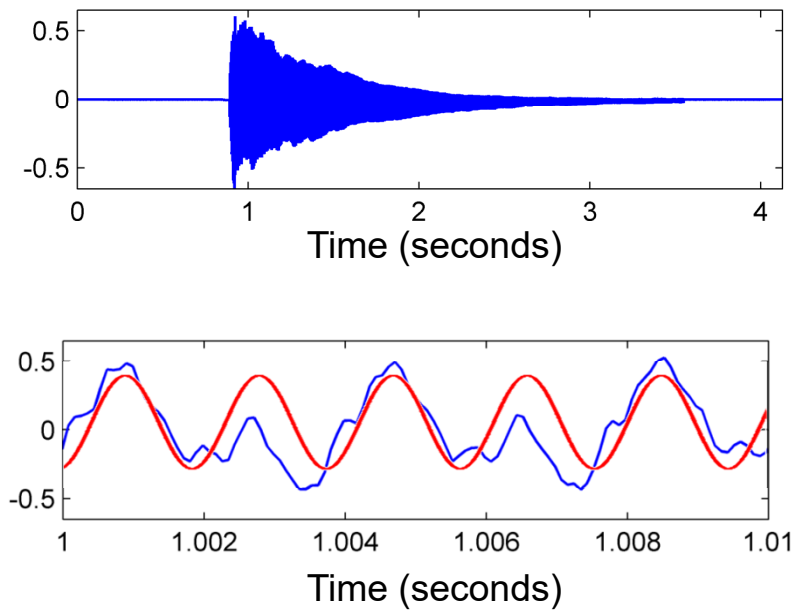
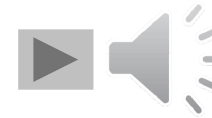
Analysis using sinusoid with **400 Hz**

→ low correlation

→ small Fourier coefficient

# Excursus: Fourier Transform

Example: Piano tone (C4, 261.6 Hz)



Analysis using sinusoid with **523 Hz**

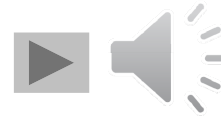
→ high correlation

→ large Fourier coefficient

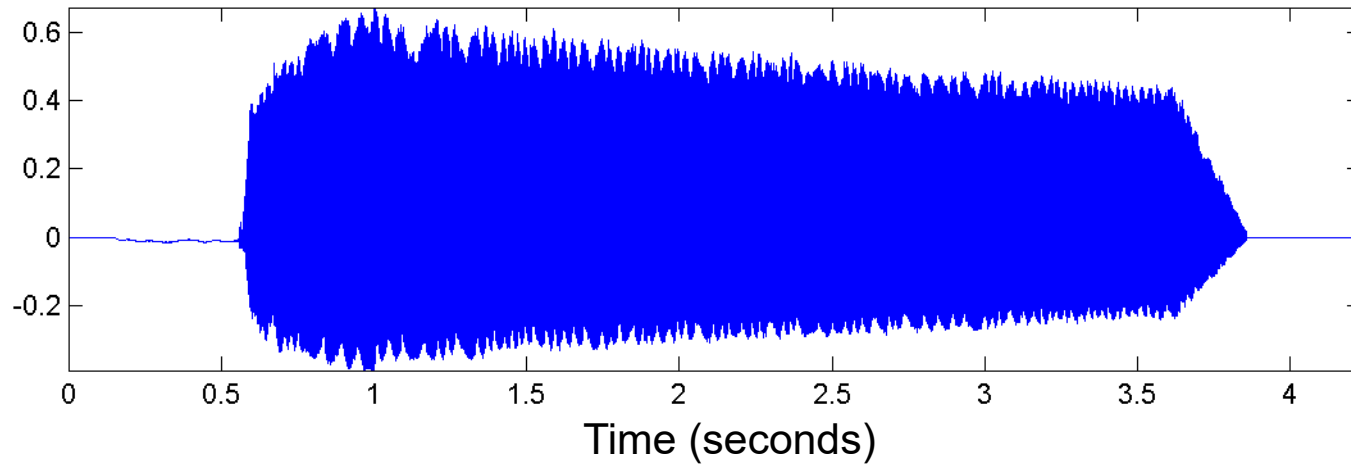


# Excursus: Fourier Transform

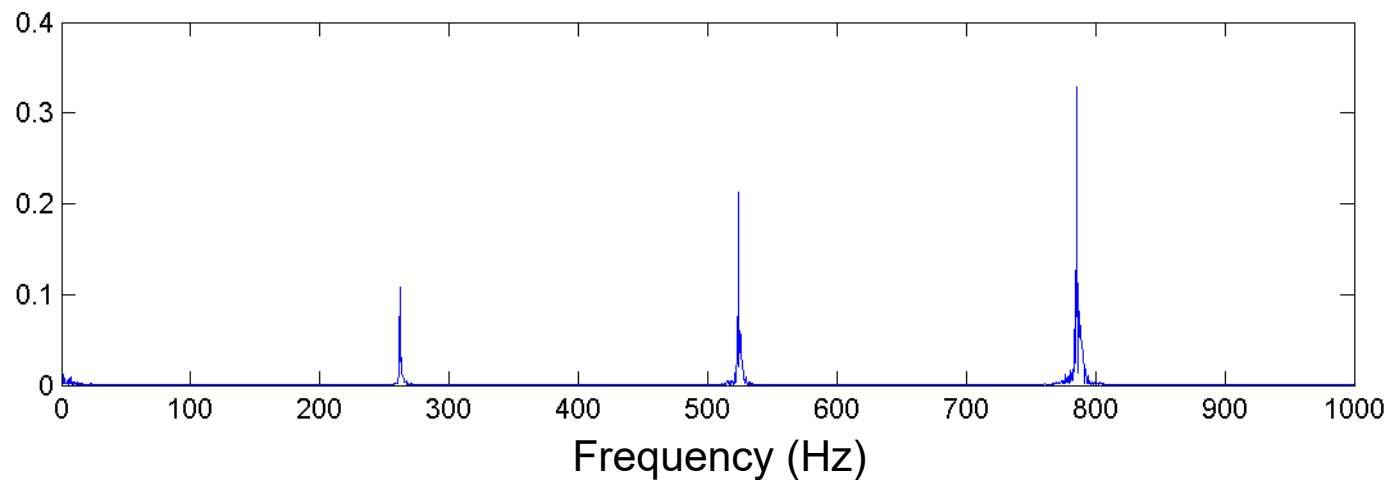
Example: C4 played by trumpet



$f$

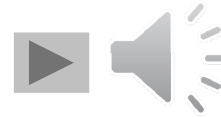


$|\hat{f}|$

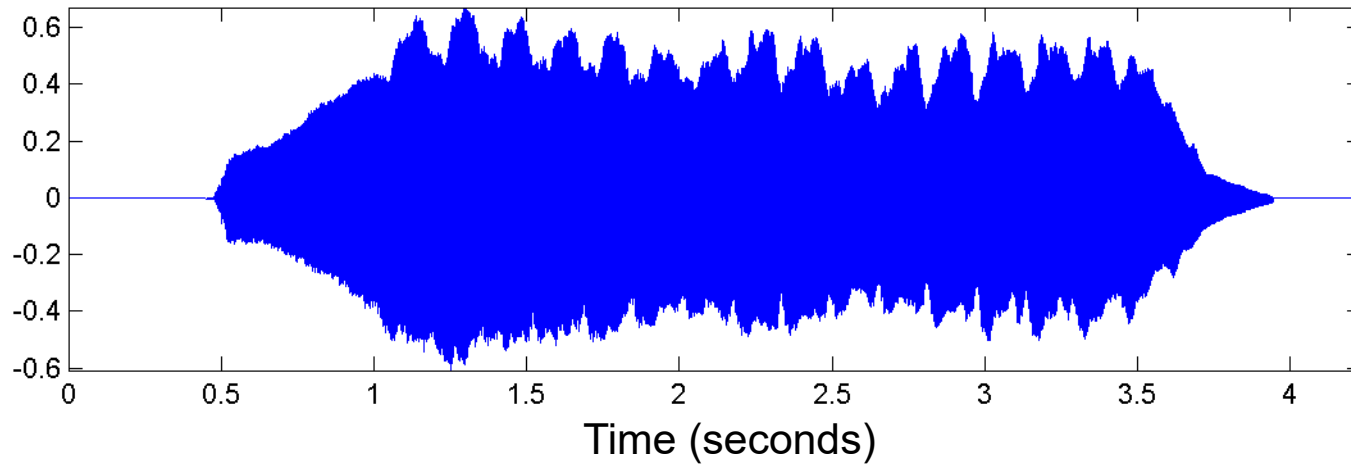


# Excursus: Fourier Transform

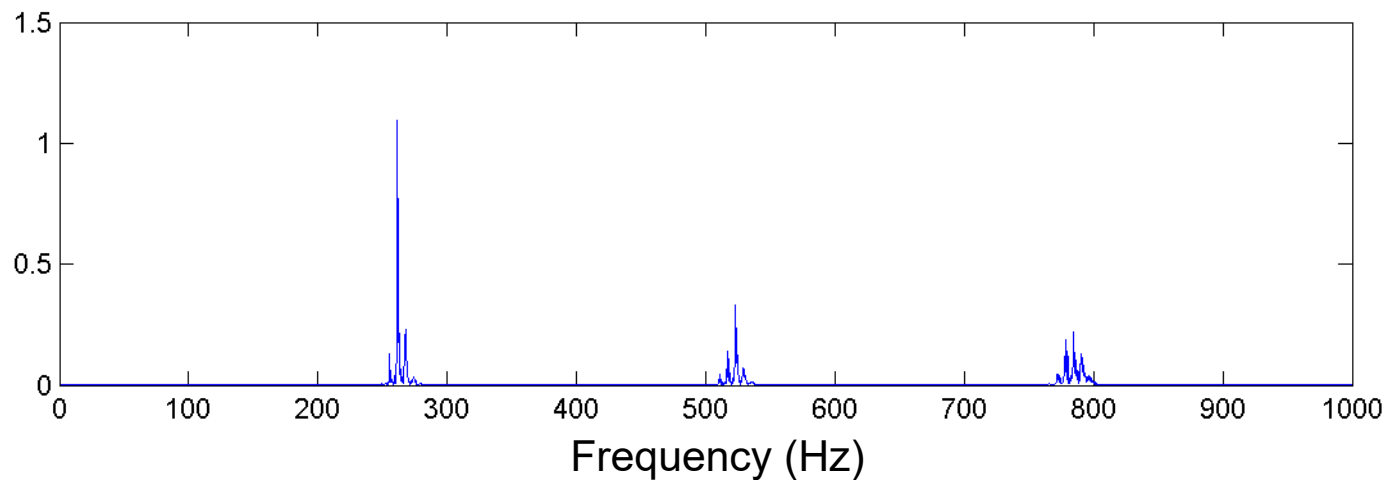
Example: C4 played by violine



$f$

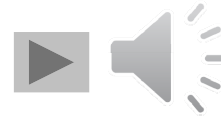


$|\hat{f}|$

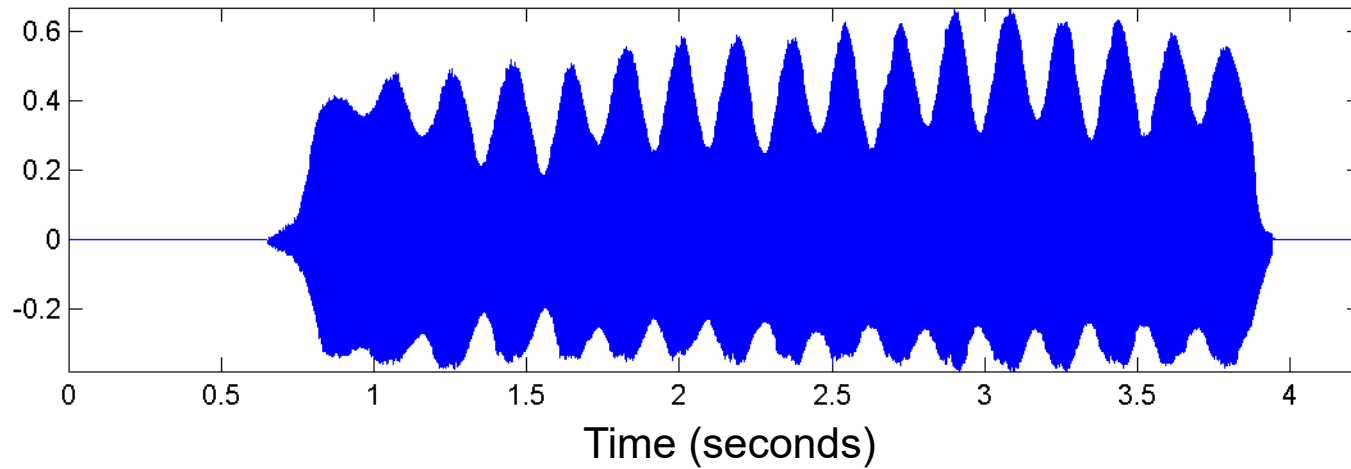


# Excursus: Fourier Transform

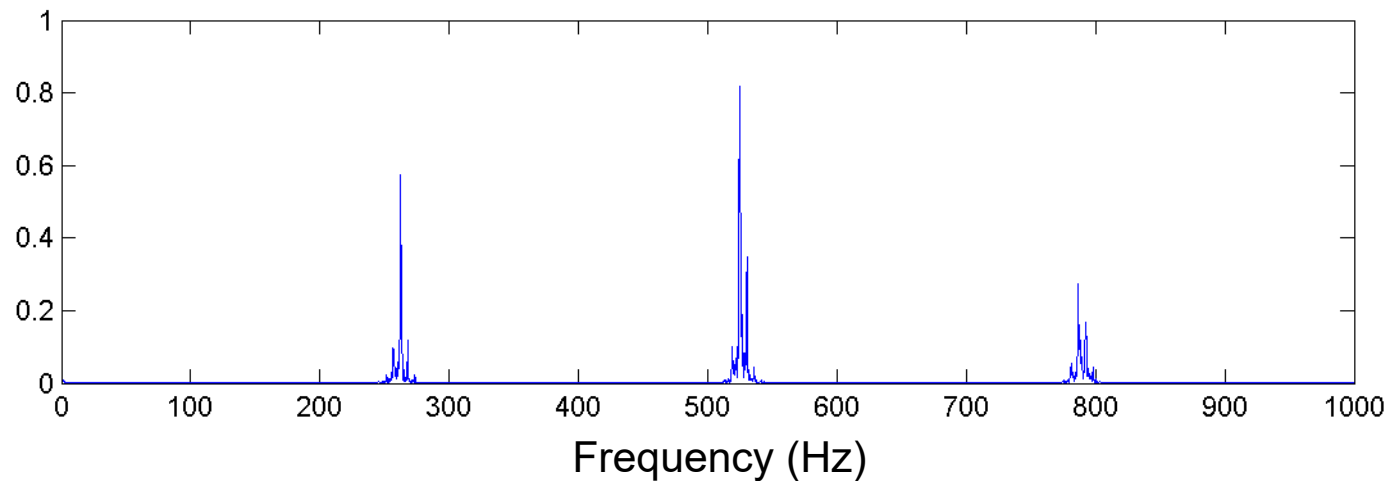
Example: C4 played by flute



$f$



$|\hat{f}|$



## Excursus: Fourier Transform

Signal

$$f: \mathbb{R} \rightarrow \mathbb{R}$$

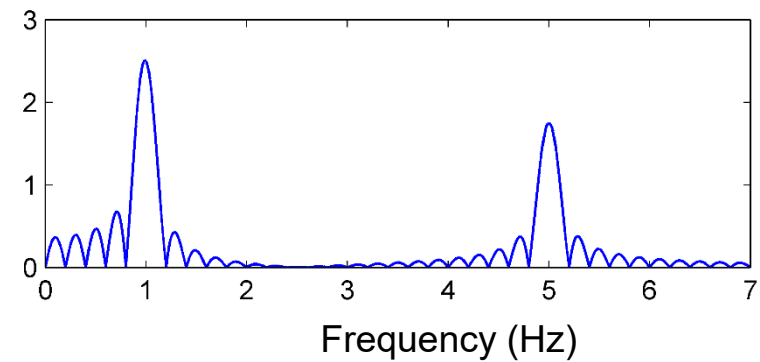
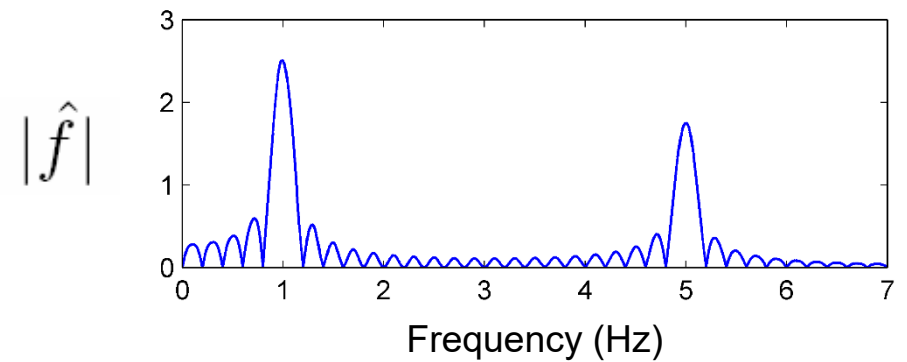
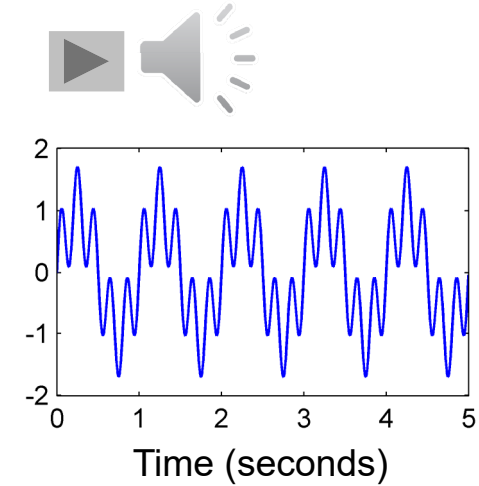
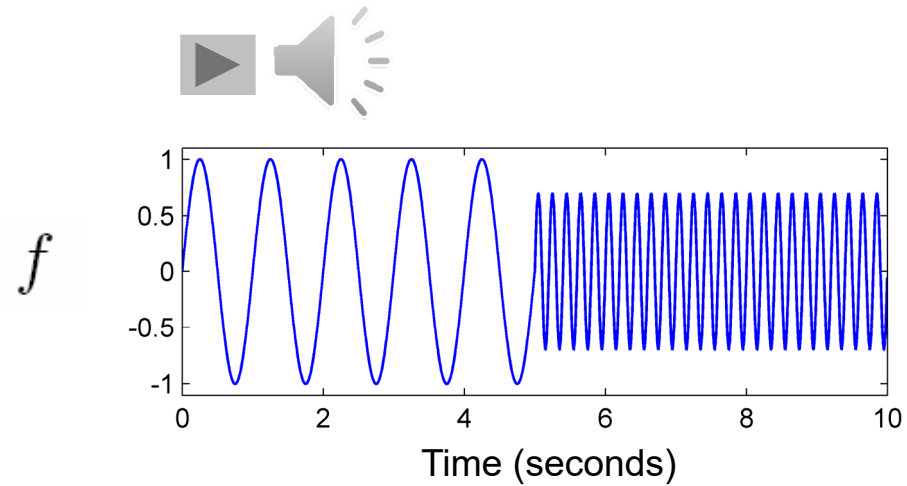
Fourier representation

$$f(t) = \int_{\omega \in \mathbb{R}} c_{\omega} \exp(2\pi i \omega t) d\omega$$

Fourier transform

$$c_{\omega} = \hat{f}(\omega) = \int_{t \in \mathbb{R}} f(t) \exp(-2\pi i \omega t) dt$$

# Excursus: Fourier Transform

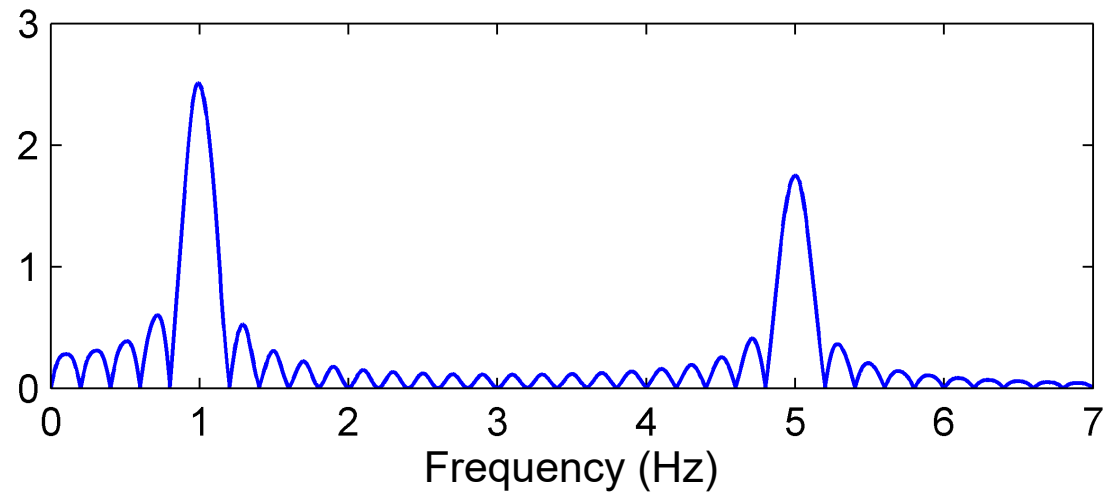
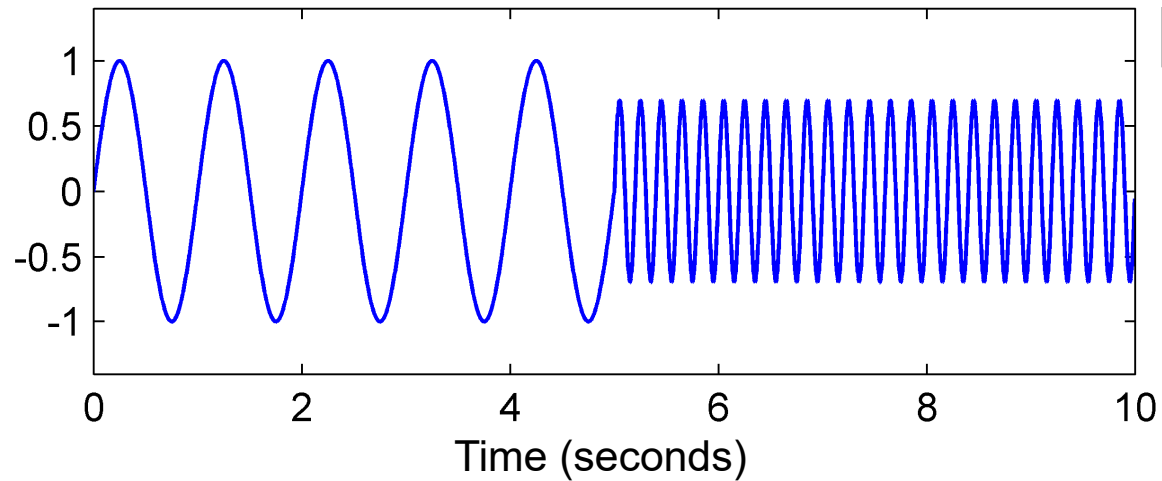


# Excursus: Short Time Fourier Transform

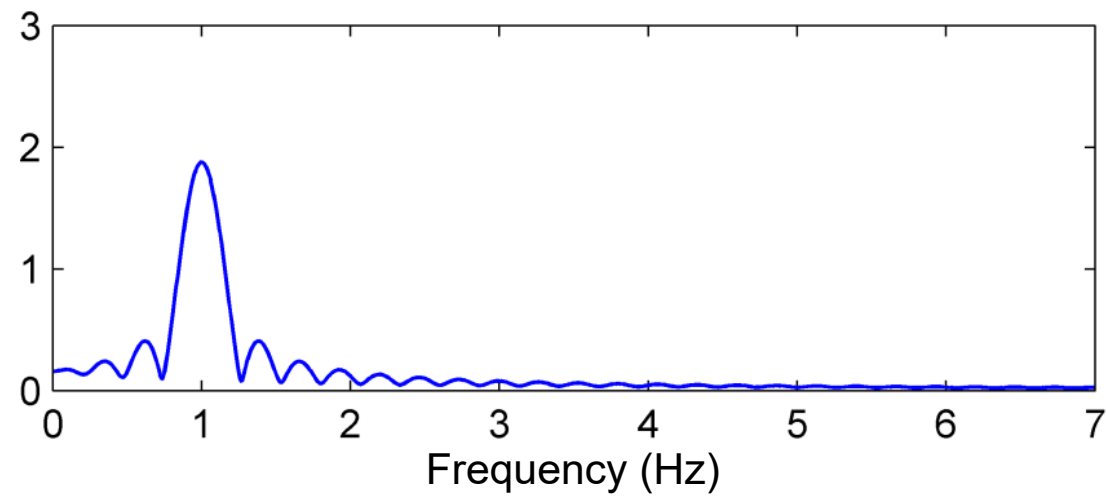
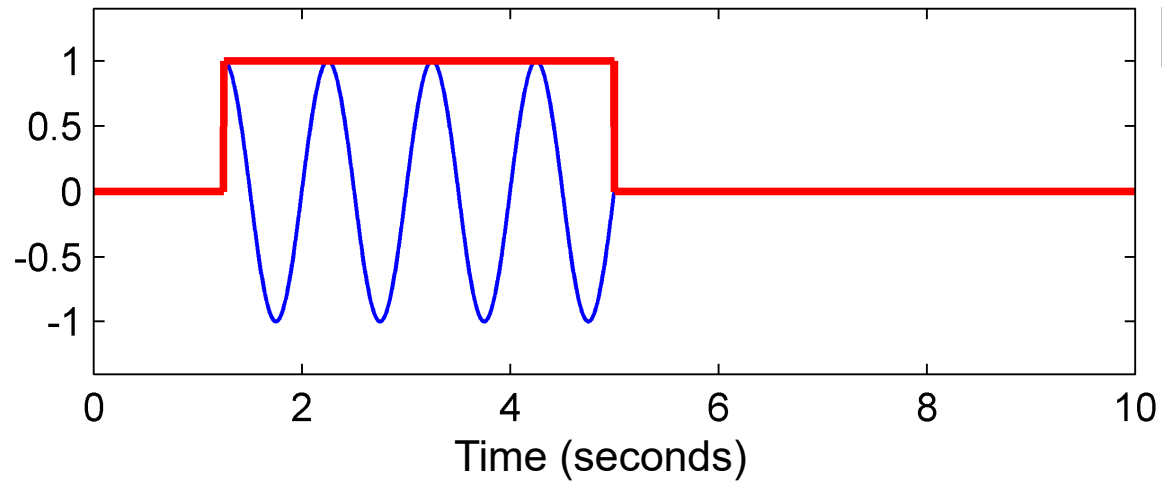
Idea (Dennis Gabor, 1946):

- Consider only a **small section** of the signal for the spectral analysis  
→ recovery of time information
- Short Time Fourier Transform (STFT)
- Section is determined by pointwise multiplication of the signal with a localizing **window function**

# Excursus: Short Time Fourier Transform

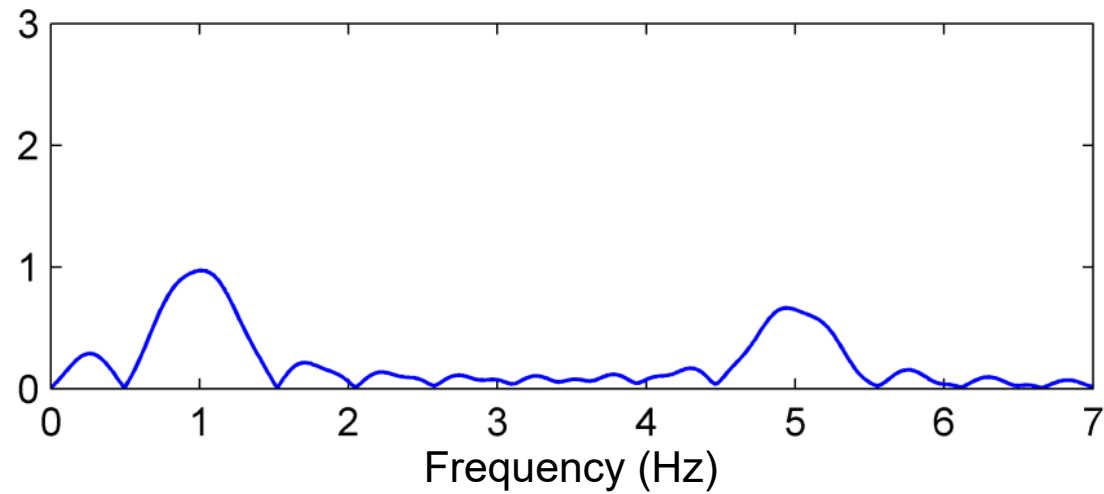
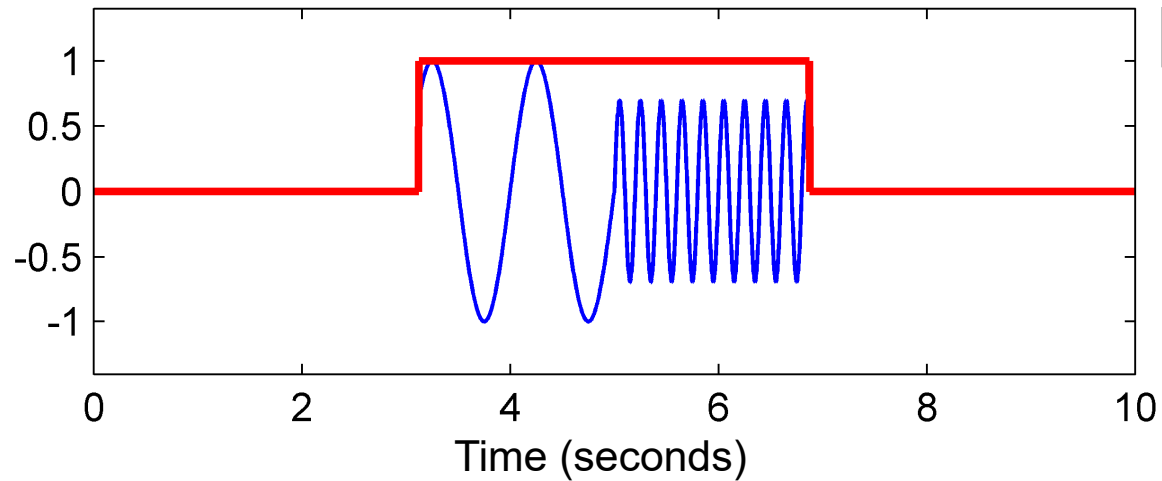


# Excursus: Short Time Fourier Transform

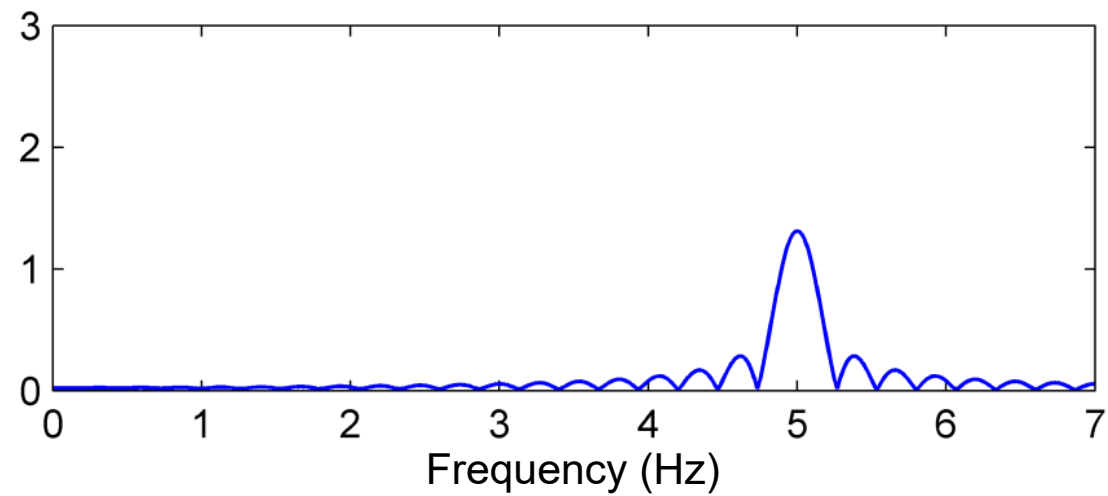
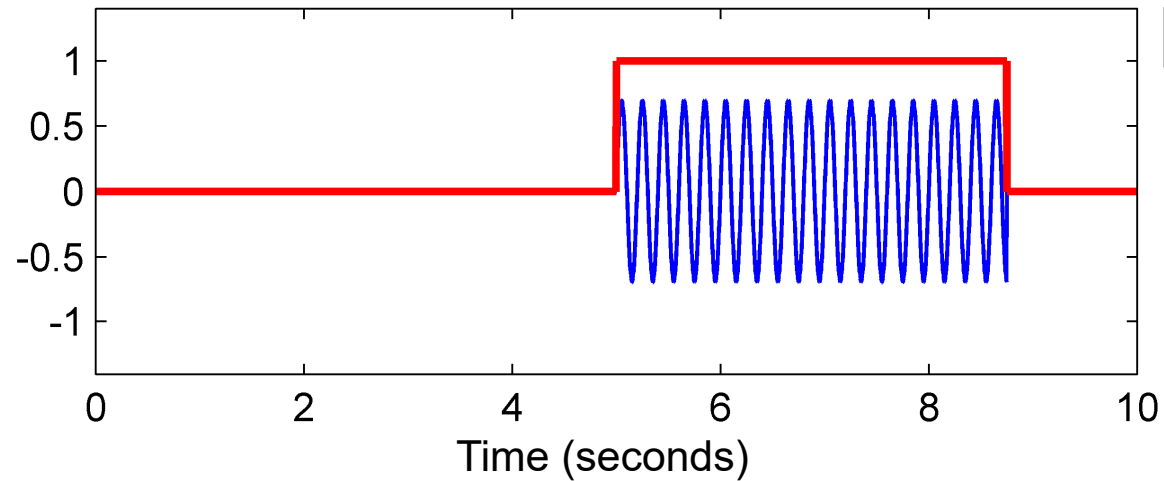




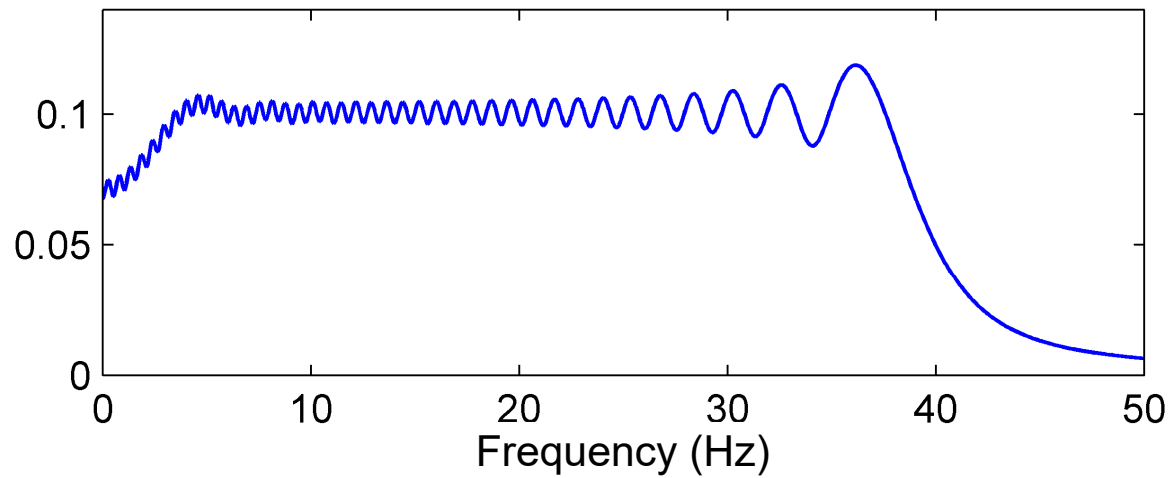
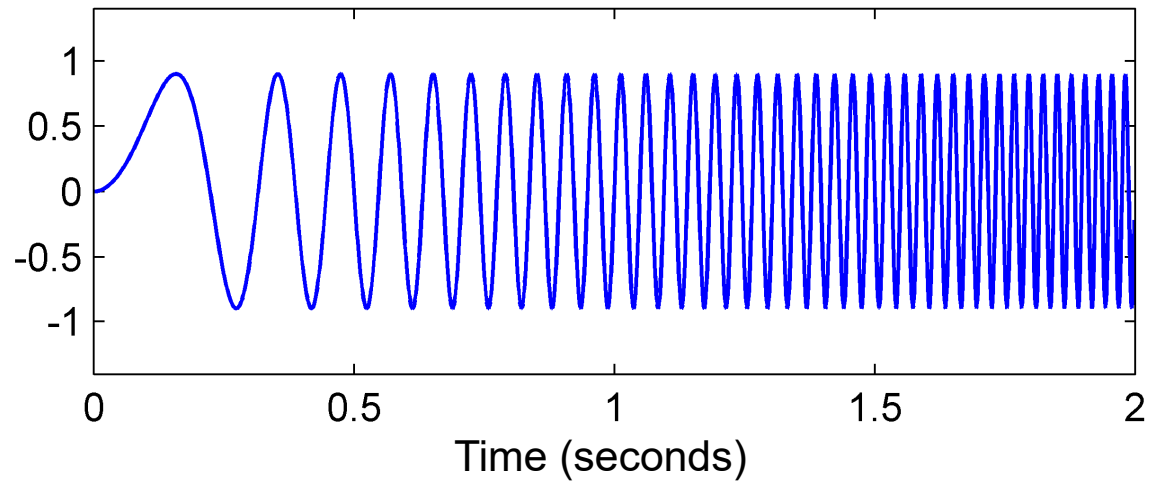
# Excursus: Short Time Fourier Transform



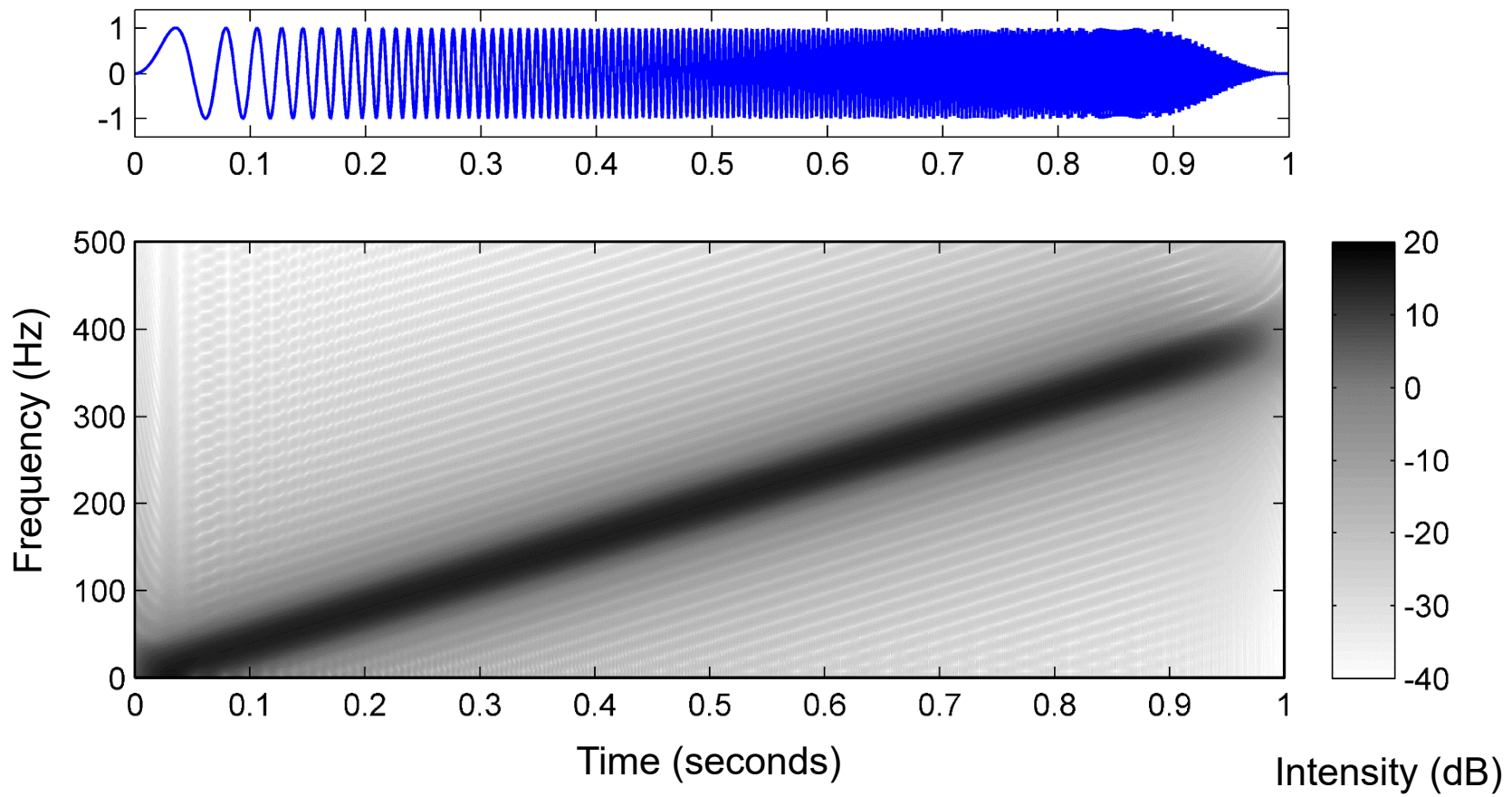
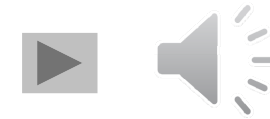
# Excursus: Short Time Fourier Transform



# Excursus: Short Time Fourier Transform

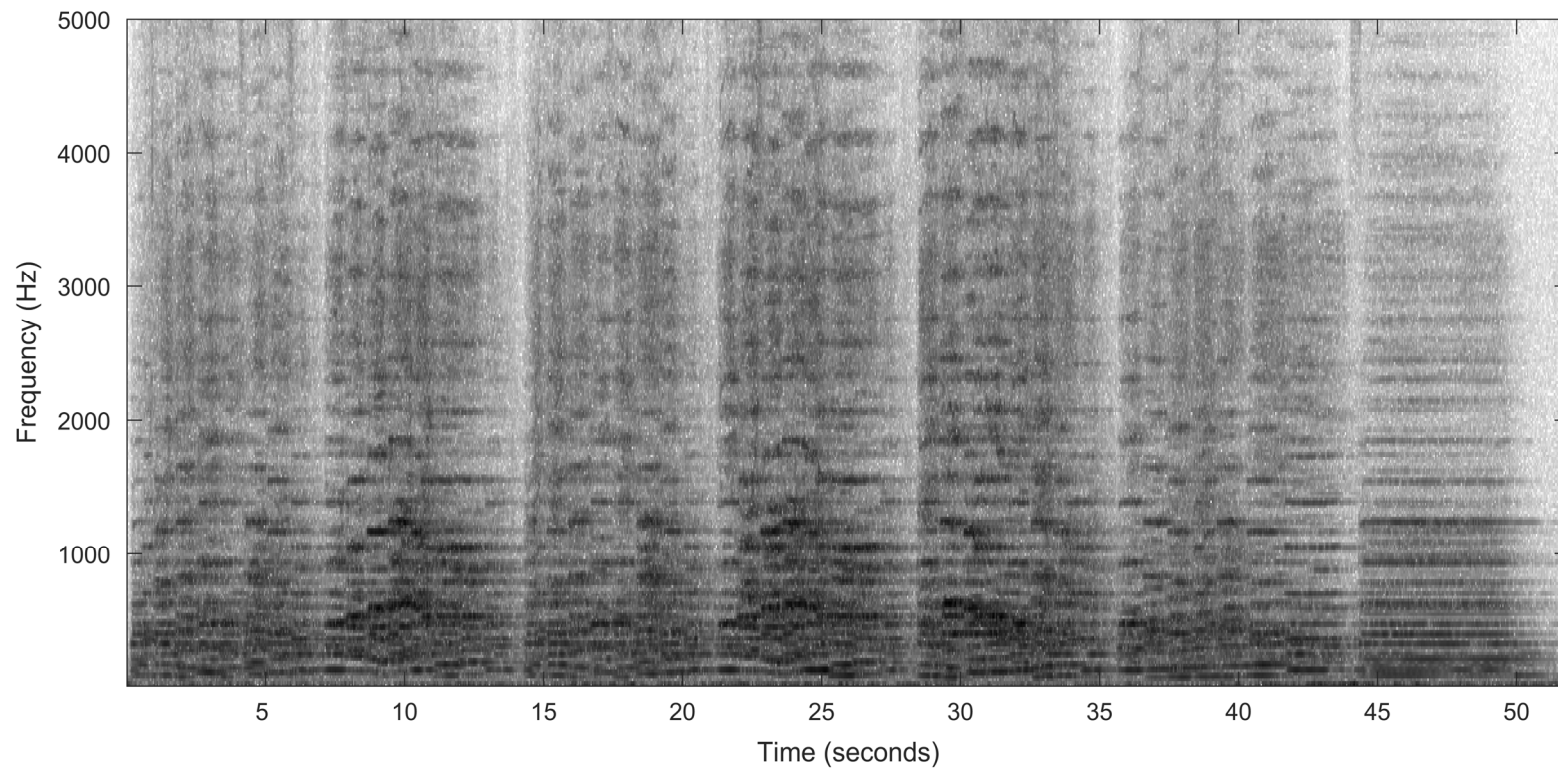


# Excursus: Spectrogram



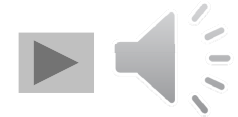
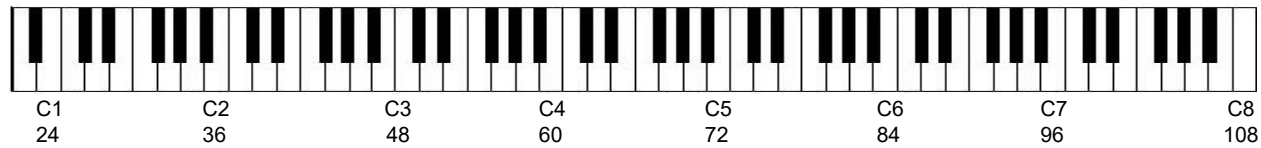
# Visualization of Diatonic Scales

- Example: J.S. Bach, Choral "Durch Dein Gefängnis" (*Johannespassion*)
- **Audio** – Spectrogram (Scholars Baroque Ensemble, Naxos 1994)

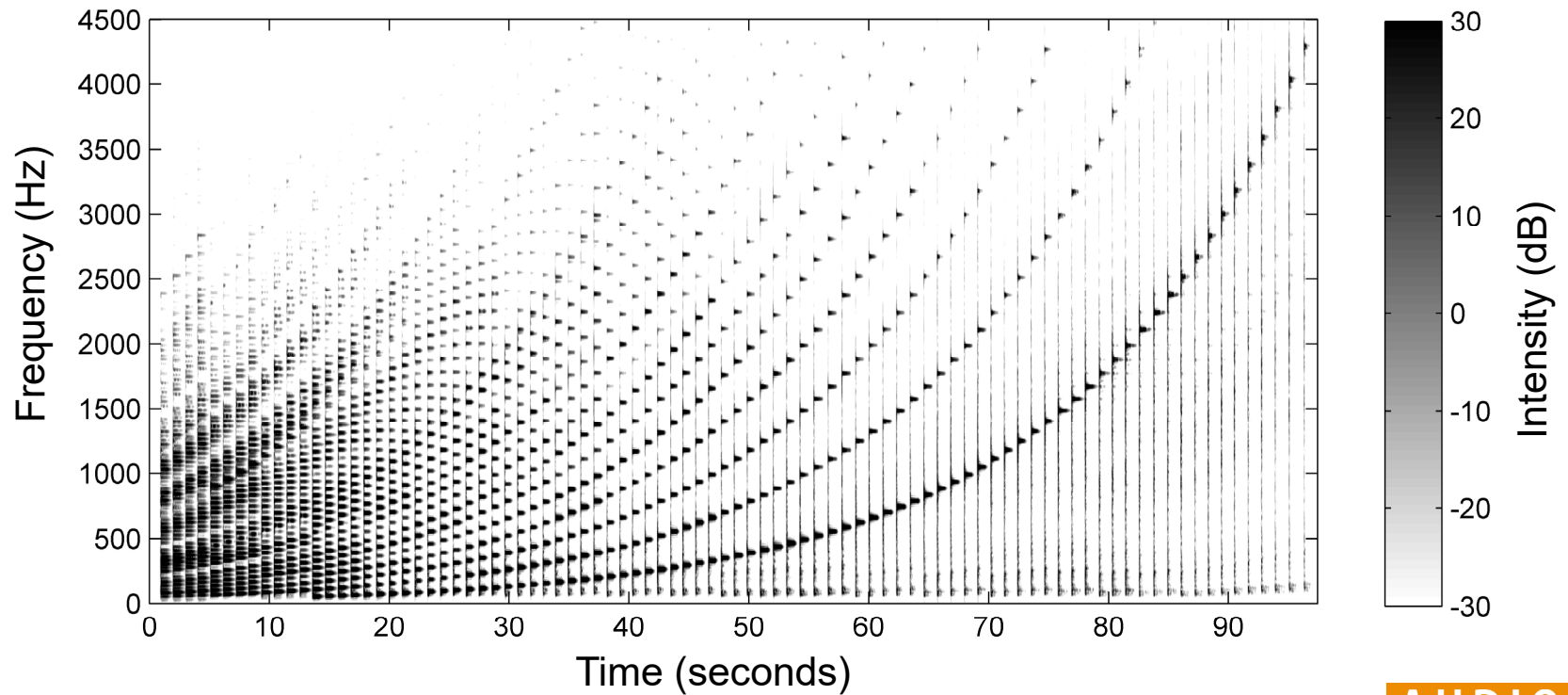


# Excursus: Spectrogram

Example: Chromatic scale

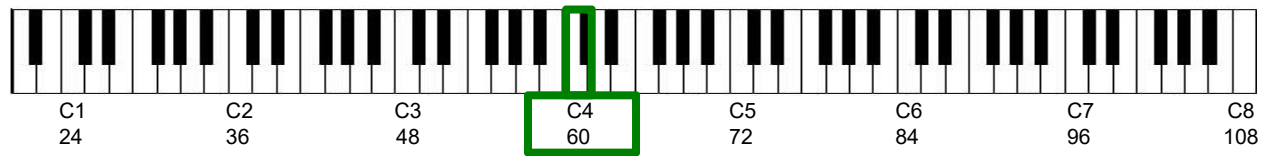


## Spectrogram

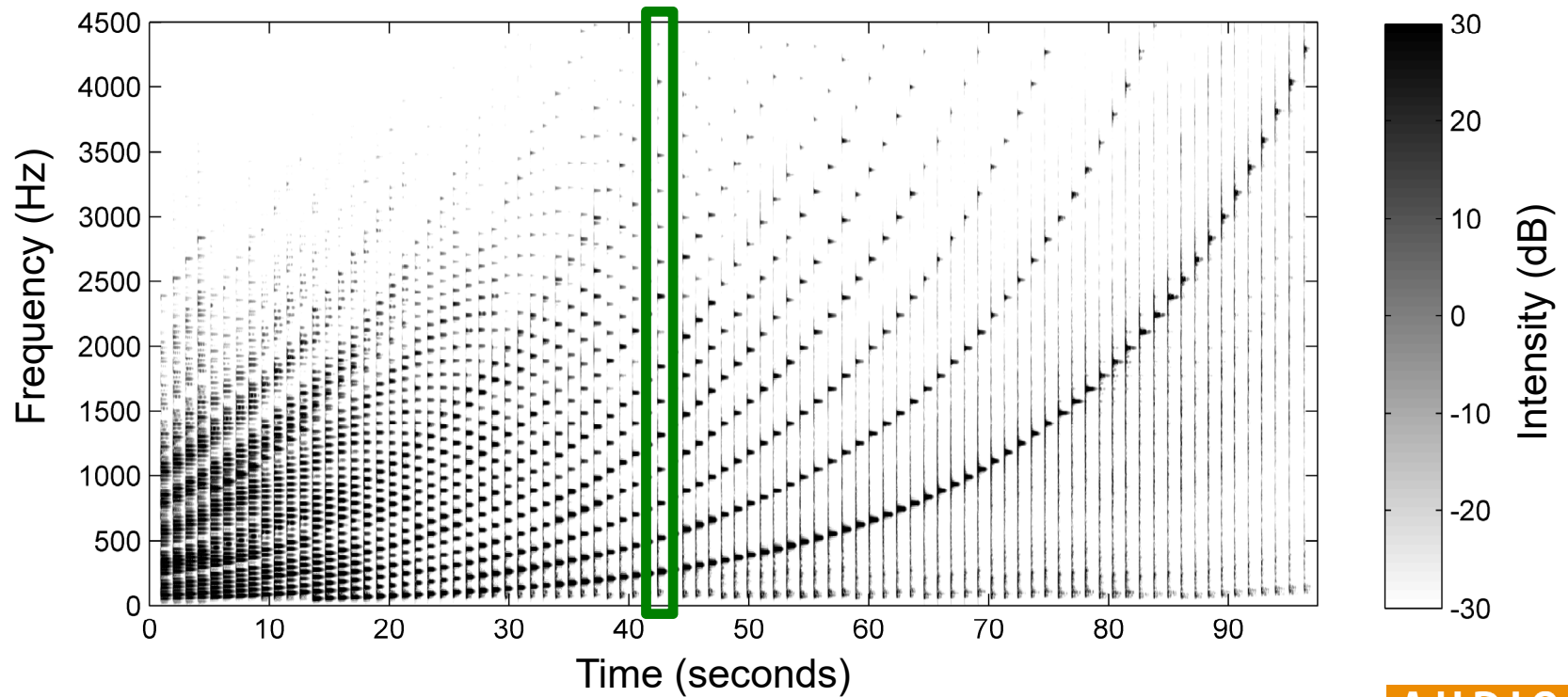


# Excursus: Spectrogram

Example: Chromatic scale

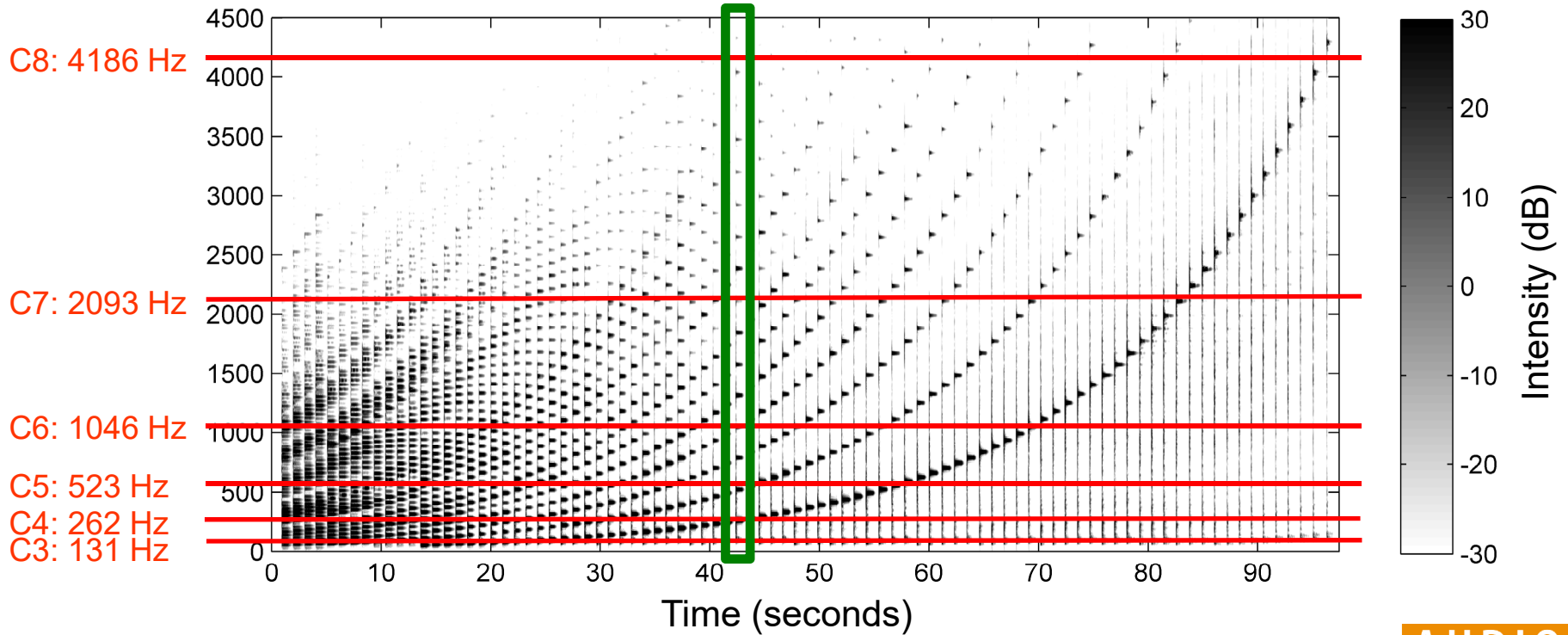
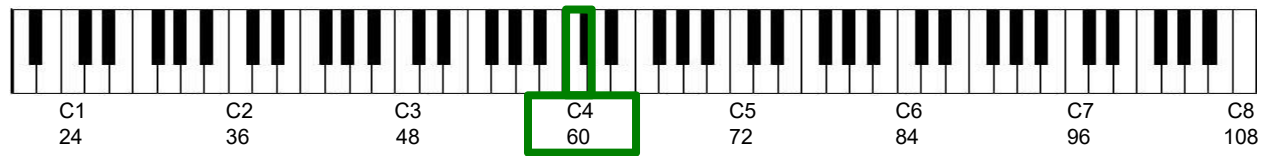


## Spectrogram



# Excursus: Spectrogram

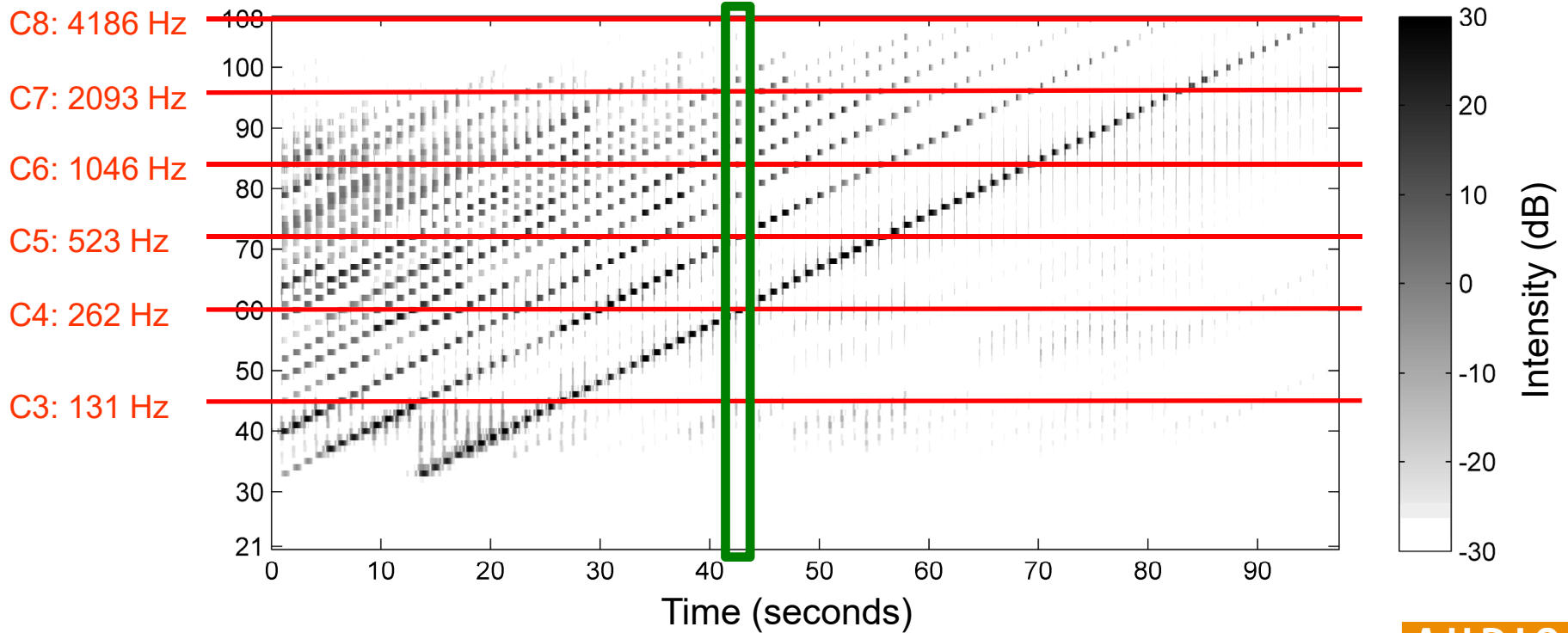
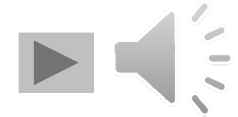
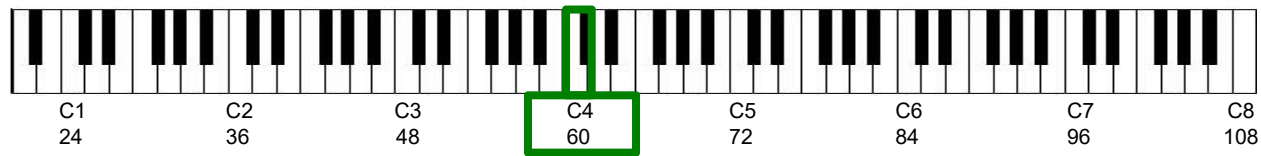
Example: Chromatic scale





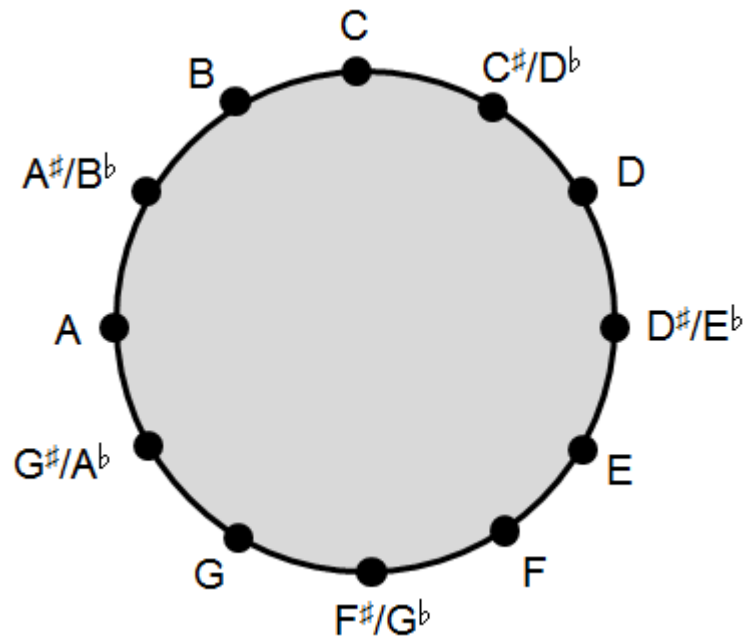
# Excursus: Log-Frequency Spectrogram

Example: Chromatic scale

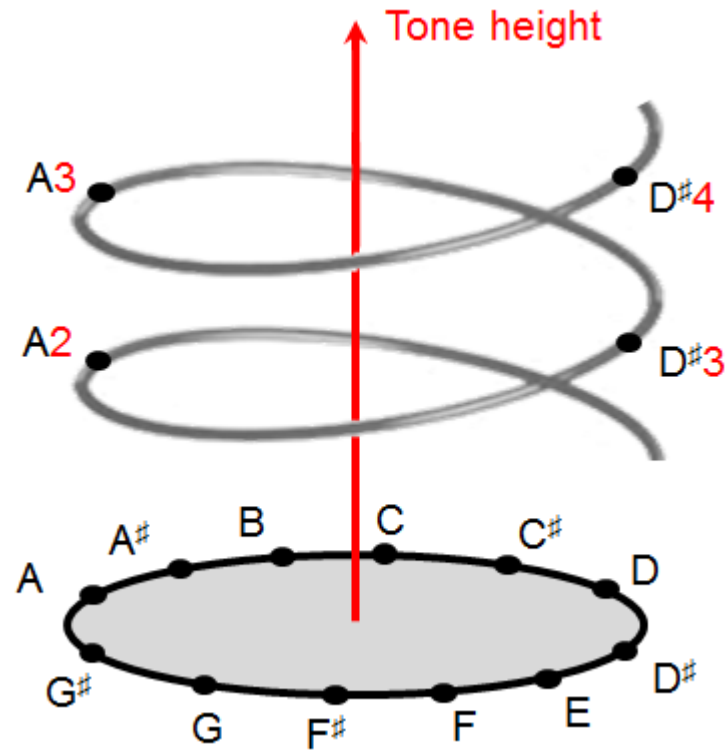


# Excursus: Chroma Features

Chromatic circle

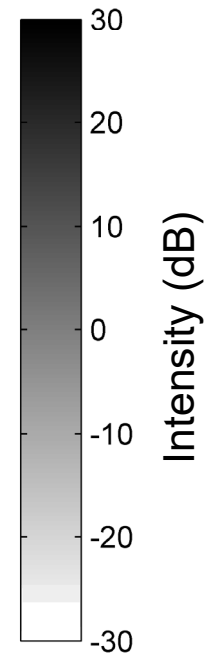
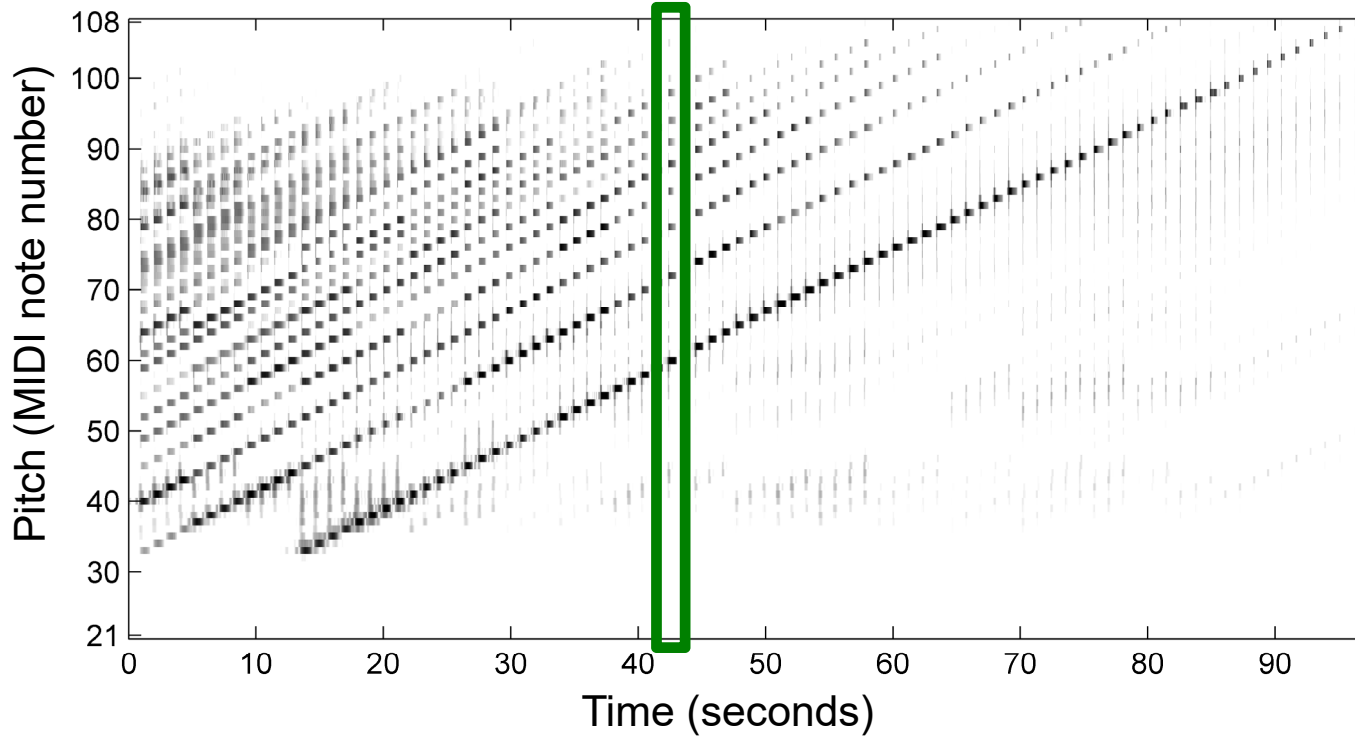
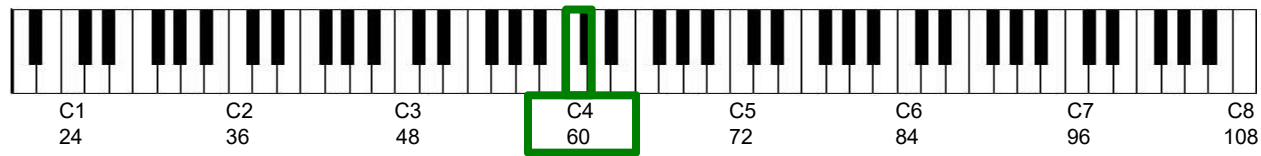


Shepard's helix of pitch



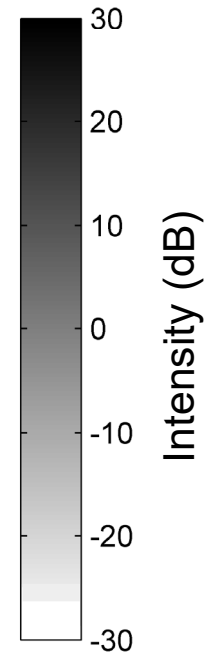
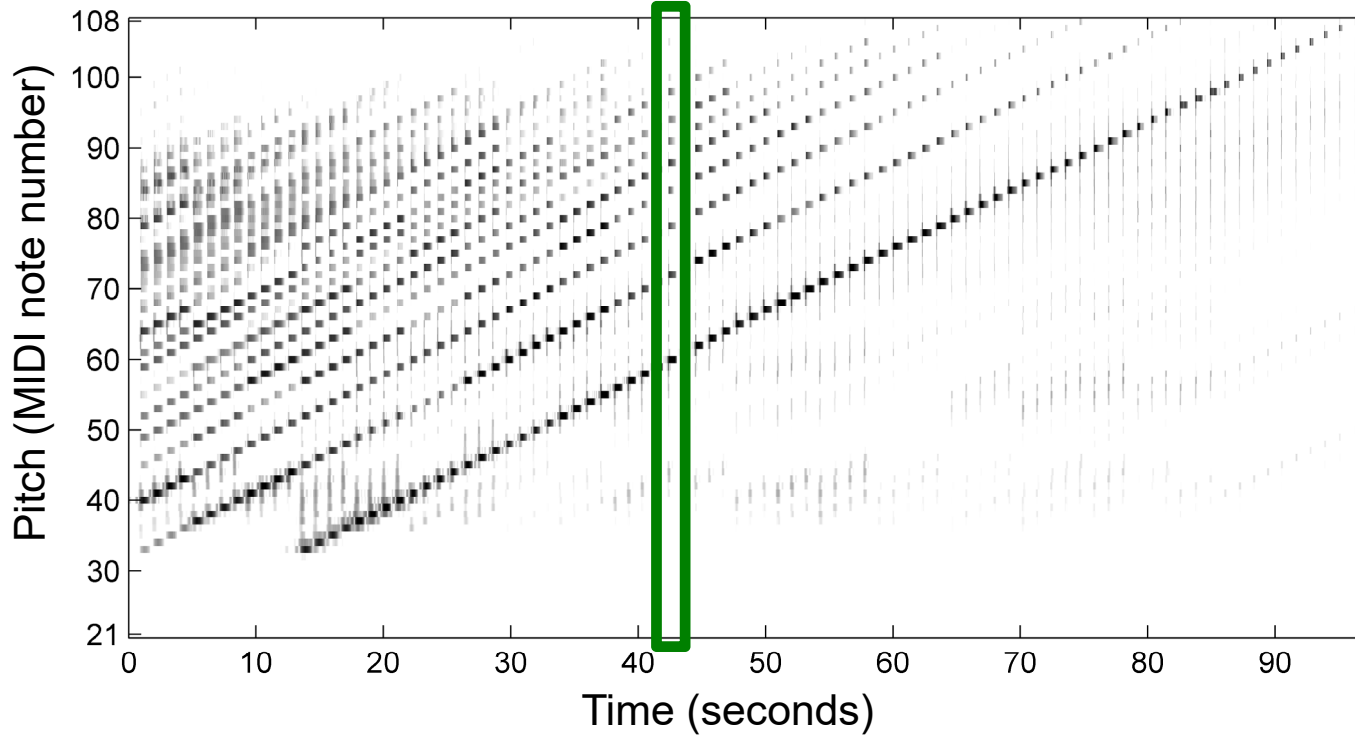
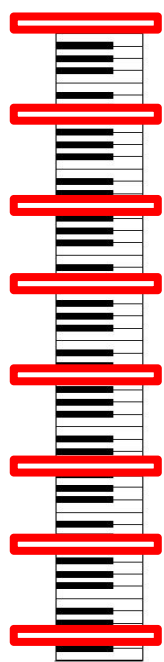
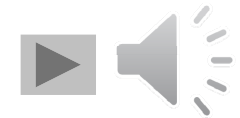
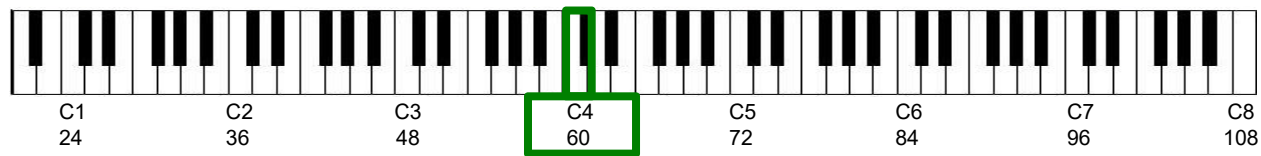
# Excursus: Chroma Features

Example: Chromatic scale



# Excursus: Chroma Features

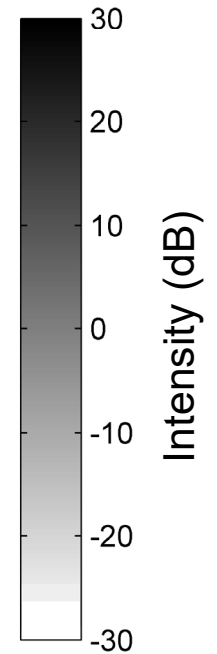
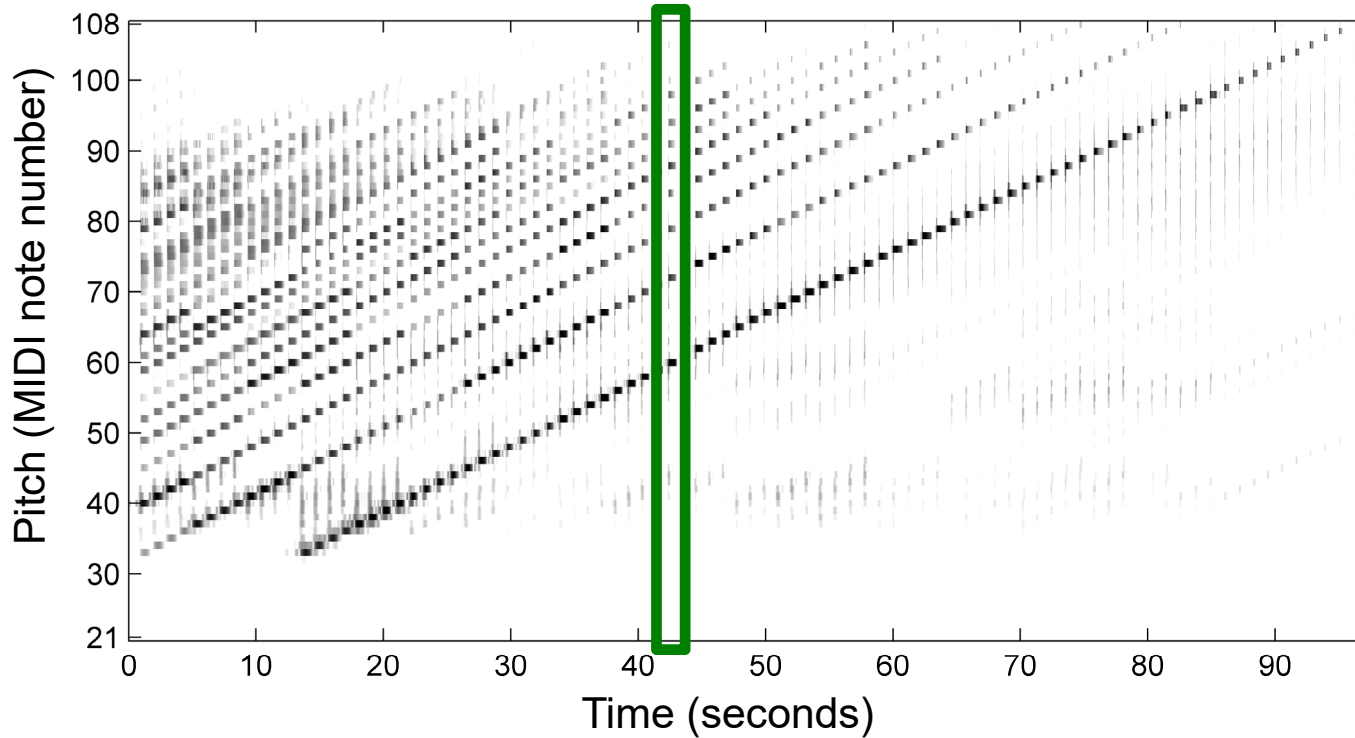
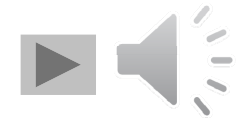
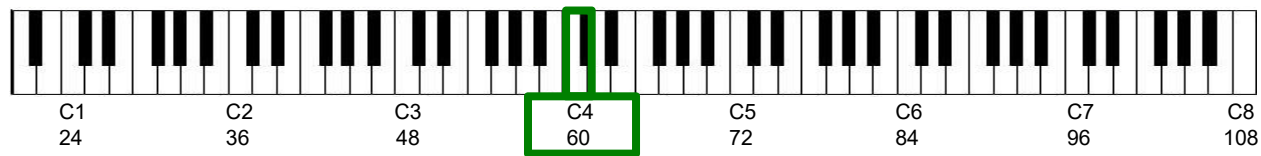
Example: Chromatic scale



Chroma C

# Excursus: Chroma Features

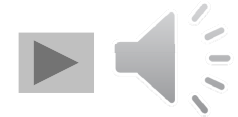
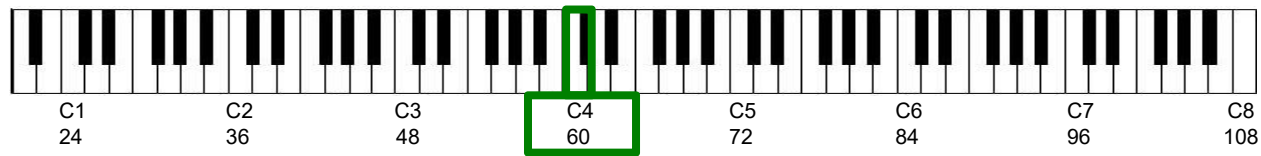
Example: Chromatic scale



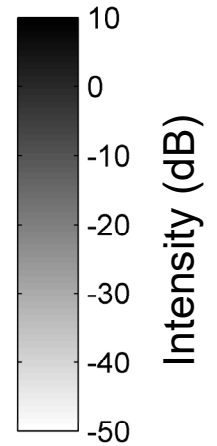
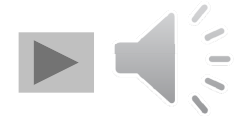
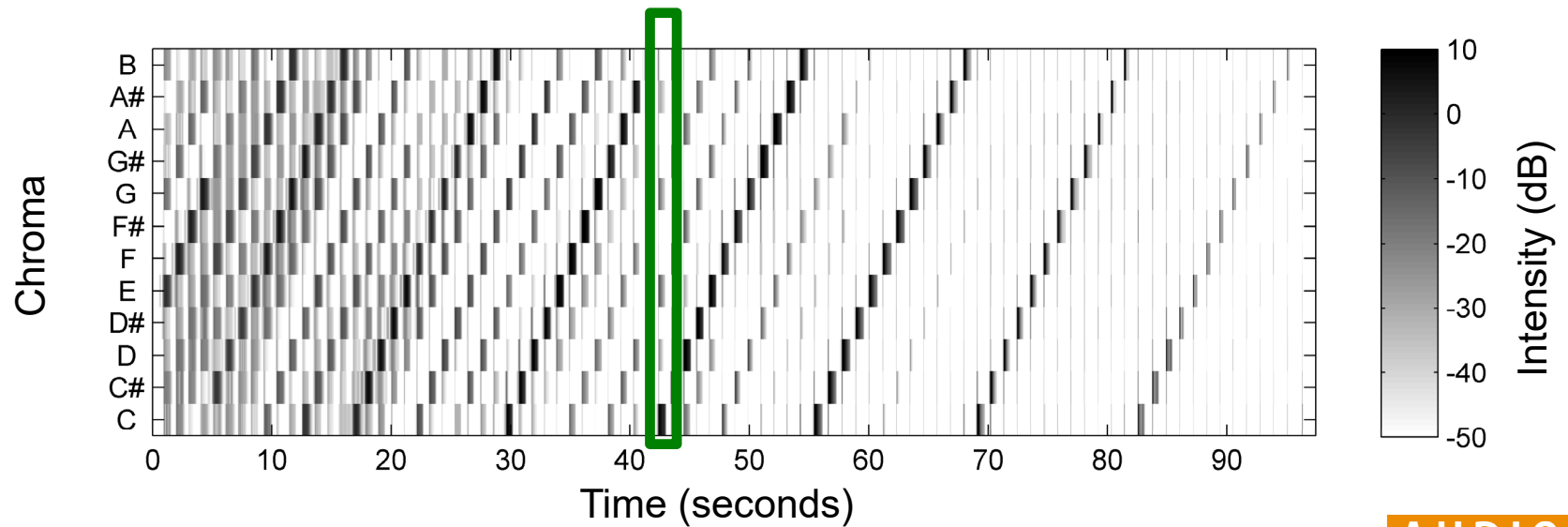
Chroma C#

# Excursus: Chroma Features


Example: Chromatic scale

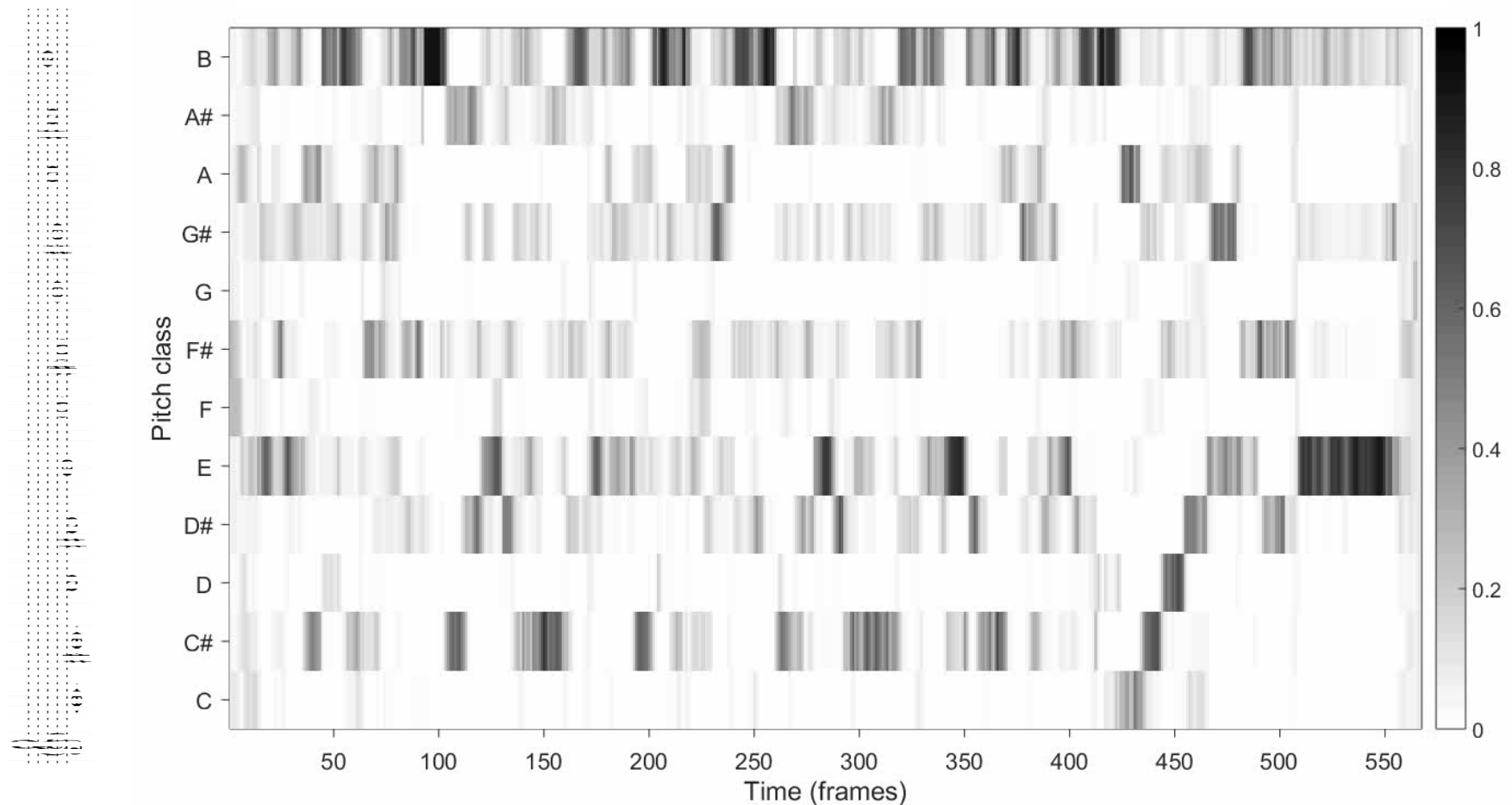


## Chromagram



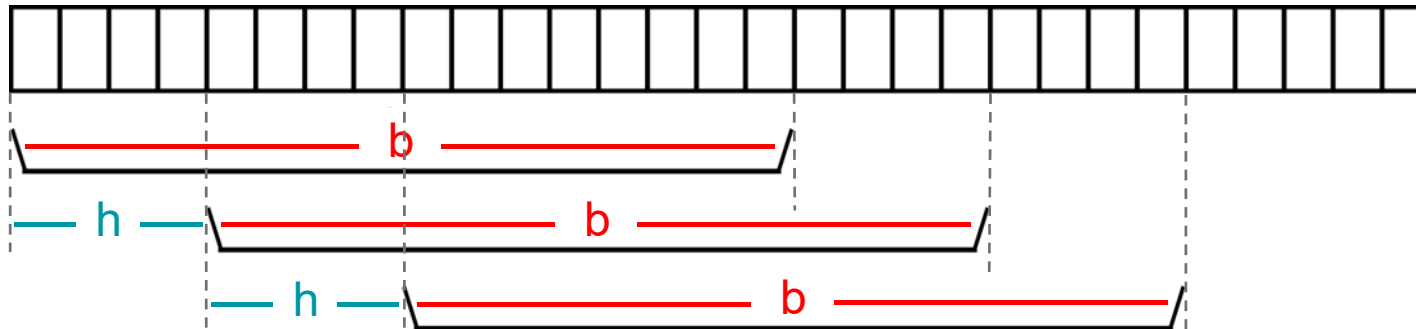
# Visualization of Diatonic Scales

- Example: J.S. Bach, Choral "Durch Dein Gefängnis" (*Johannespassion*)
- **Audio** – Chroma features (Scholars Baroque Ensemble, Naxos 1994) 



# Visualization of Diatonic Scales

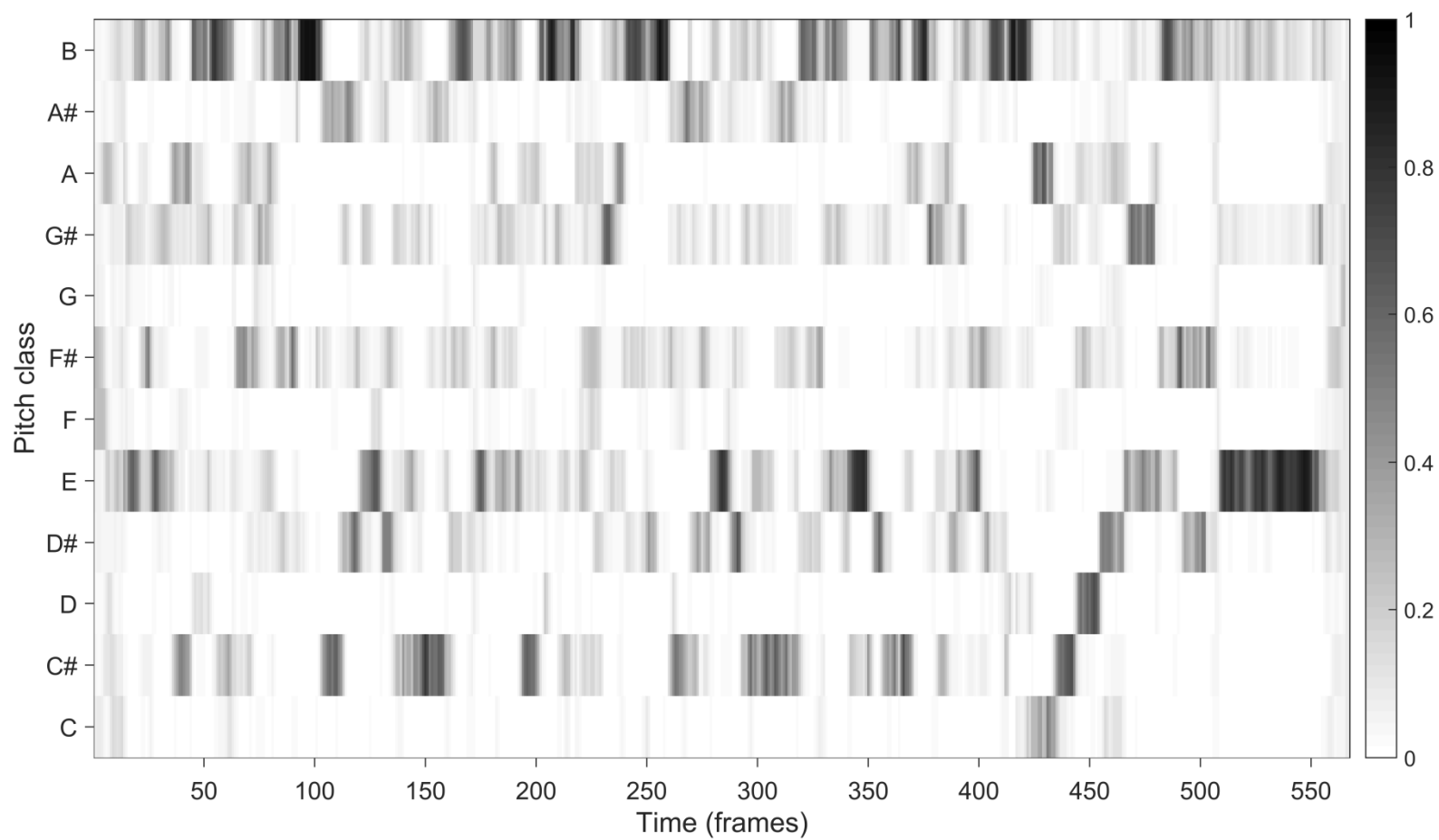
- Summarize pitch classes over a certain time
  - **Chroma smoothing**
  - Parameters: blocksize **b** and hopsize **h**





# Visualization of Diatonic Scales

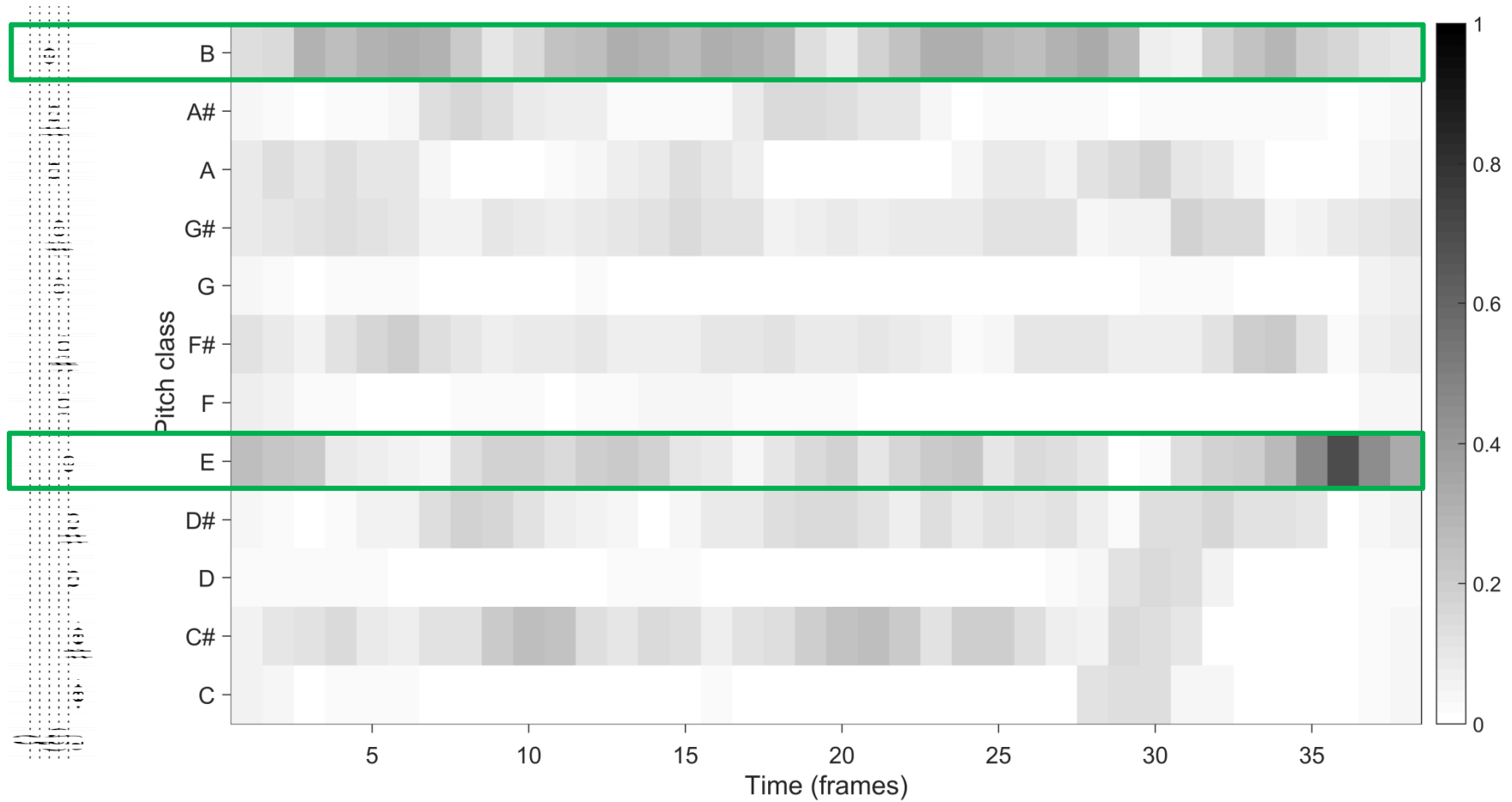
- Choral (Bach)





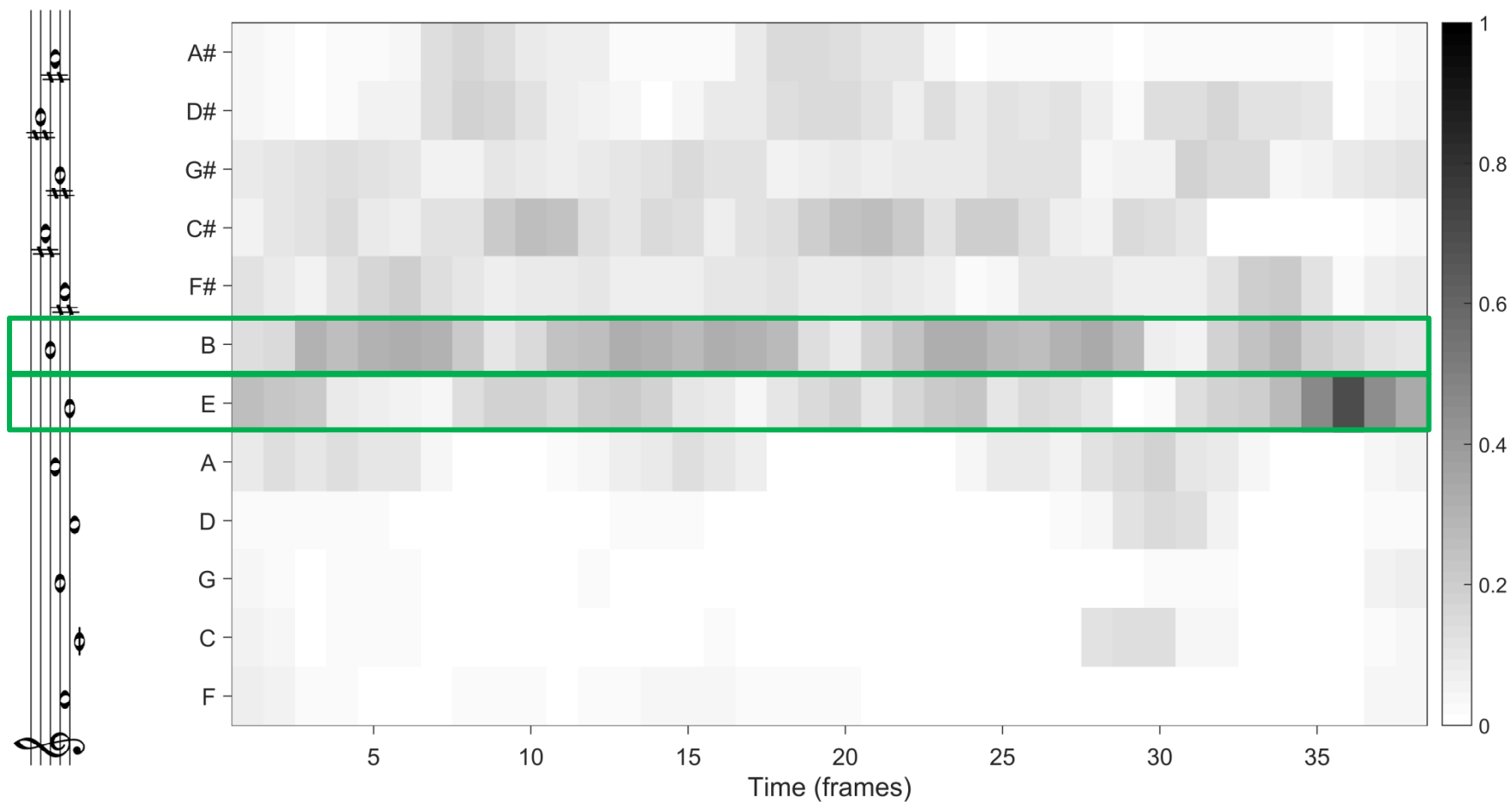
# Visualization of Diatonic Scales

- Choral (Bach) — Re-ordering to **perfect fifth** series



# Visualization of Diatonic Scales

- Choral (Bach) — Re-ordering to **perfect fifth** series



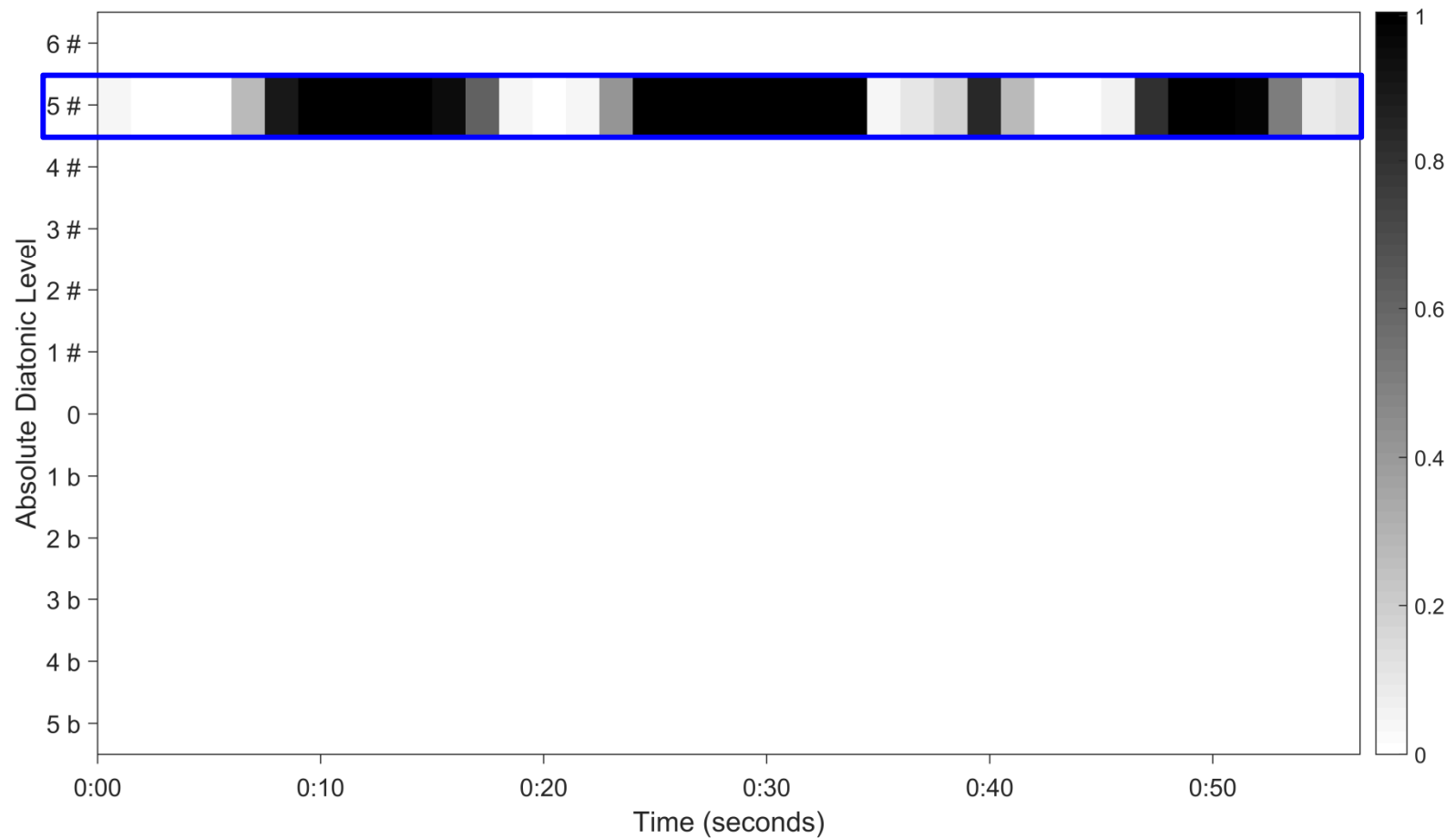






# Visualization of Diatonic Scales

- Choral (Bach) — Diatonic Scale Estimation: [Multiply chroma values](#)



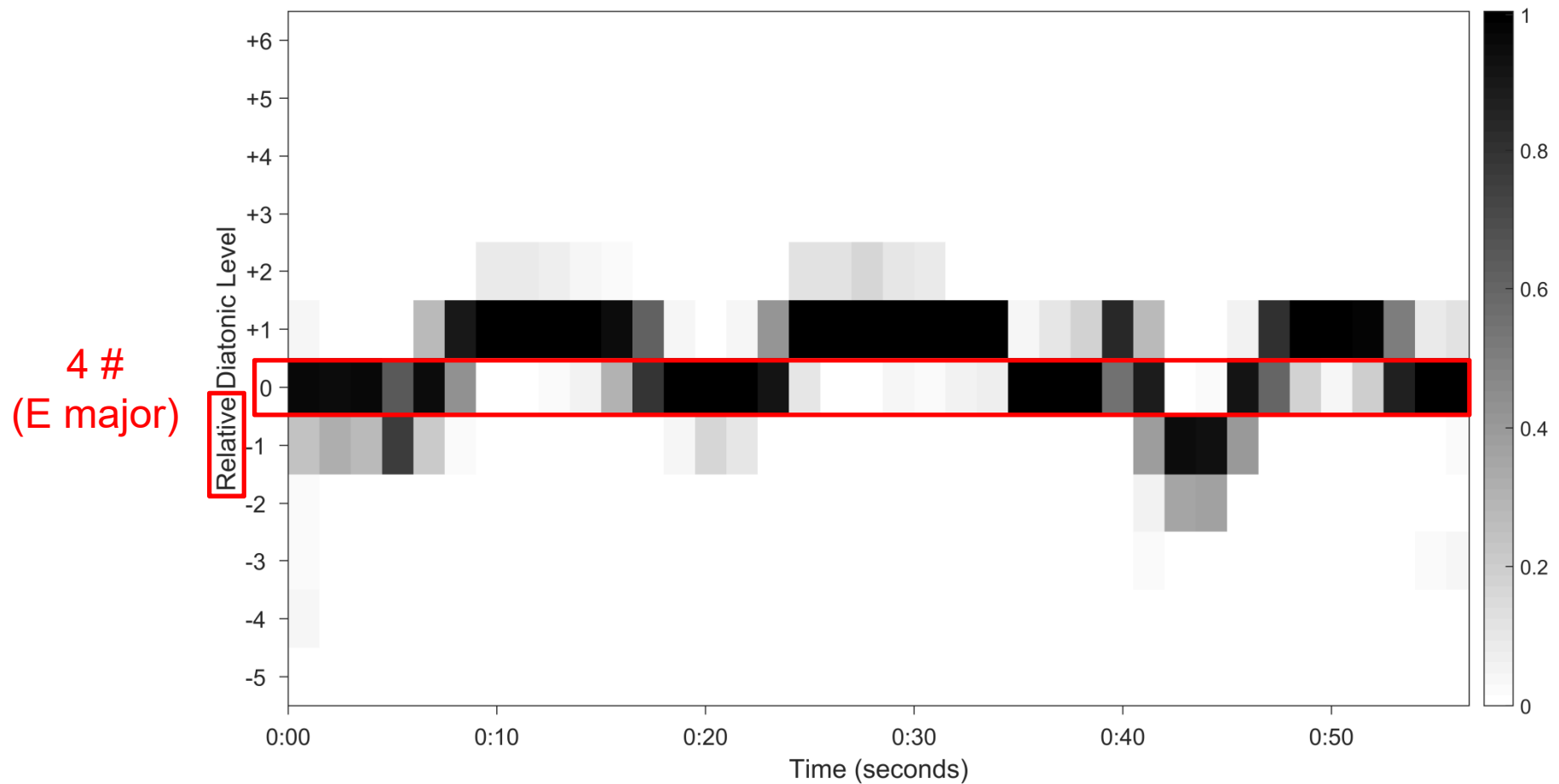






# Visualization of Diatonic Scales

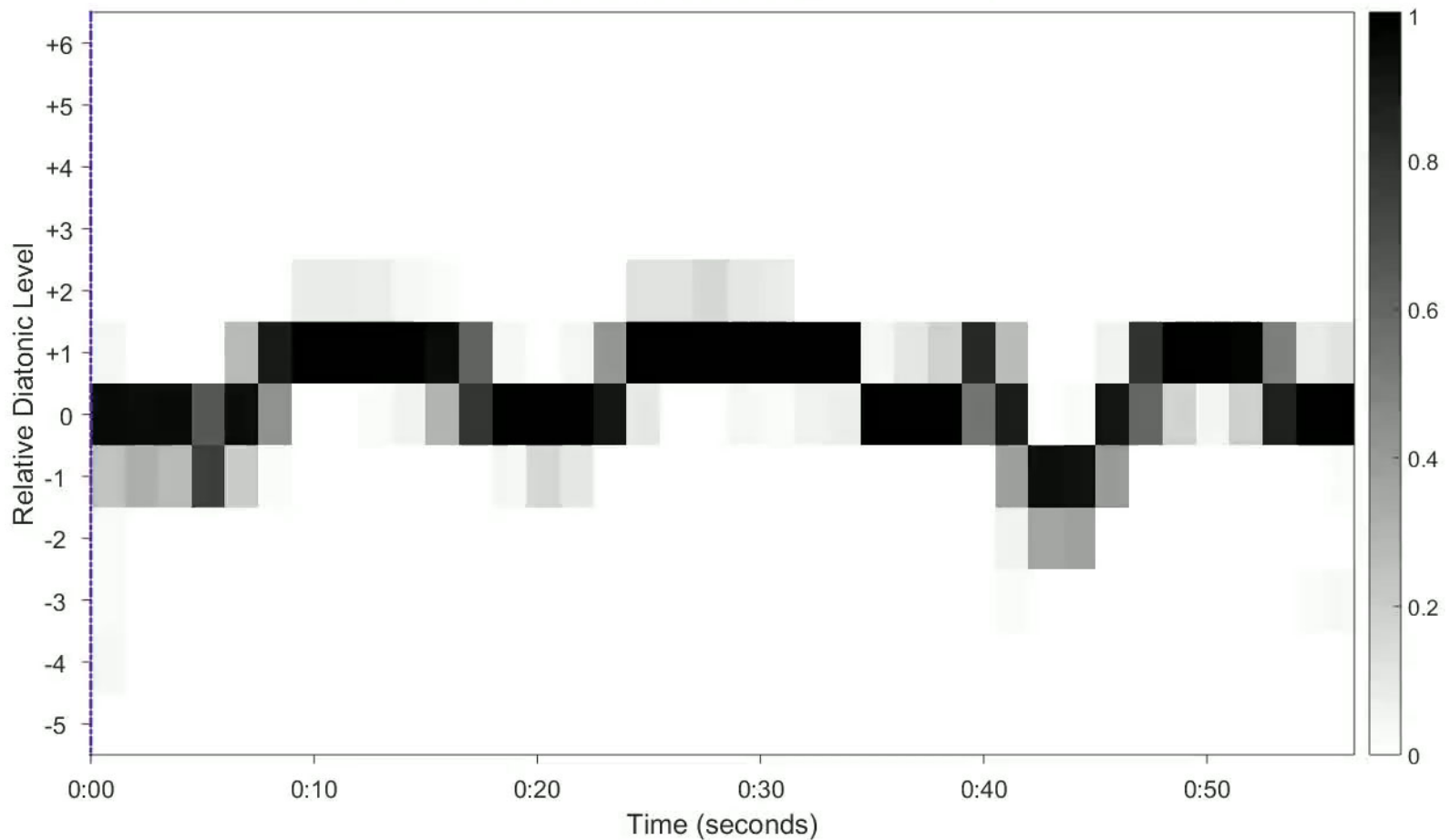
- Choral (Bach) — Diatonic Scale Estimation: **Shift to global key**



# Visualization of Diatonic Scales

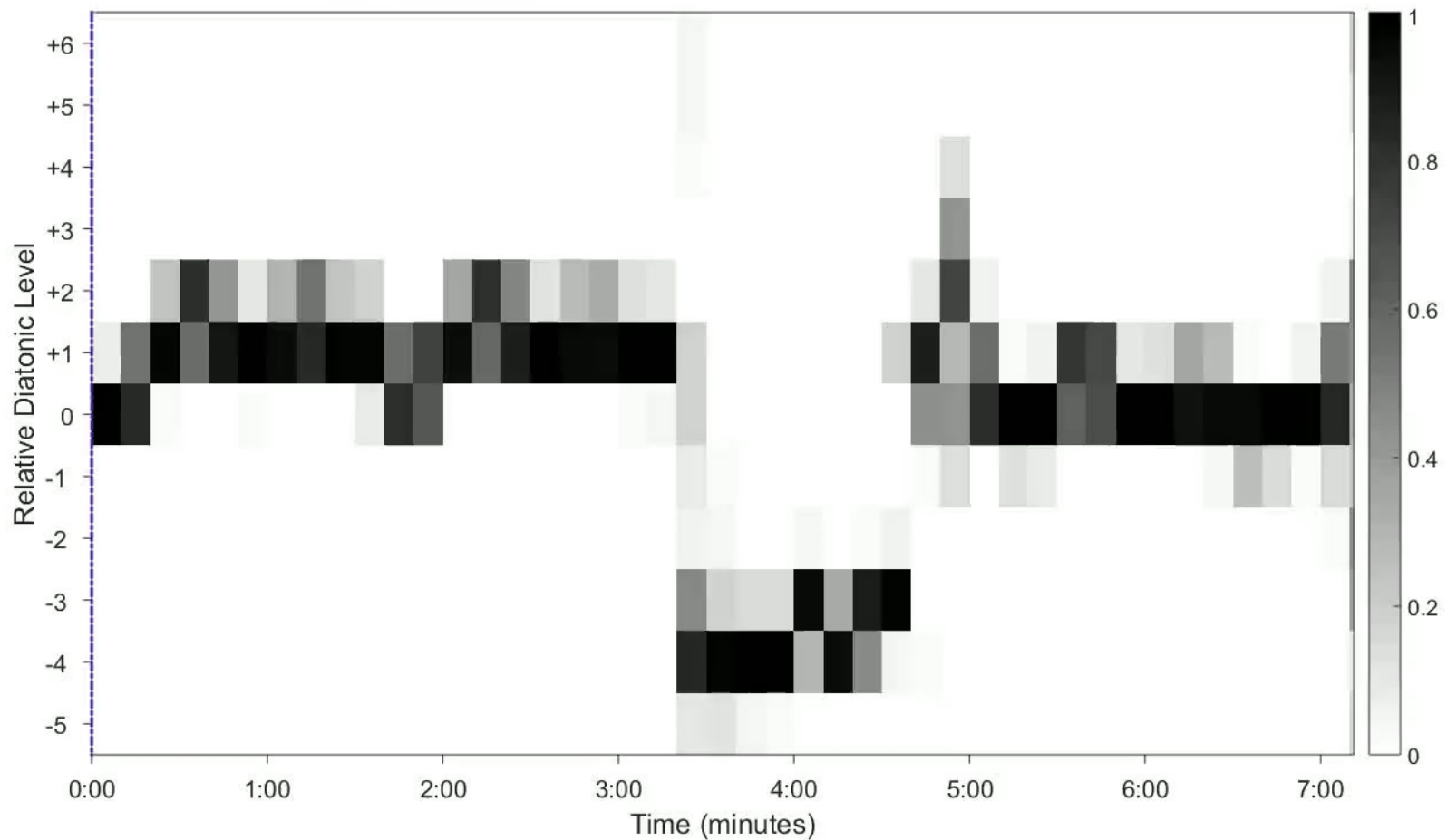
- Choral (Bach) — 0  $\hat{=}$  4#

C. Weiß, J. Habryka, "Chroma-Based Scale Matching for Audio Tonality Analysis" In: *Proceedings of the 9th Conference on Interdisciplinary Musicology*, Berlin 2014.



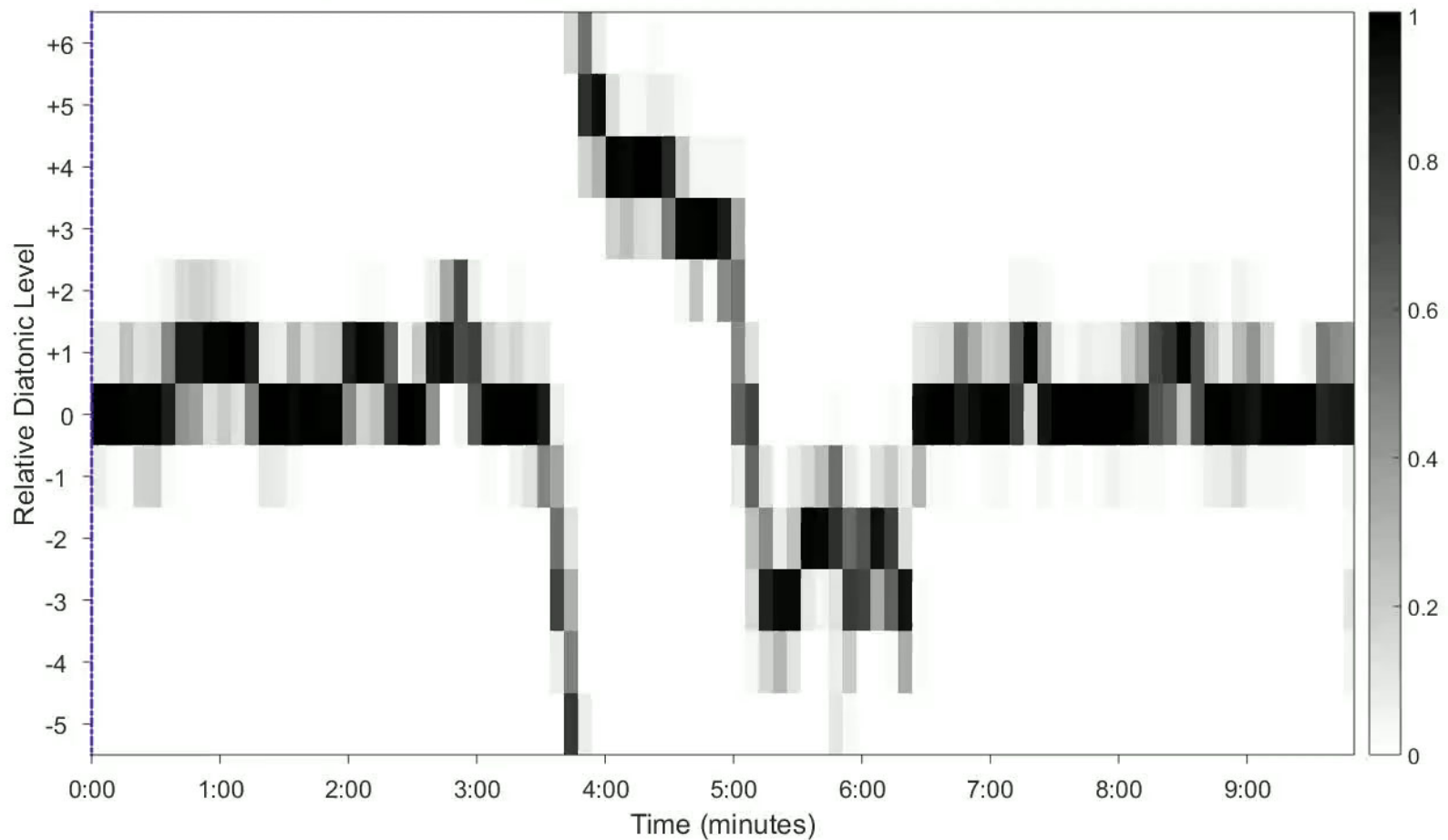
# Visualization of Diatonic Scales

- L. v. Beethoven – Sonata No. 10 op. 14 Nr. 2, 1. Allegro — 0  $\hat{=}$  1  
(Barenboim, EMI 1998)



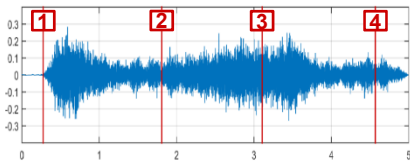
# Visualization of Diatonic Scales

- R. Wagner, *Die Meistersinger von Nürnberg*, Vorspiel — 0  $\hat{=}$  0  
(Polish National Radio Symphony Orchestra, J. Wildner, Naxos 1993)



# Cross-Version Analysis

- Up to 18 versions
- 3 versions manually annotated



No.	Conductor	Recording	hh:mm:ss
→ 1	Barenboim	1991–92	14:54:55
2	Boulez	1980–81	13:44:38
3	Böhm	1967–71	13:39:28
4	Furtwängler	1953	15:04:22
→ 5	Haitink	1988–91	14:27:10
6	Janowski	1980–83	14:08:34
→ 7	Karajan	1967–70	14:58:08
8	Keilberth/Furtwängler	1952–54	14:19:56
9	Krauss	1953	14:12:27
10	Levine	1987–89	15:21:52
11	Neuhold	1993–95	14:04:35
12	Sawallisch	1989	14:06:50
13	Solti	1958–65	14:36:58
14	Swarowsky	1968	14:56:34
15	Thielemann	2011	14:31:13
16	Weigle	2010–12	14:48:46

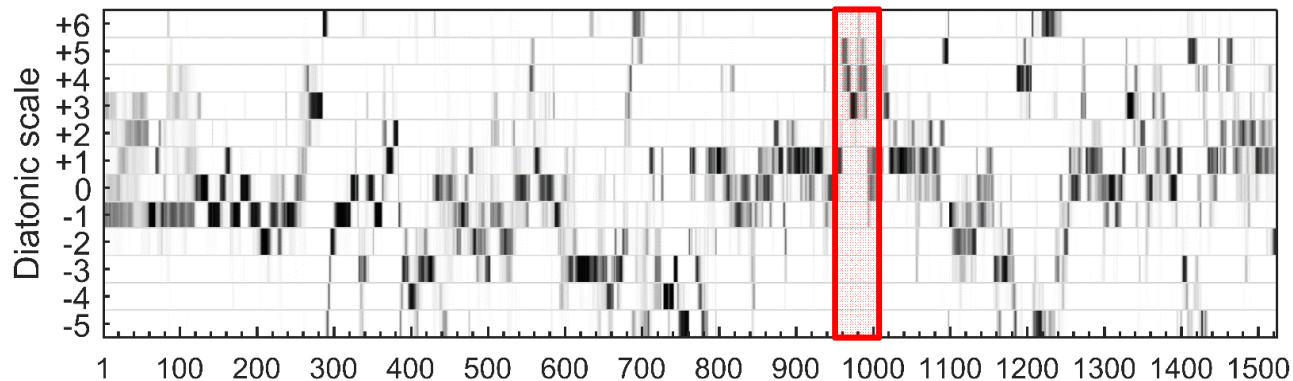
# Cross-Version Analysis

- Idea: Use analysis results based on different interpretations (versions)
- Tonal characteristics should not depend on interpretation  
→ Test reliability of the method
- Visualize consistency with gray scheme

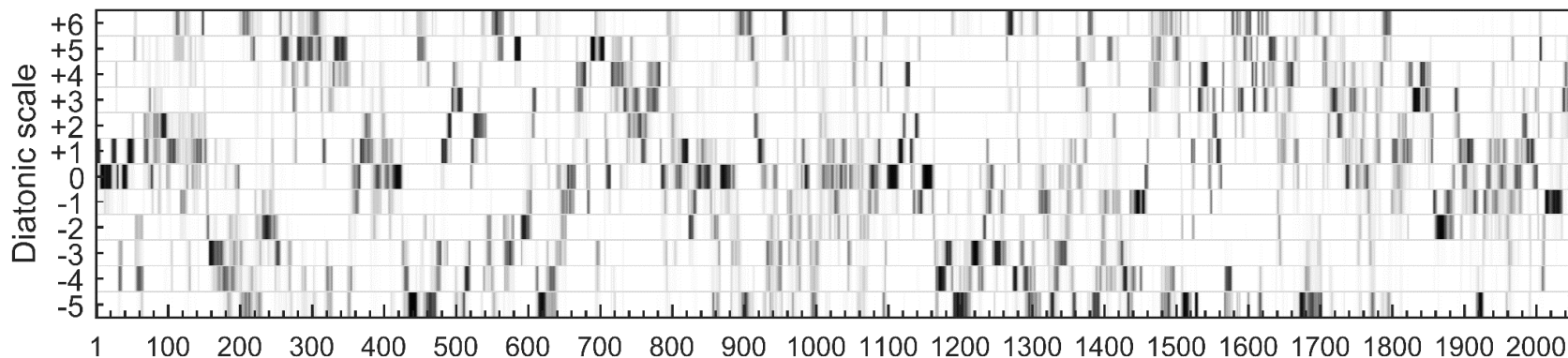


# Die Walküre WWV 86 B

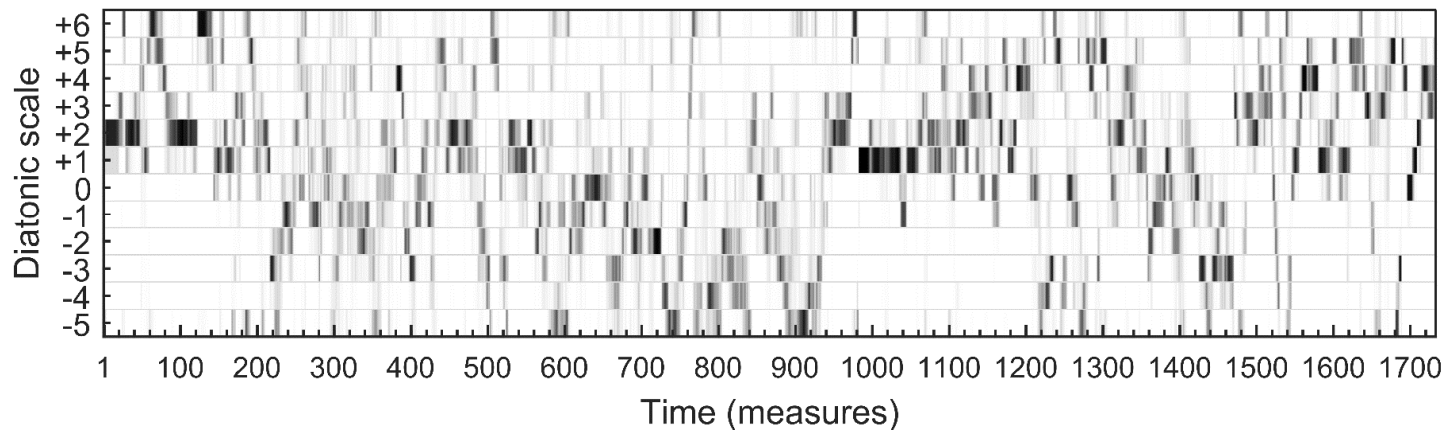
Act 1



Act 2

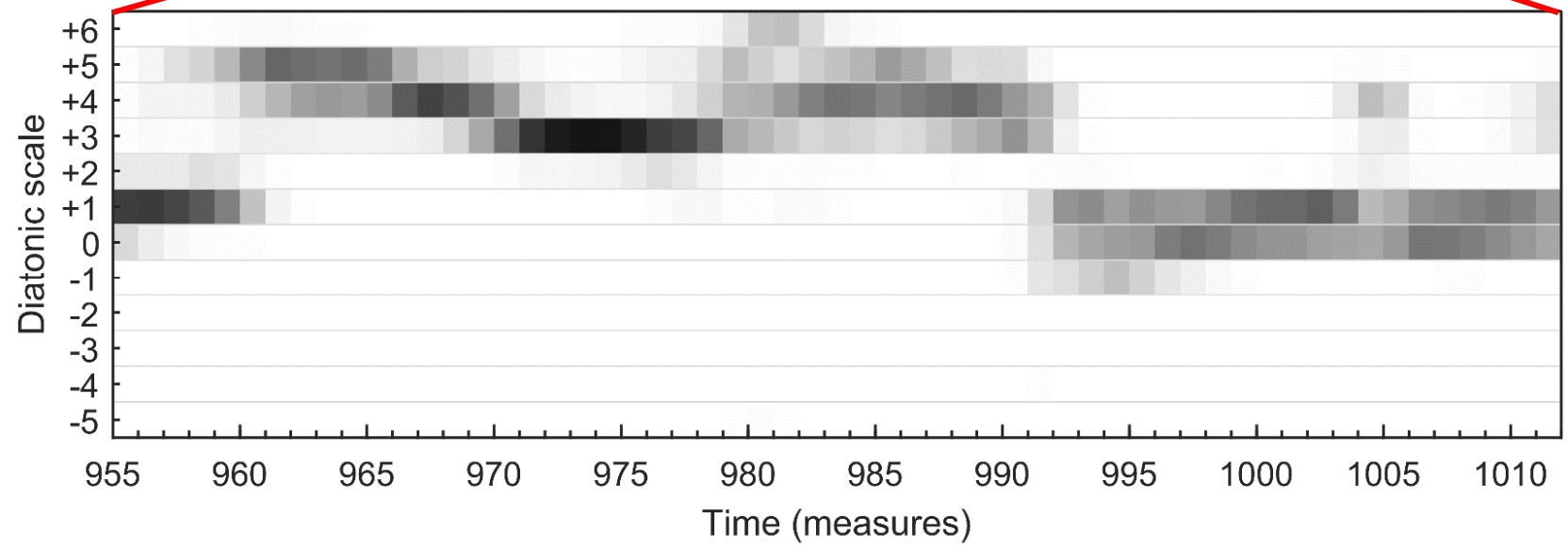
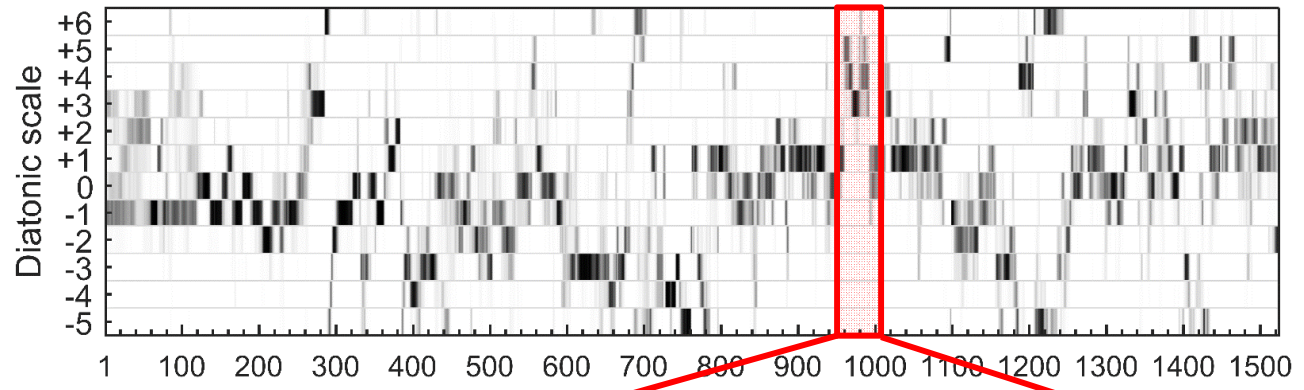


Act 3



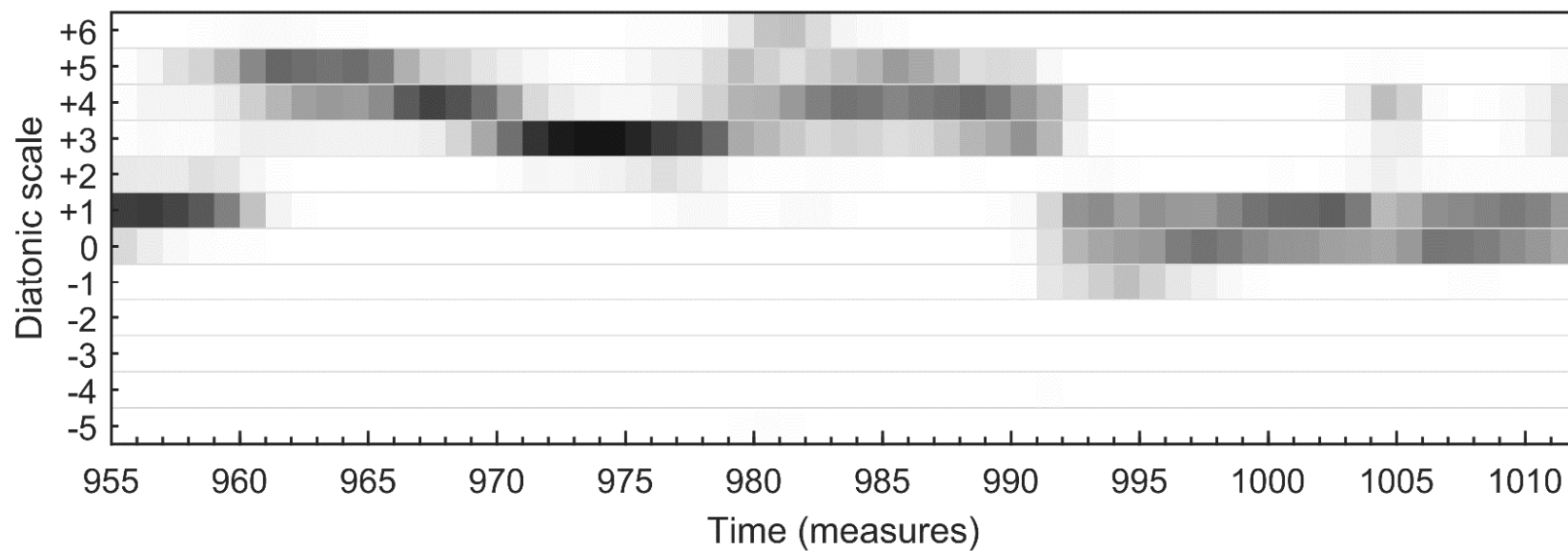
# Die Walküre WWV 86 B

Act 1



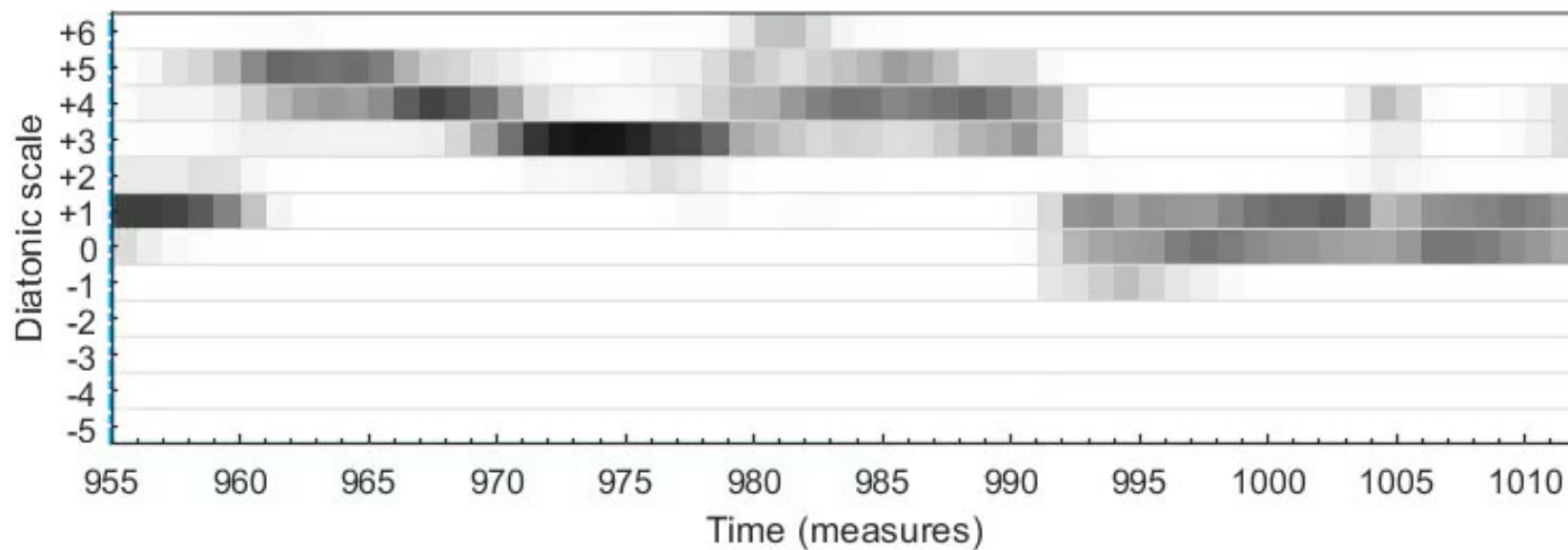
## Die Walküre WWV 86 B

- Act 1, measures 955–1012
- Sieglinde's narration



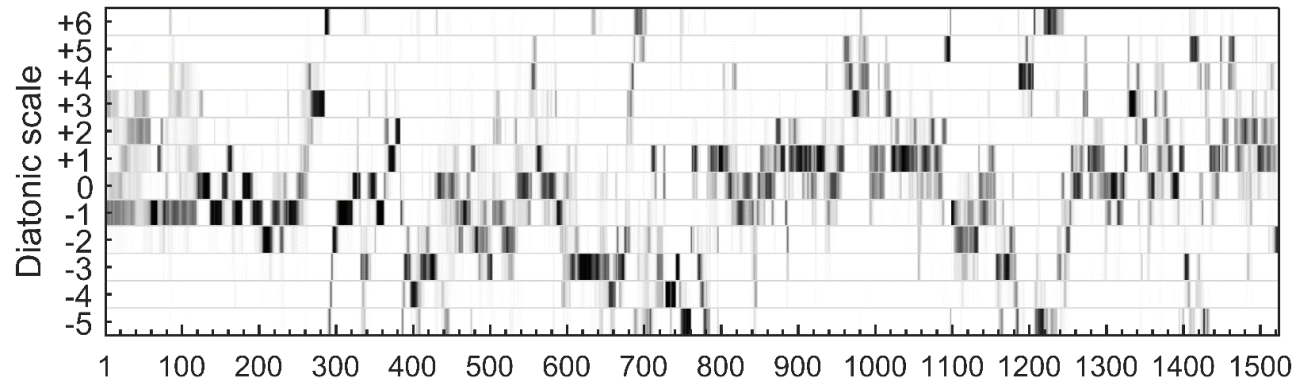
## Die Walküre WWV 86 B

- Act 1, measures 955–1012
- Sieglinde's narration

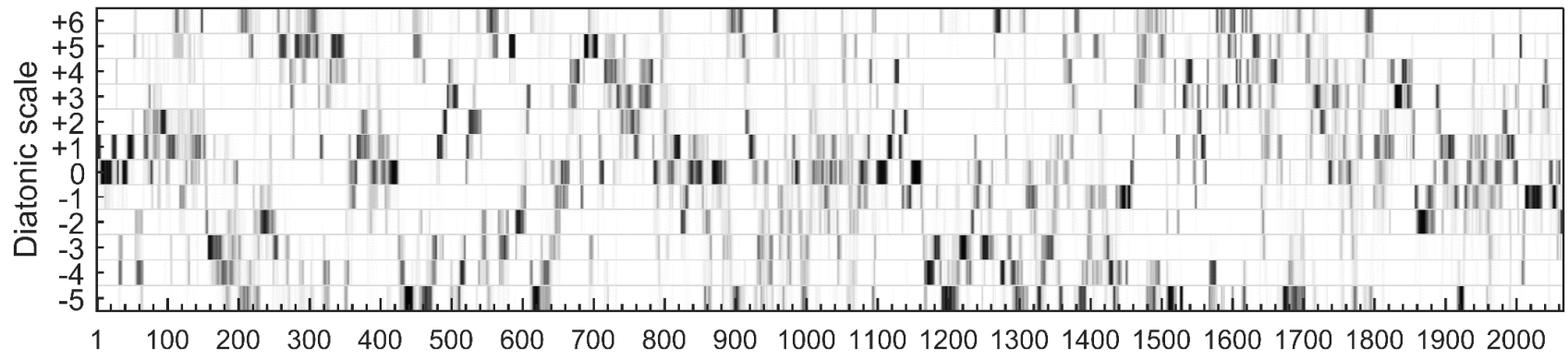


# Die Walküre WWV 86 B

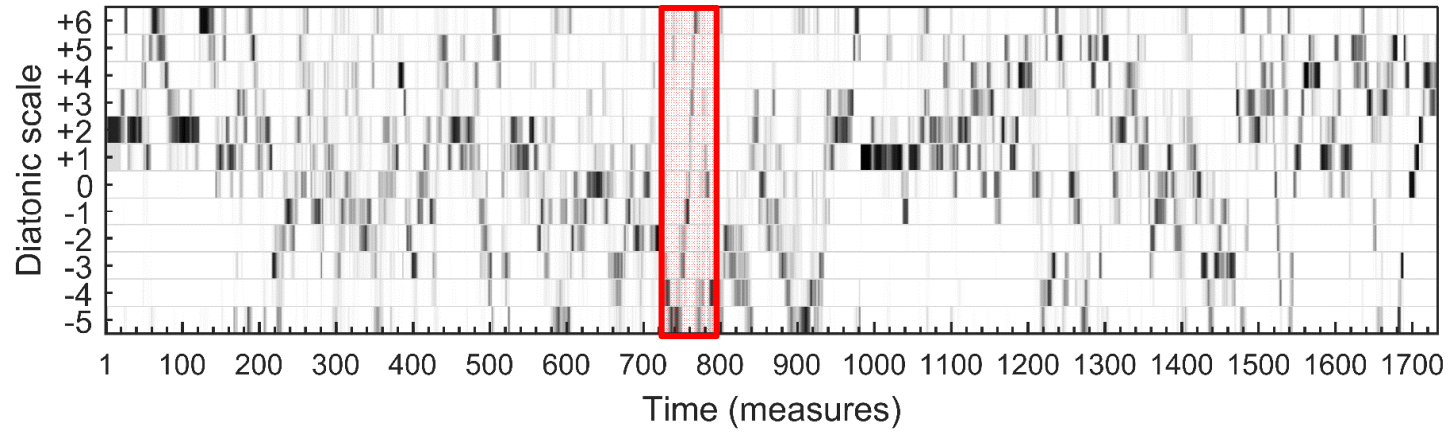
Act 1



Act 2

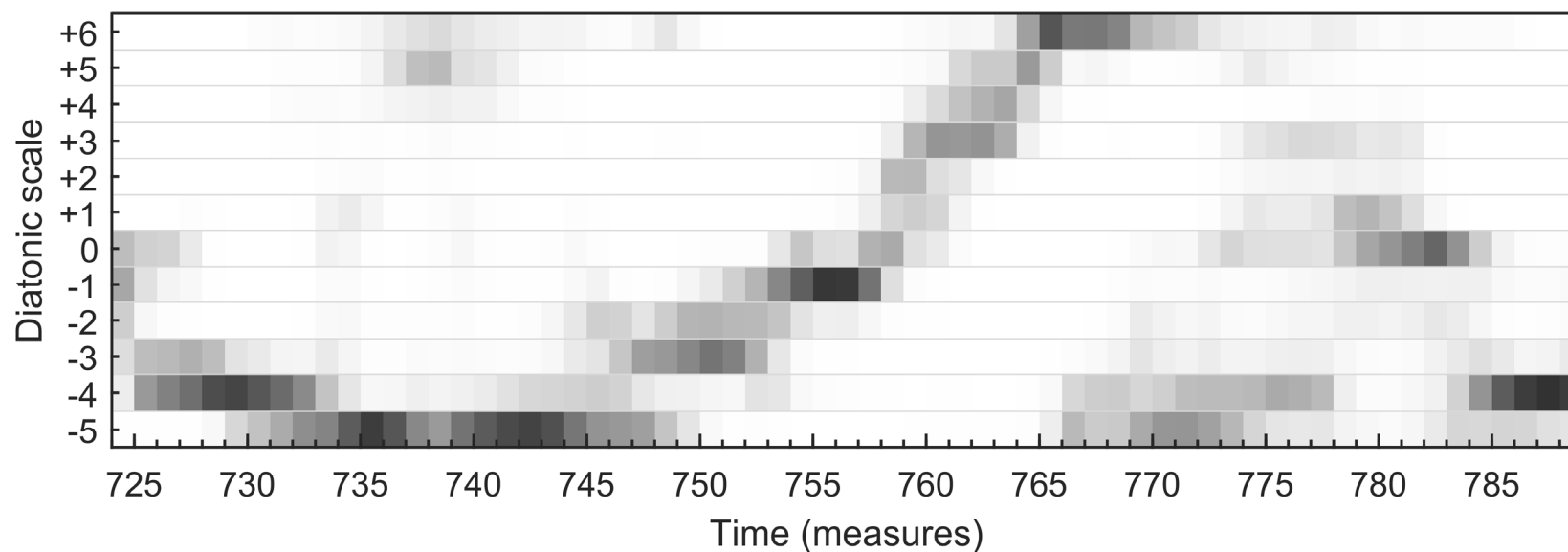


Act 3



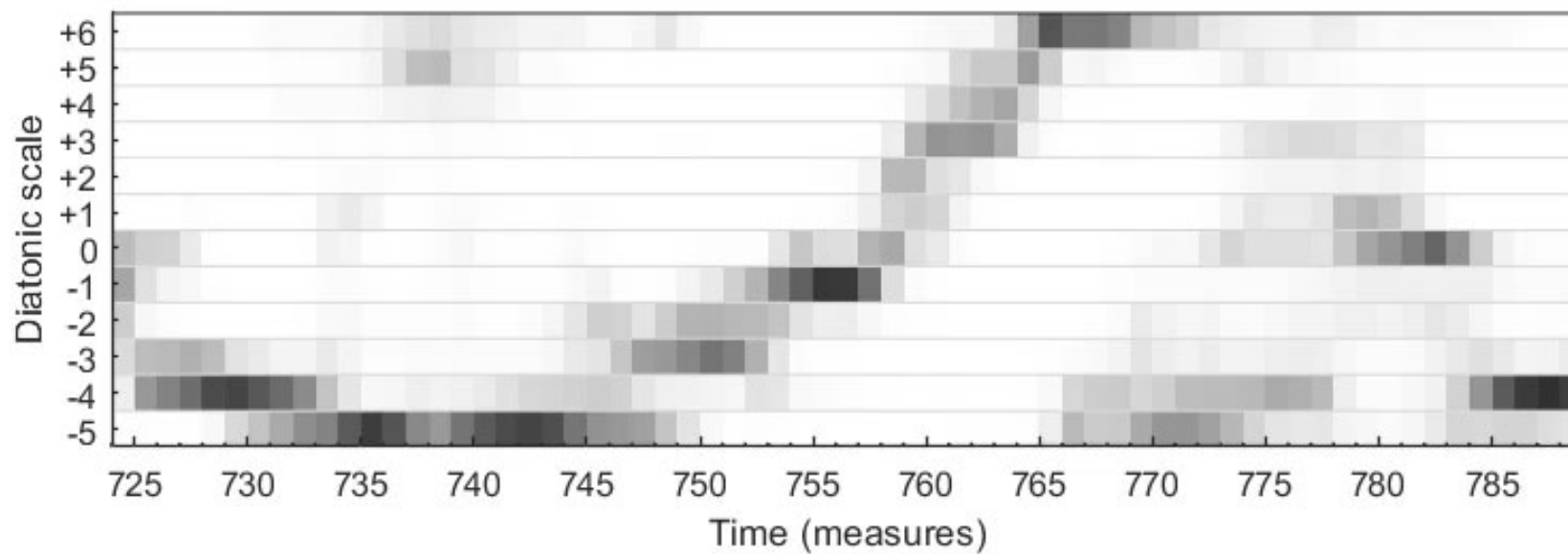
## Die Walküre WWV 86 B

- Act 3, measures 724–789
- Wotan's punishment

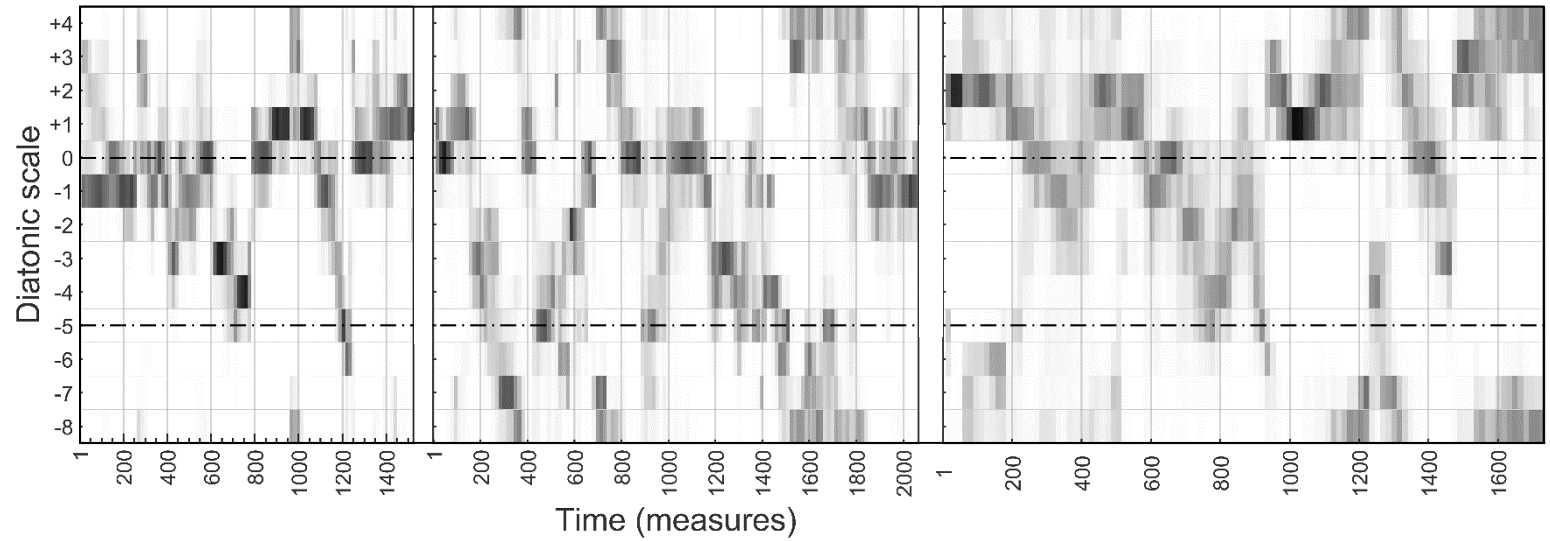
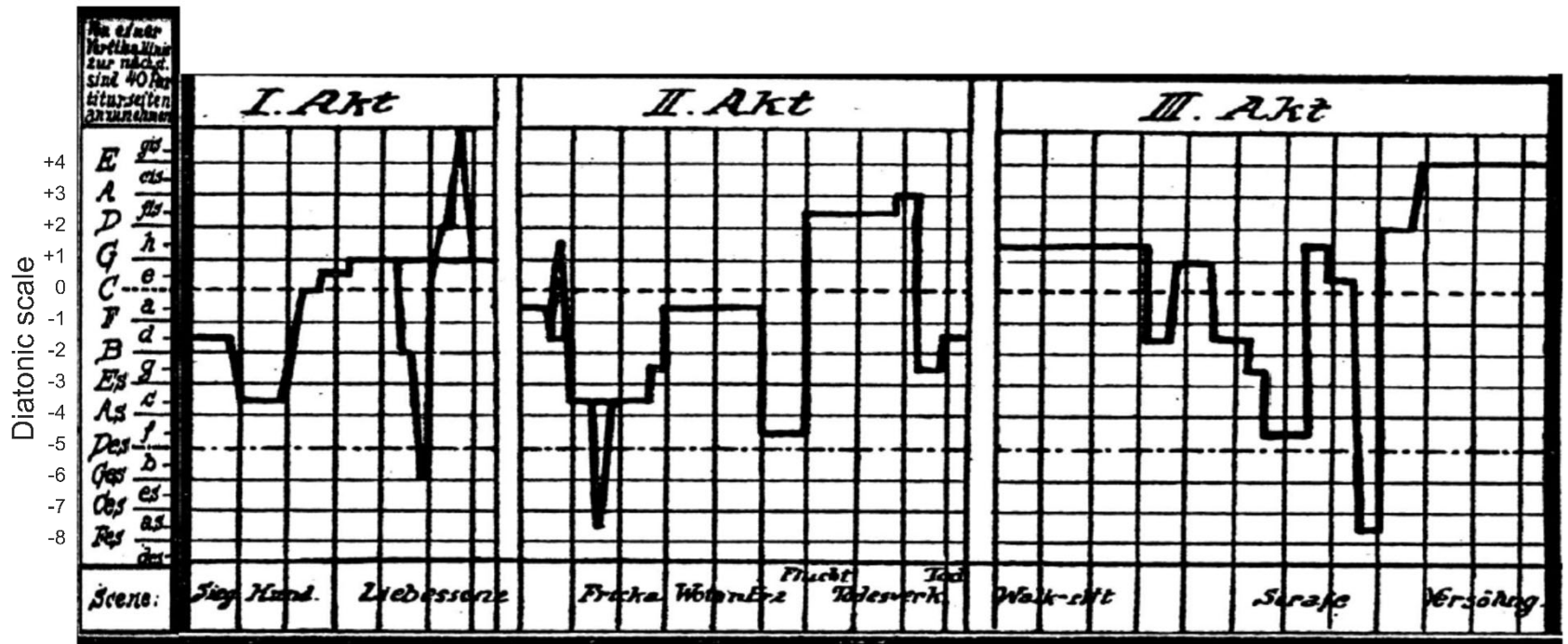


## Die Walküre WWV 86 B

- Act 3, measures 724–789
- Wotan's punishment



# Die Walküre WWV 86 B

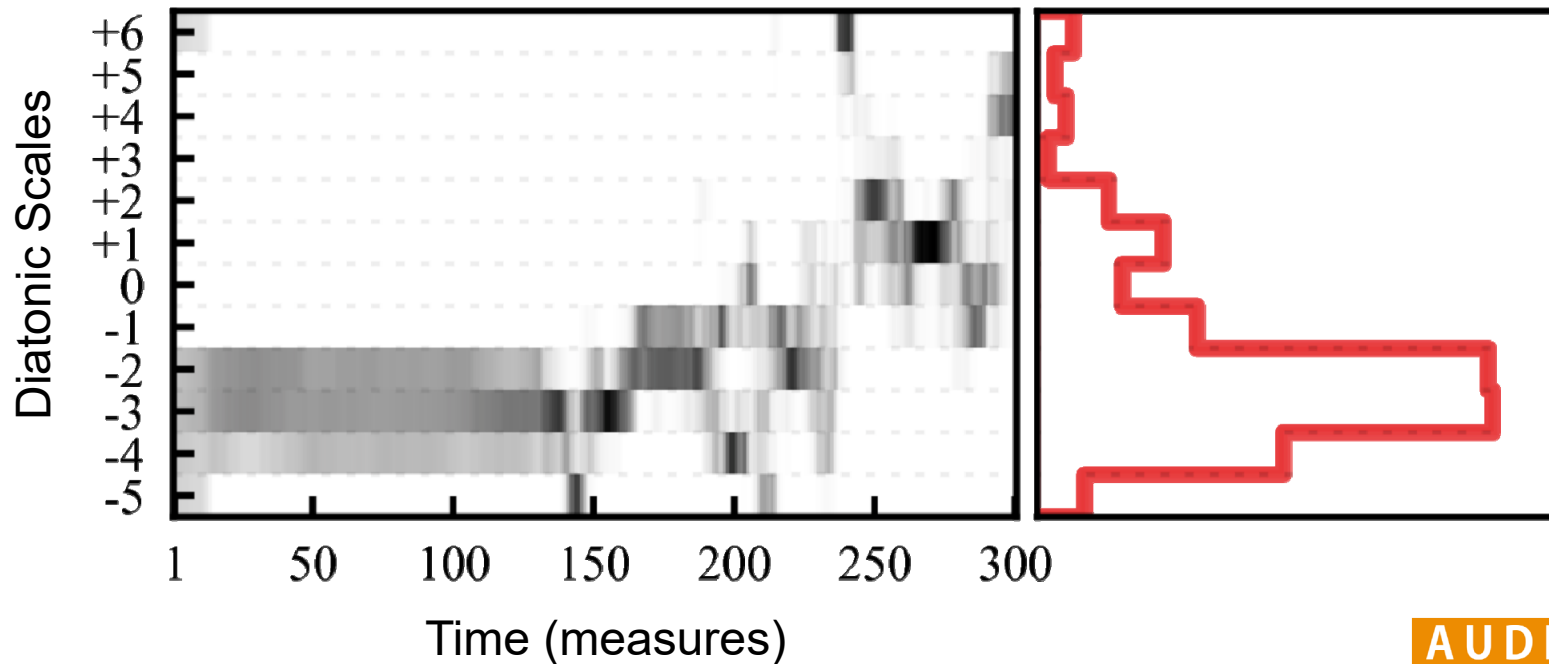




# Exploring Tonal-Dramatic Relationships

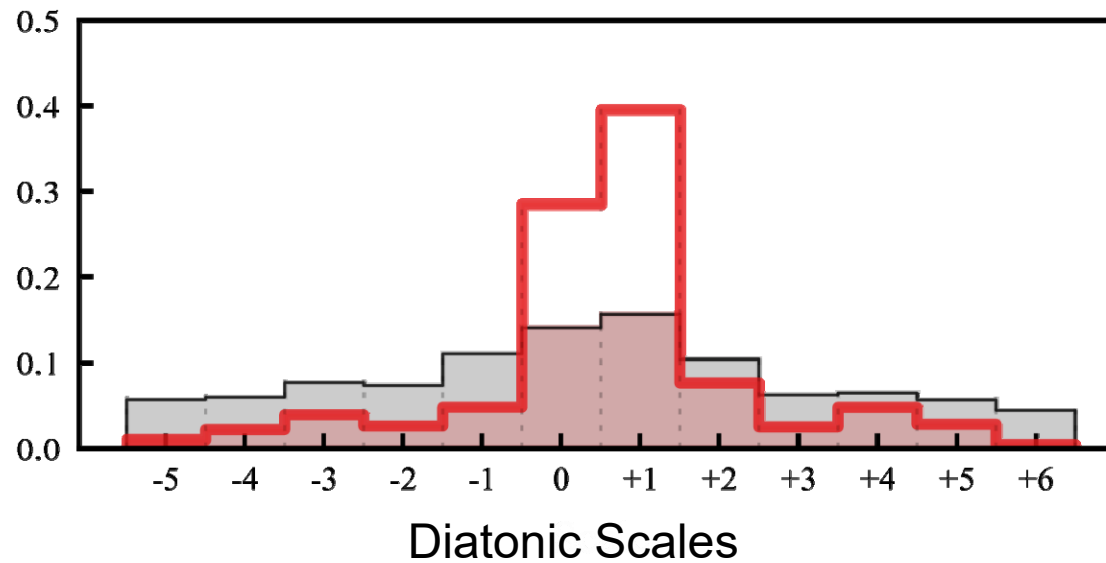
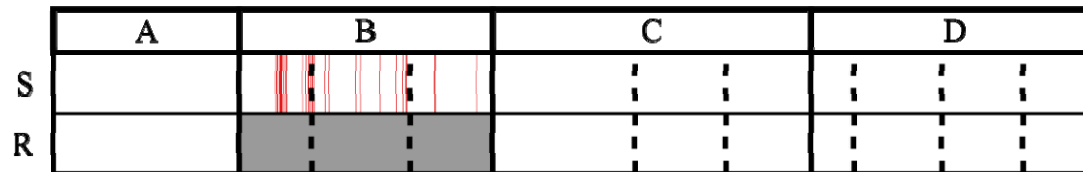
- Histograms of Analysis over time

Das Rheingold WWV 86 A 3897 measures	Die Walküre WWV 86 B 5322 measures	Siegfried WWV 86 C 6682 measures	Götterdämmerung WWV 86 D 6040 measures
--	--	--	--



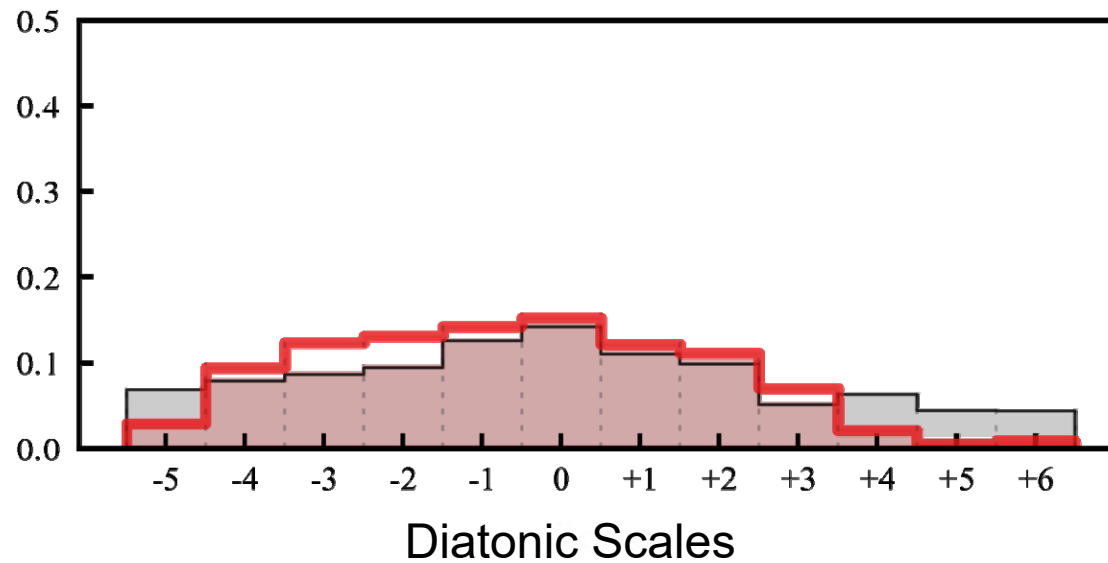
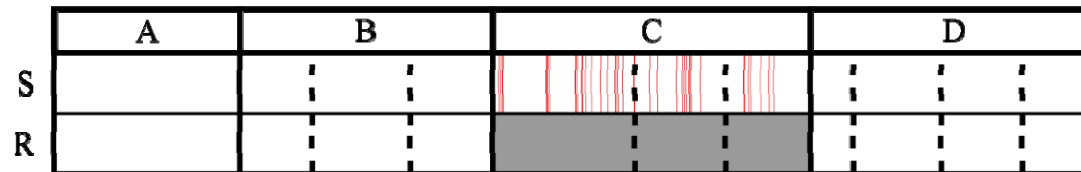
# Exploring Tonal-Dramatic Relationships

## Sword motif – *Die Walküre*



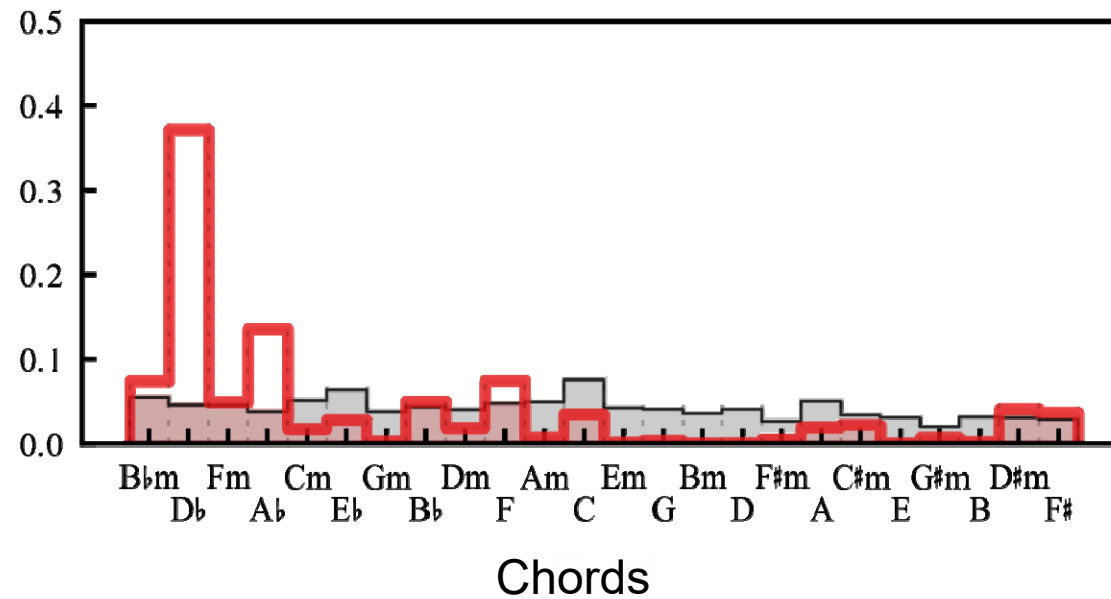
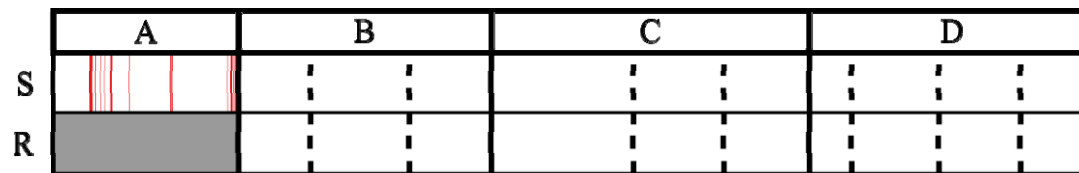
# Exploring Tonal-Dramatic Relationships

## Sword motif – *Siegfried*



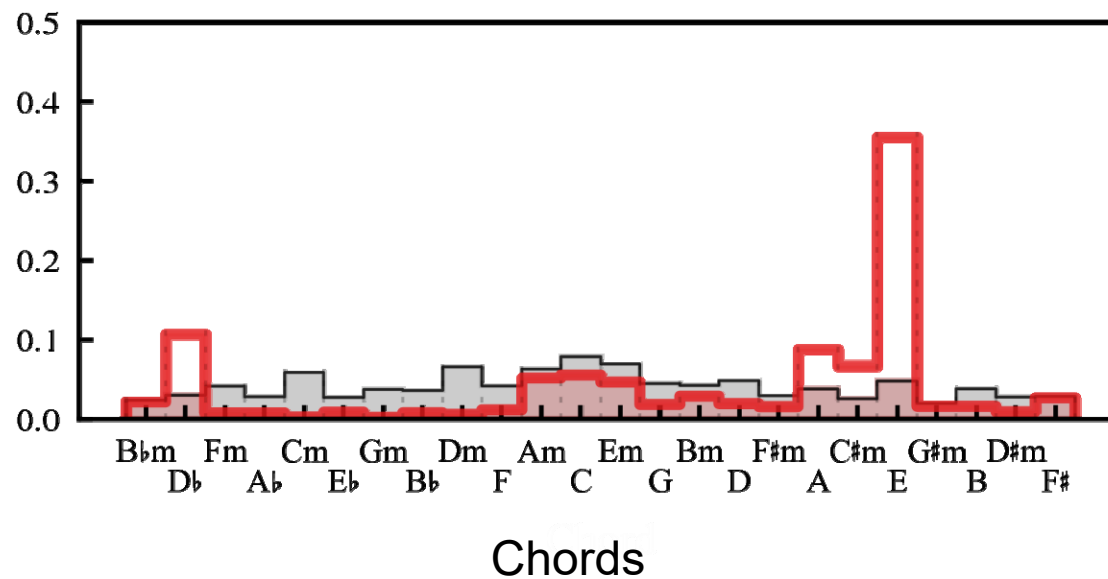
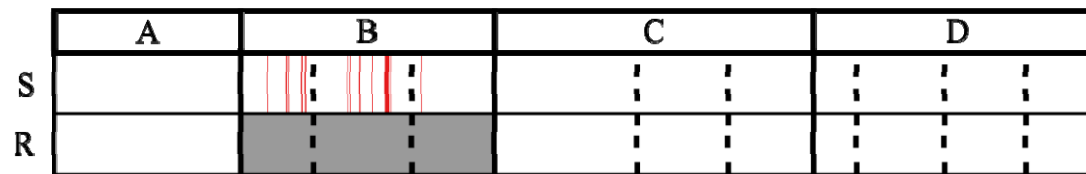
# Exploring Tonal-Dramatic Relationships

## Valhalla motif – *Das Rheingold*

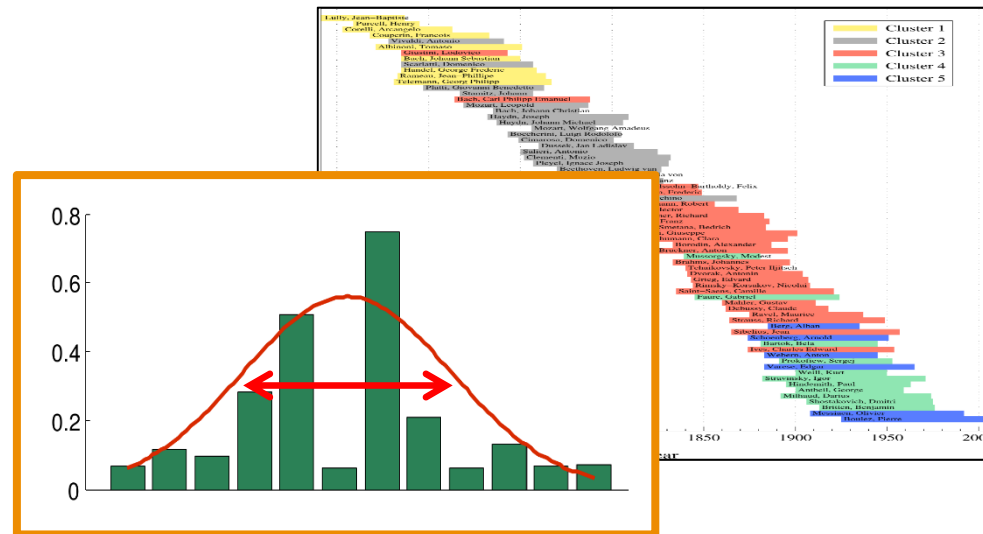


# Exploring Tonal-Dramatic Relationships

## Valhalla motif – *Die Walküre*

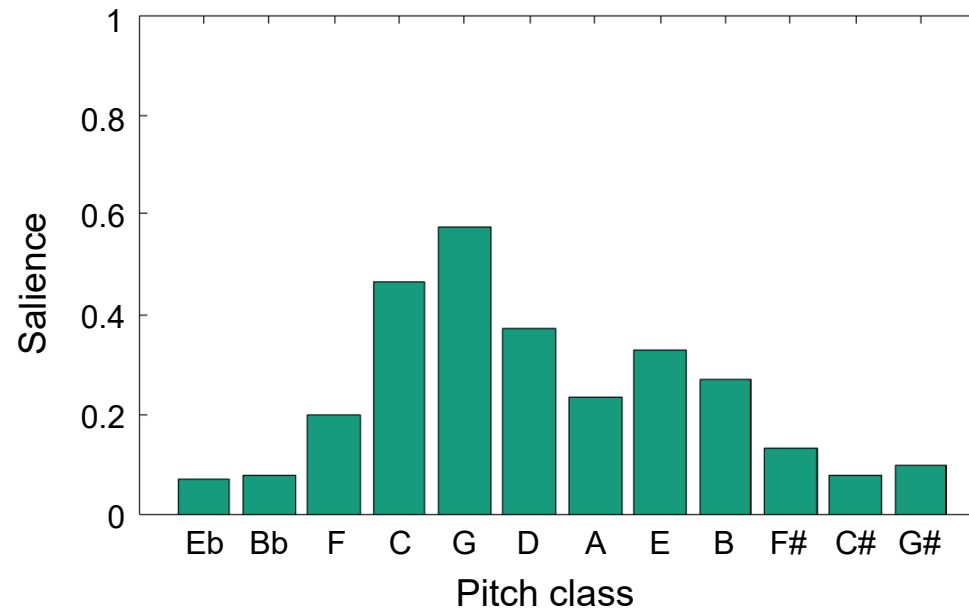


### 3. Machine Learning and Corpus Analyses in Classical Music and Jazz



# Tonal Complexity

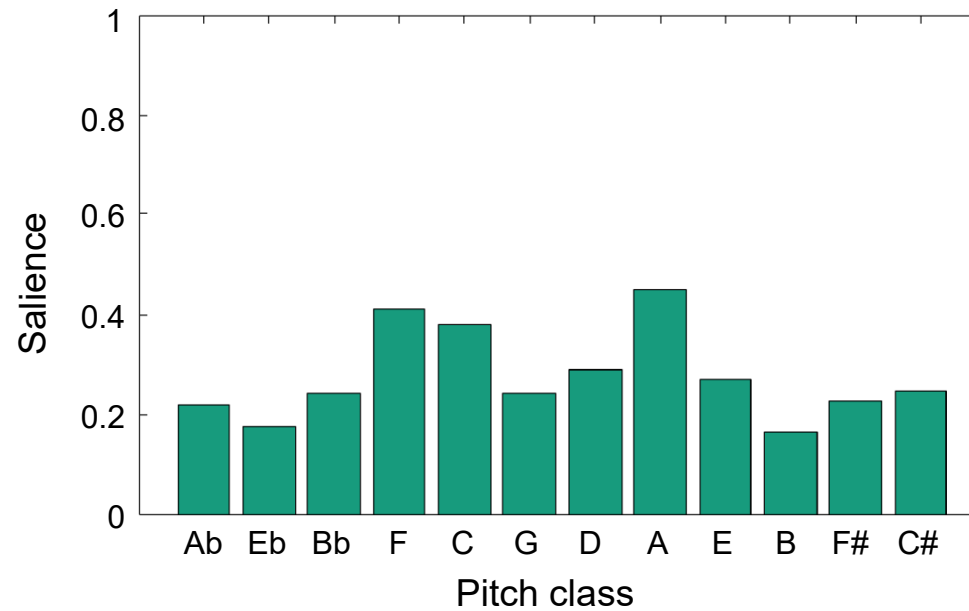
- Global chroma statistics (audio)
- **1783** – W. A. Mozart, „Linz“ symphony KV 425, 1. Adagio / Allegro (C major)



*Circle of fifths →*

# Tonal Complexity

- Global chroma statistics (audio)
- **1883** – J. Brahms, Symphony No. 3, 1. Allegro con brio (F major)

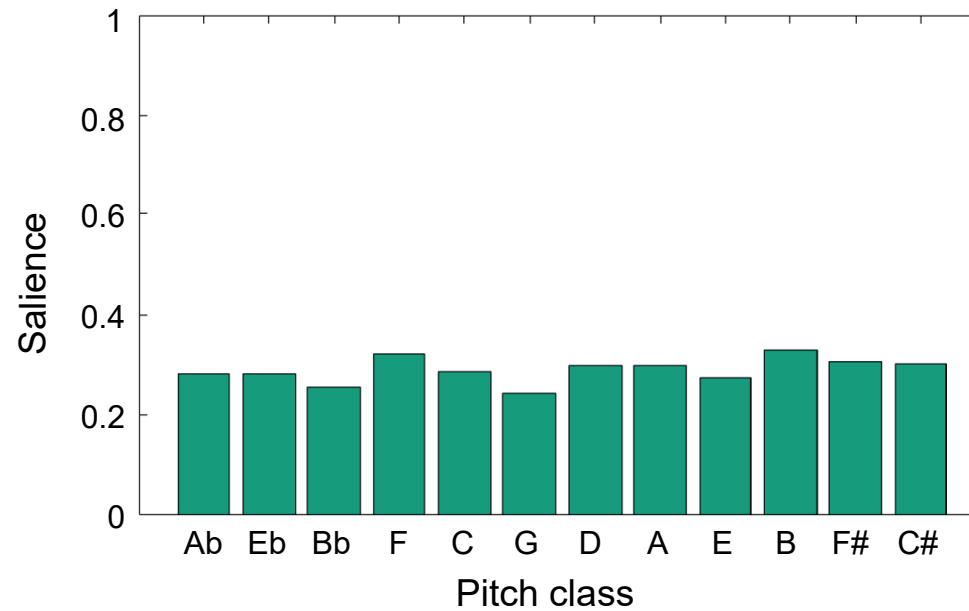


*Circle of fifths →*



# Tonal Complexity

- Global chroma statistics (audio)
- **1940** – A. Webern, Variations for Orchestra op. 30



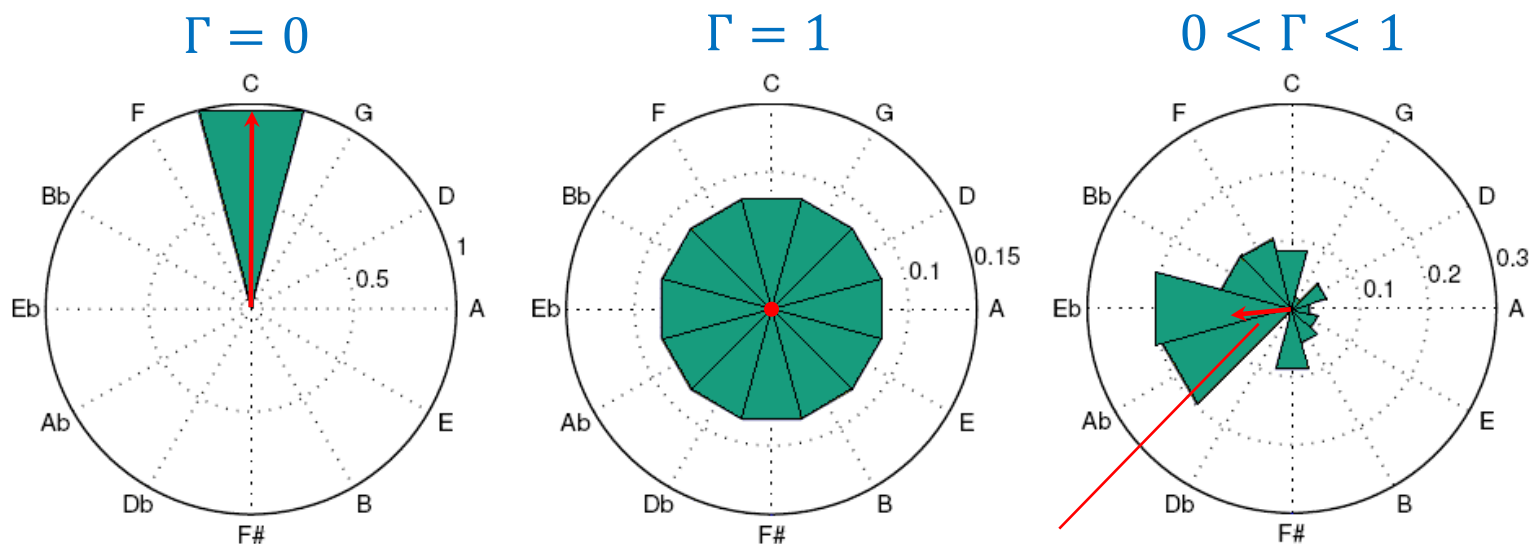
*Circle of fifths* →

# Tonal Complexity

- Realization of complexity measure  $\Gamma$

Entropy / Flatness measures

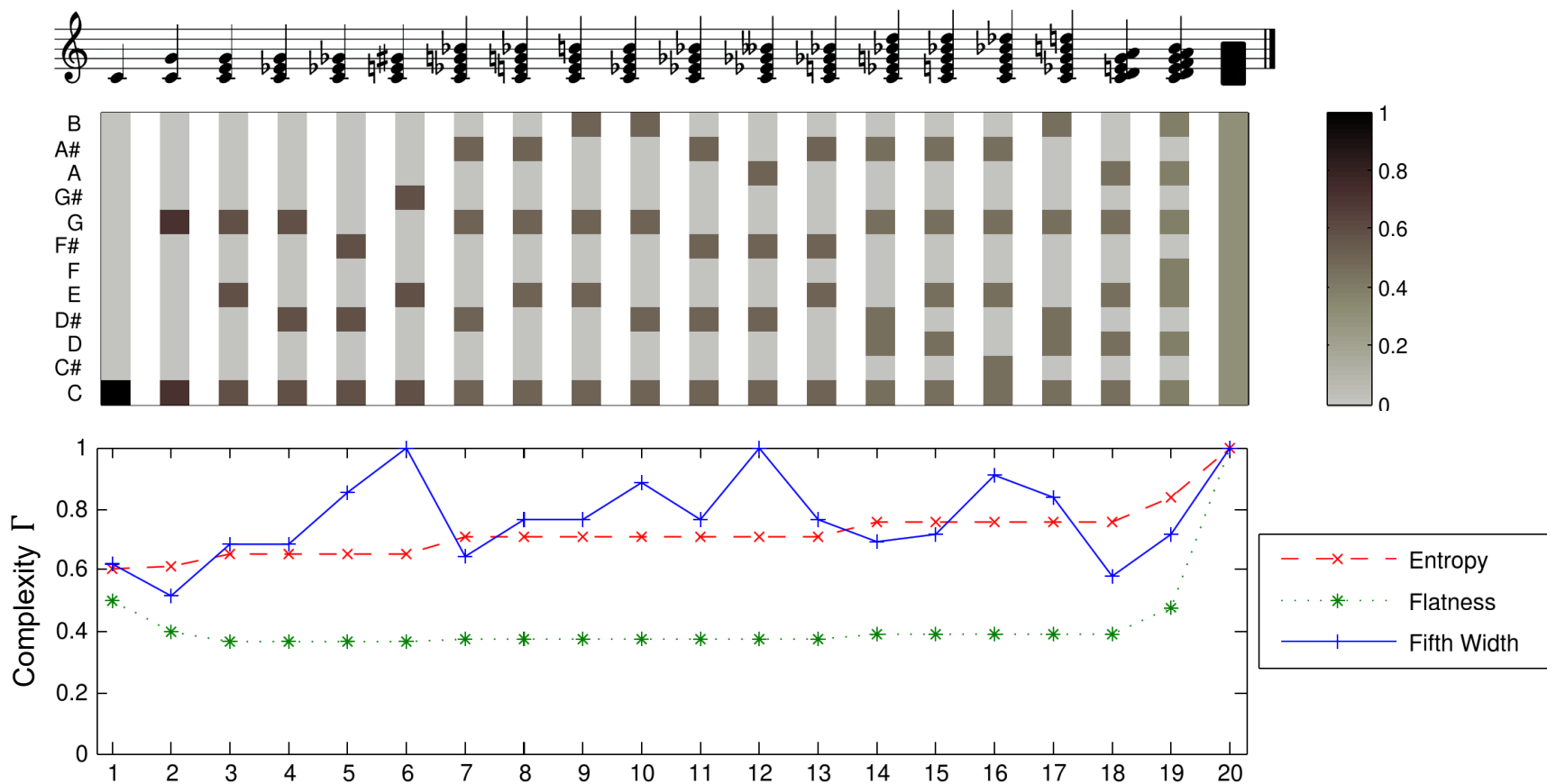
Distribution over *Circle of Fifths*



$$\Gamma = \sqrt{1 - r}$$

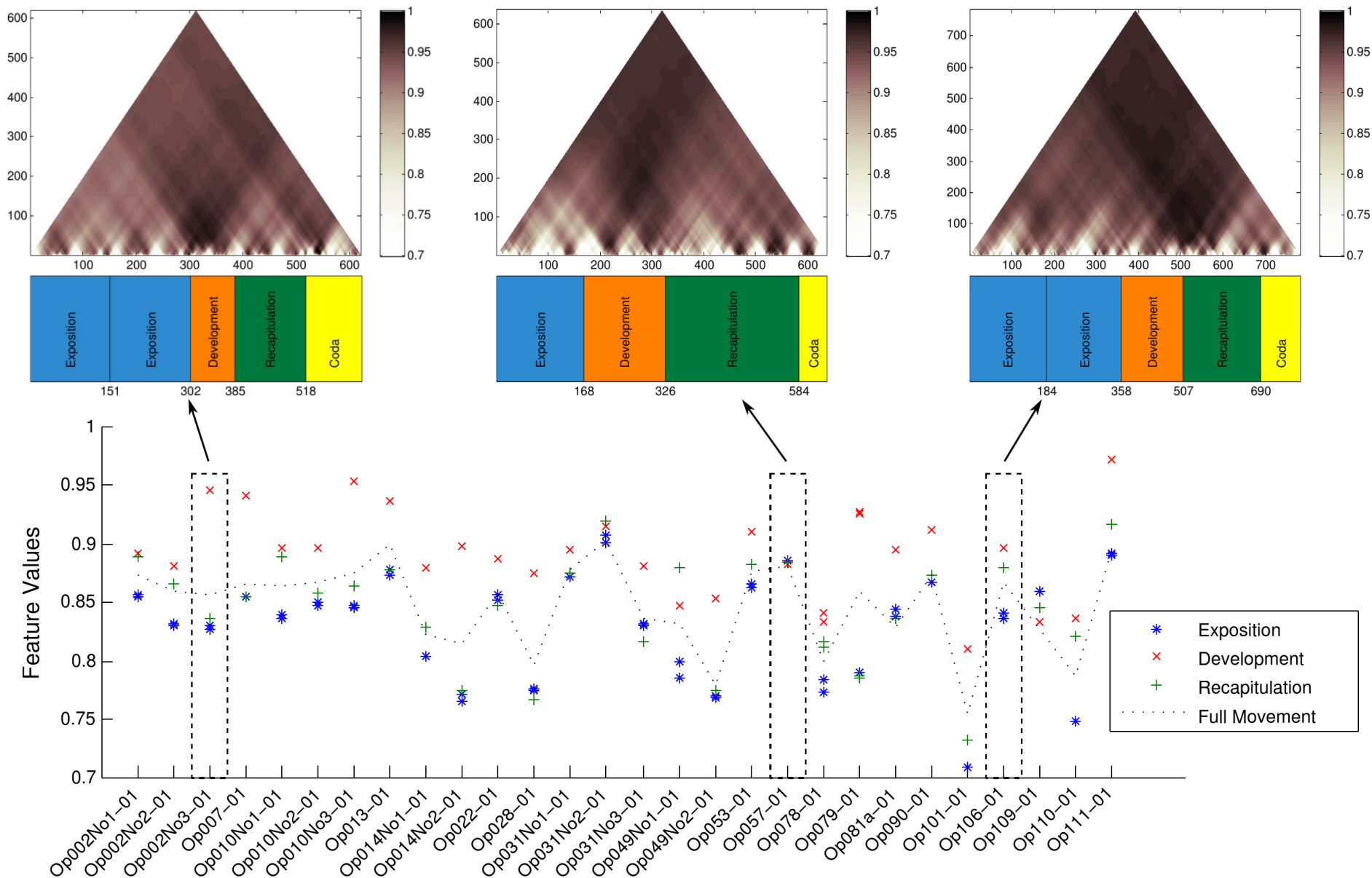
- Relating to different time scales!

# Tonal Complexity – Chords



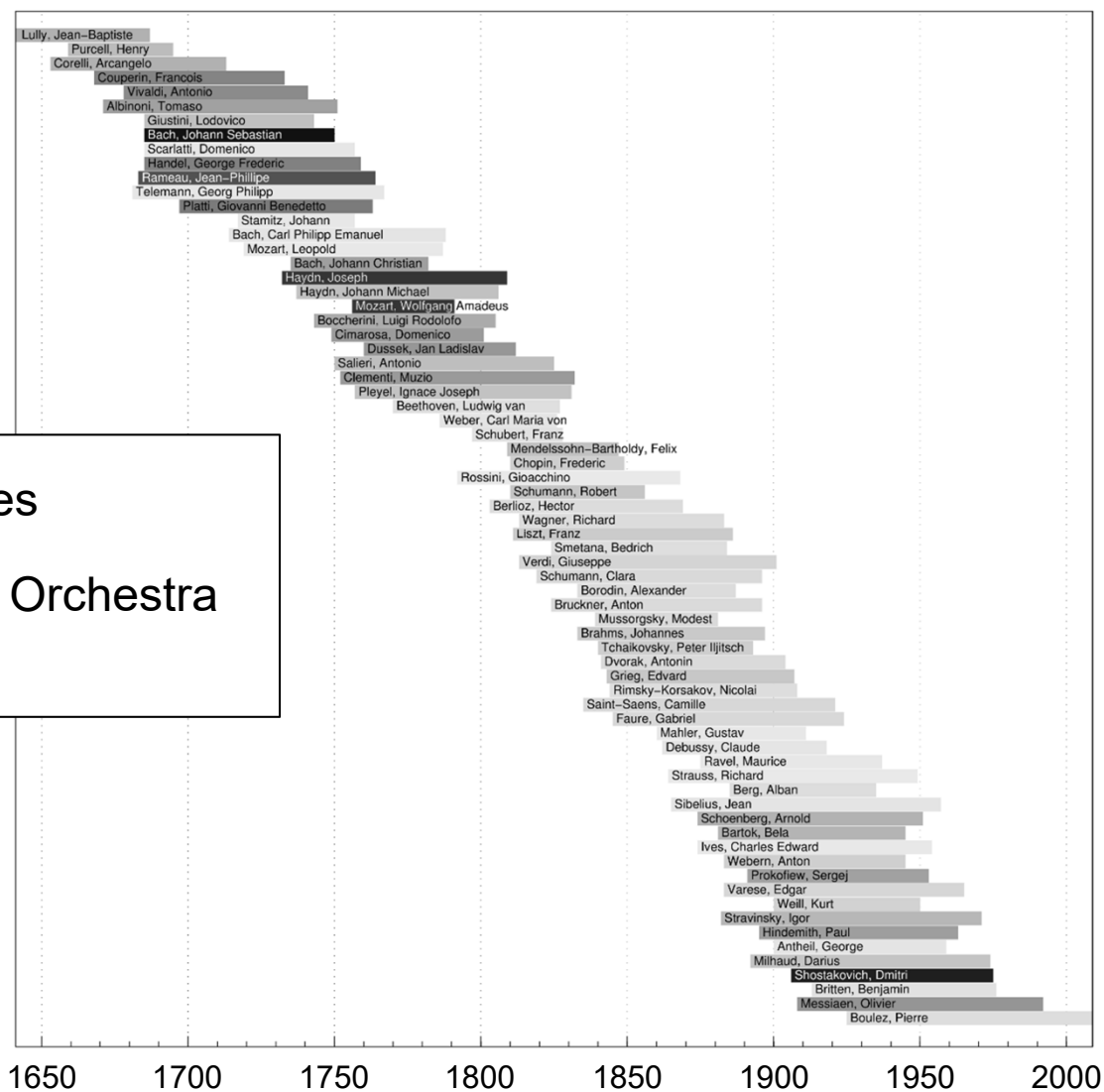
[8] Weiss / Müller, *Quantifying and Visualizing Tonal Complexity*, CIM 2014

# Tonal Complexity – Beethoven's Sonatas

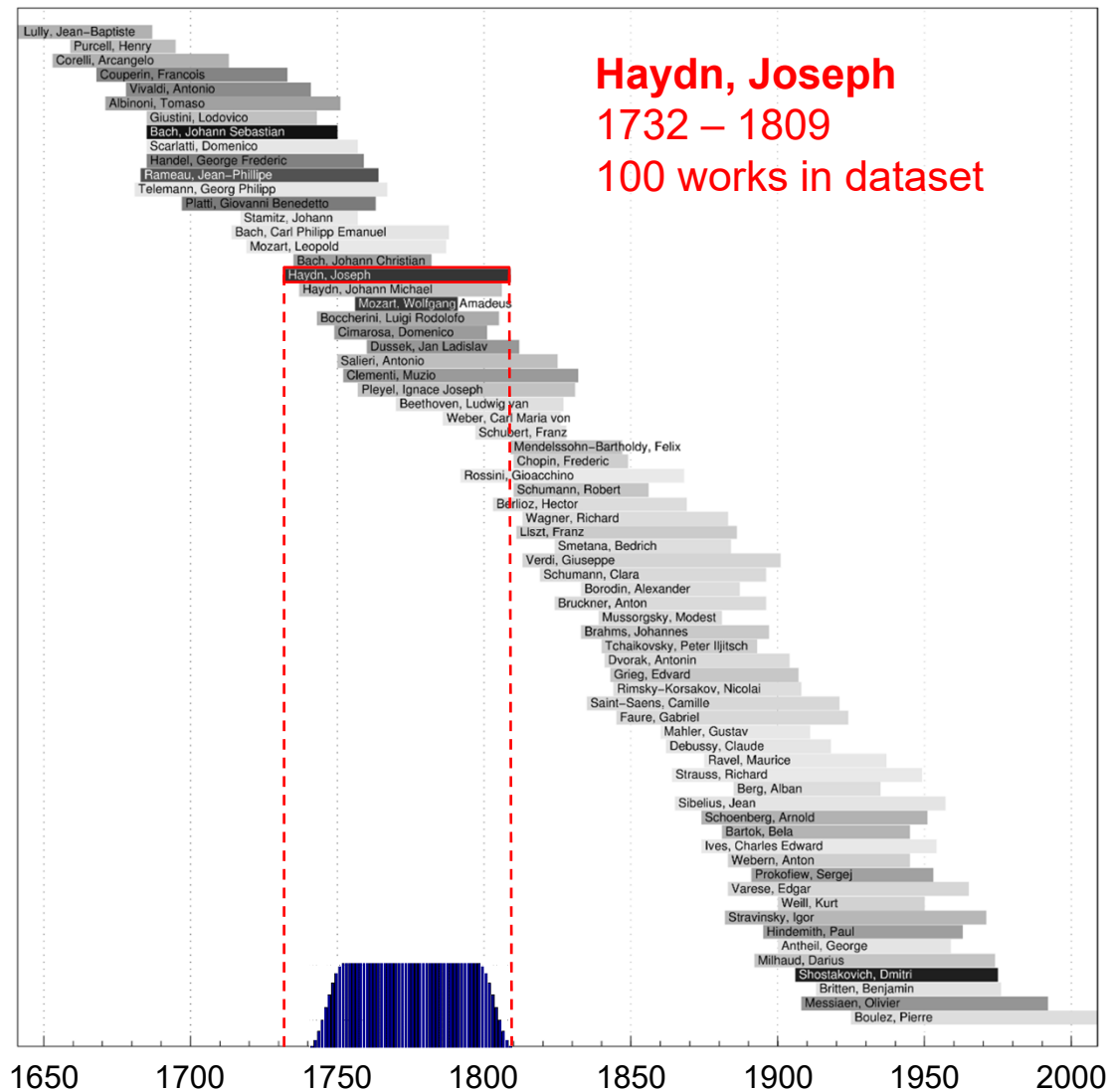


# Analyzing Composer Styles

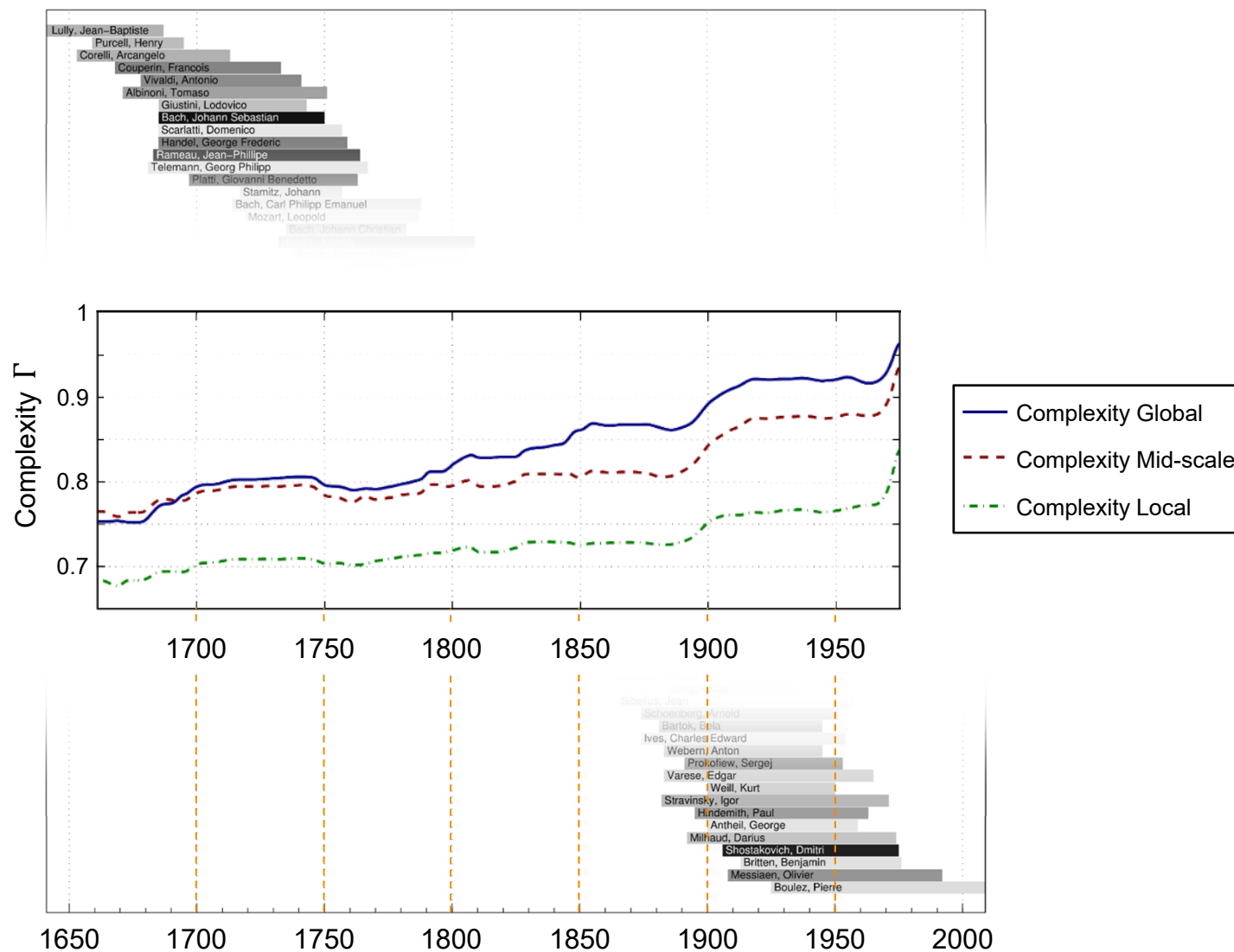
- 2000 pieces
- Piano and Orchestra music



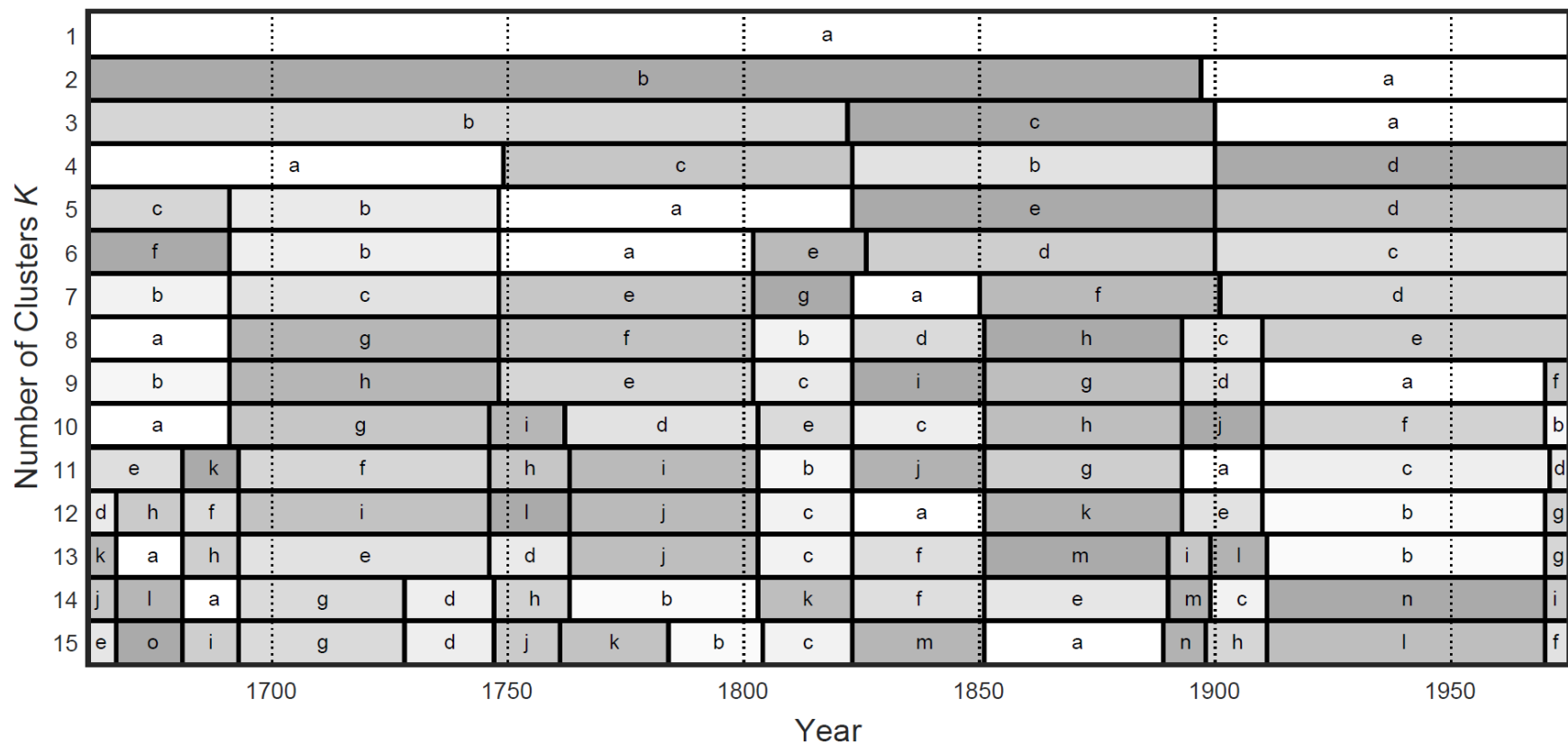
# Analyzing Composer Styles



# Analyzing Composer Styles

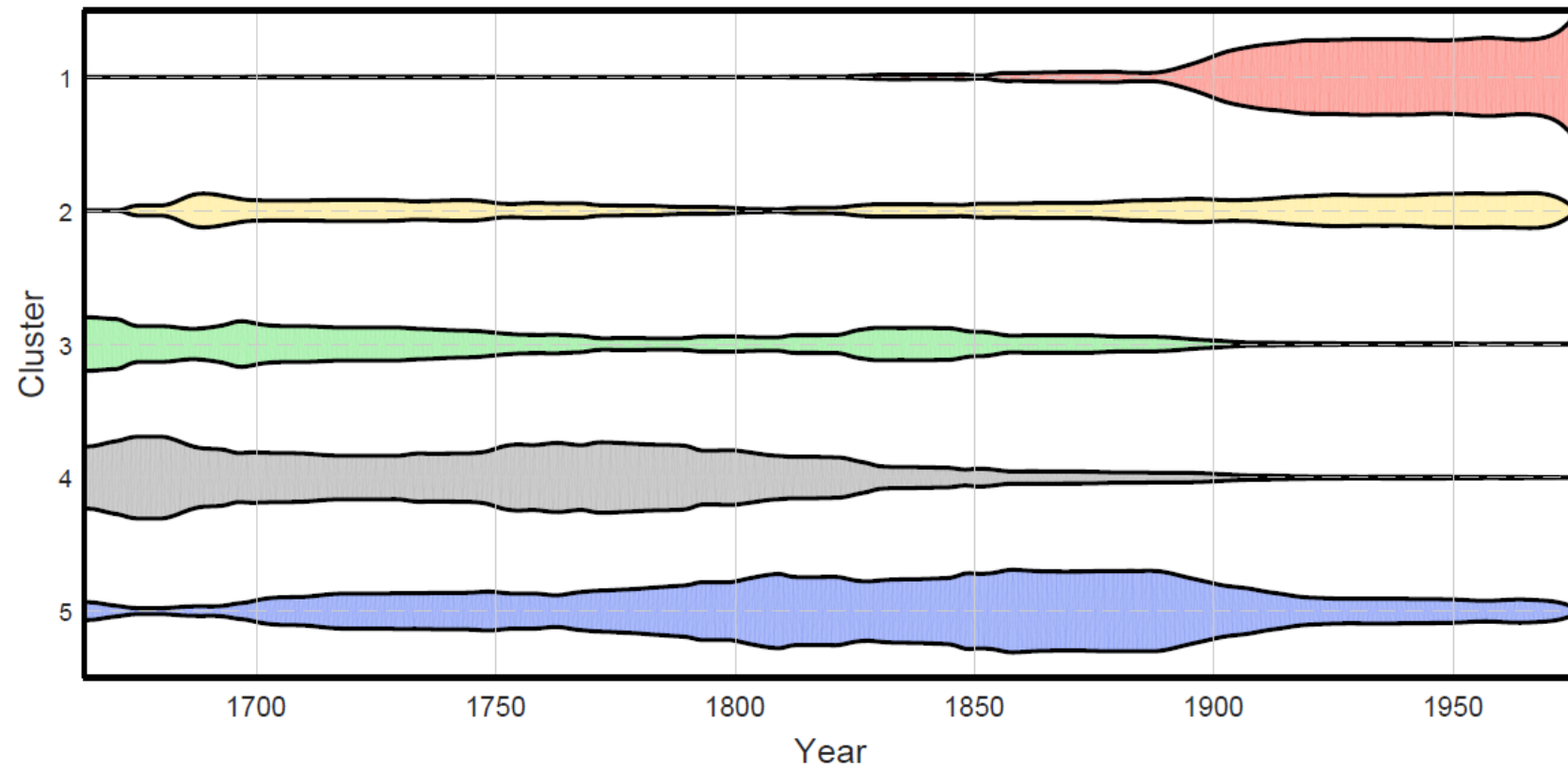


# Clustering Composition Years

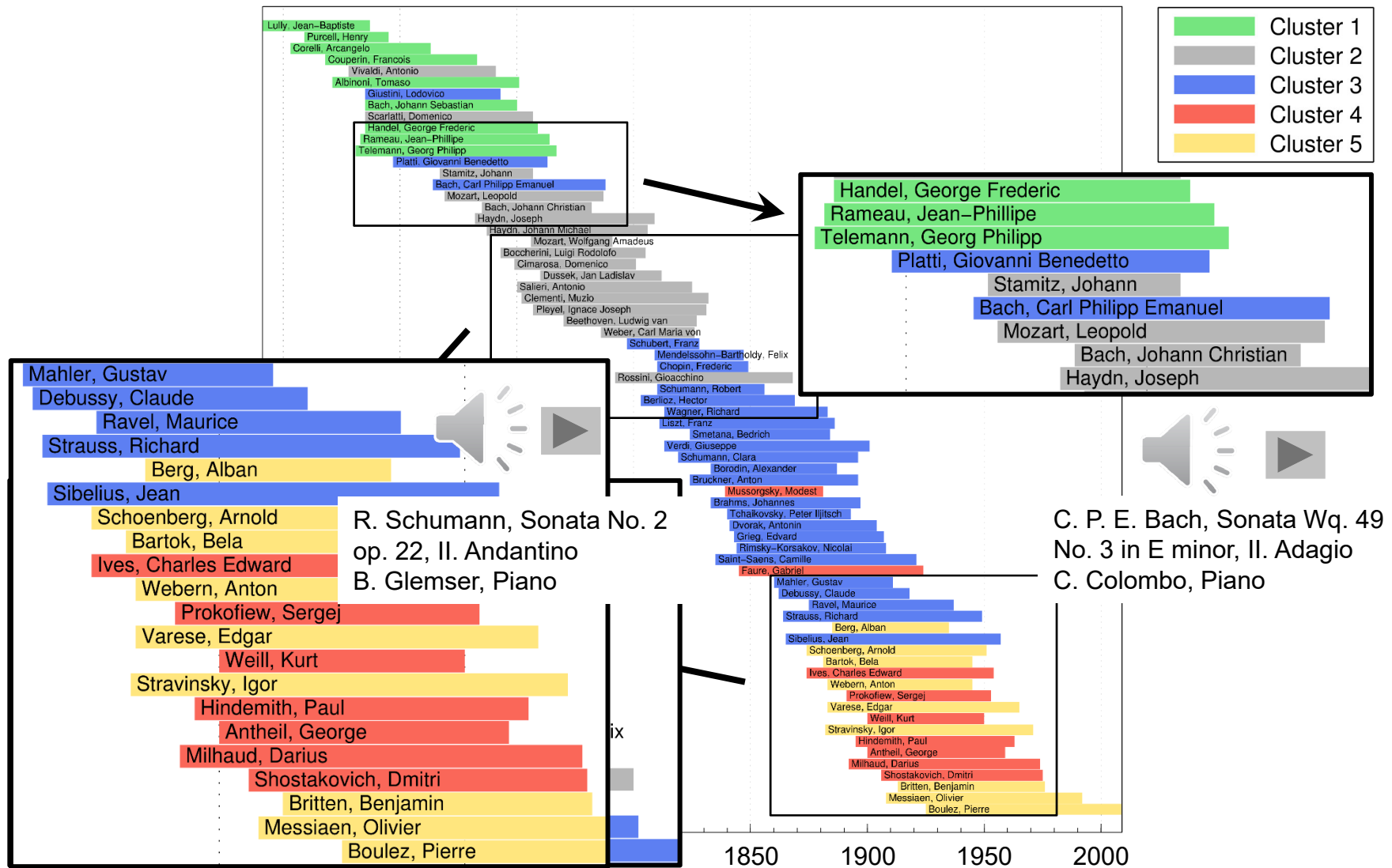


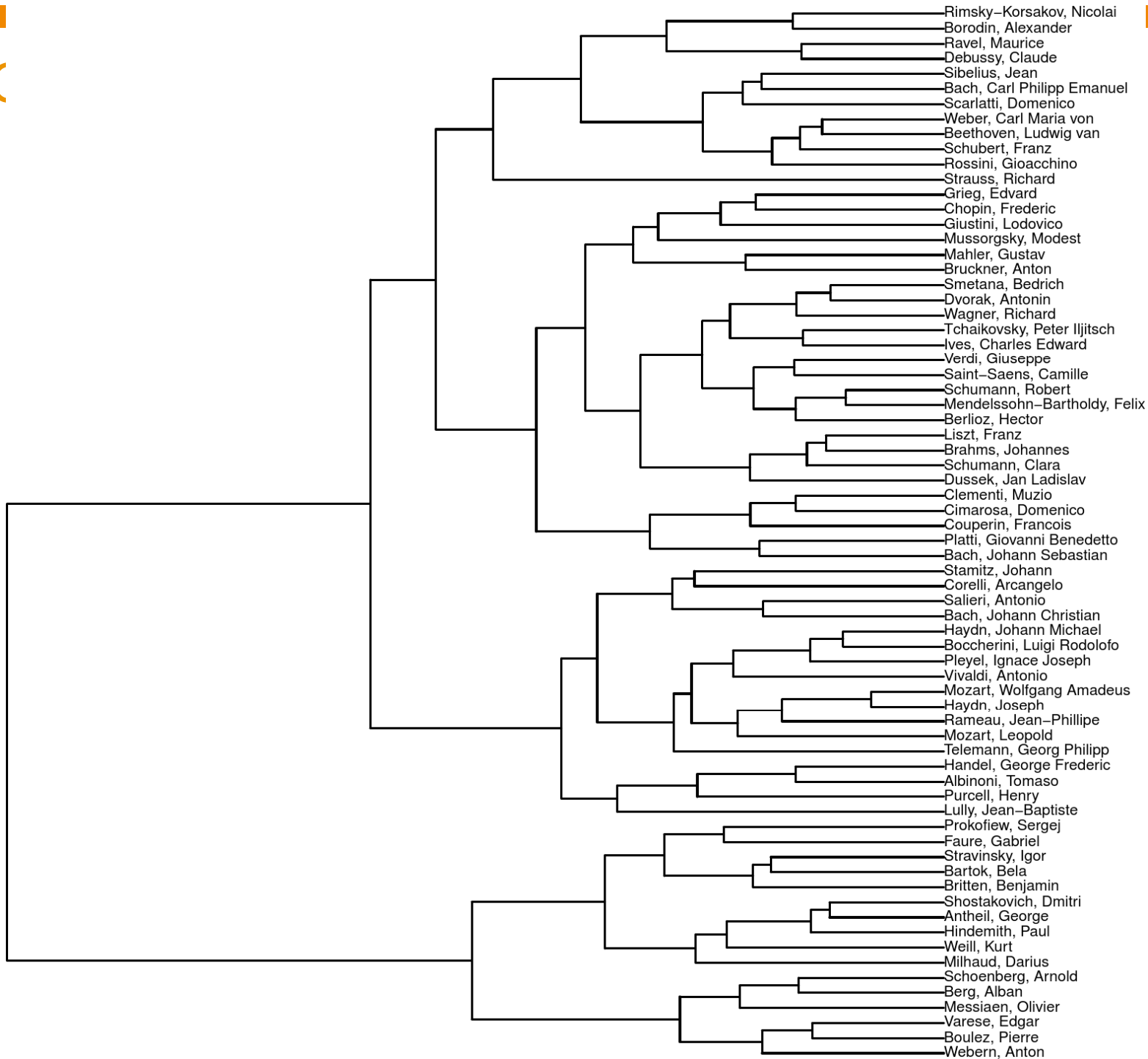


# Clustering Individual Pieces



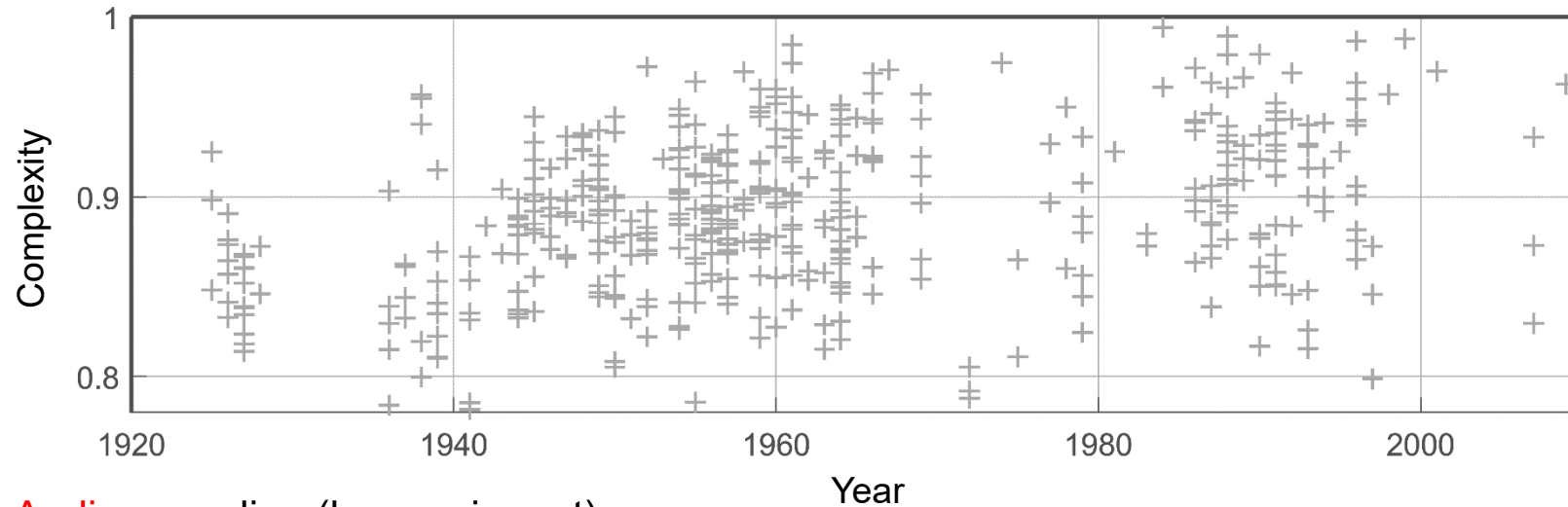
# Clustering Composers



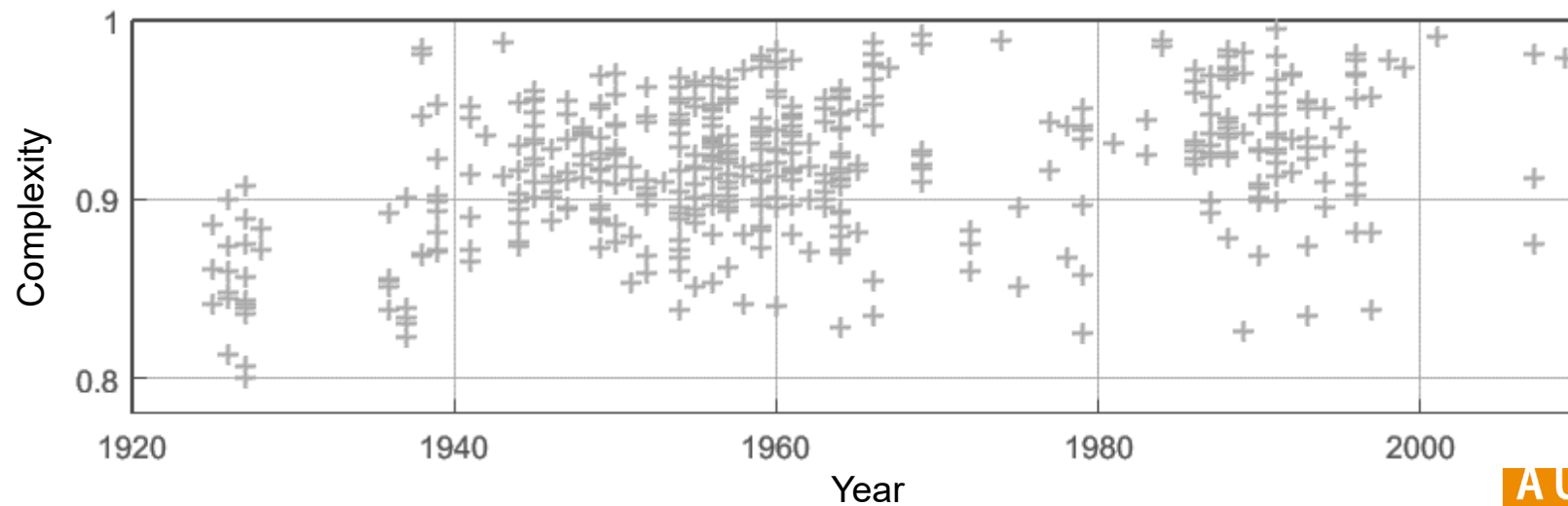


# Tonal Complexity: Jazz Solos

- **Symbolic** transcription

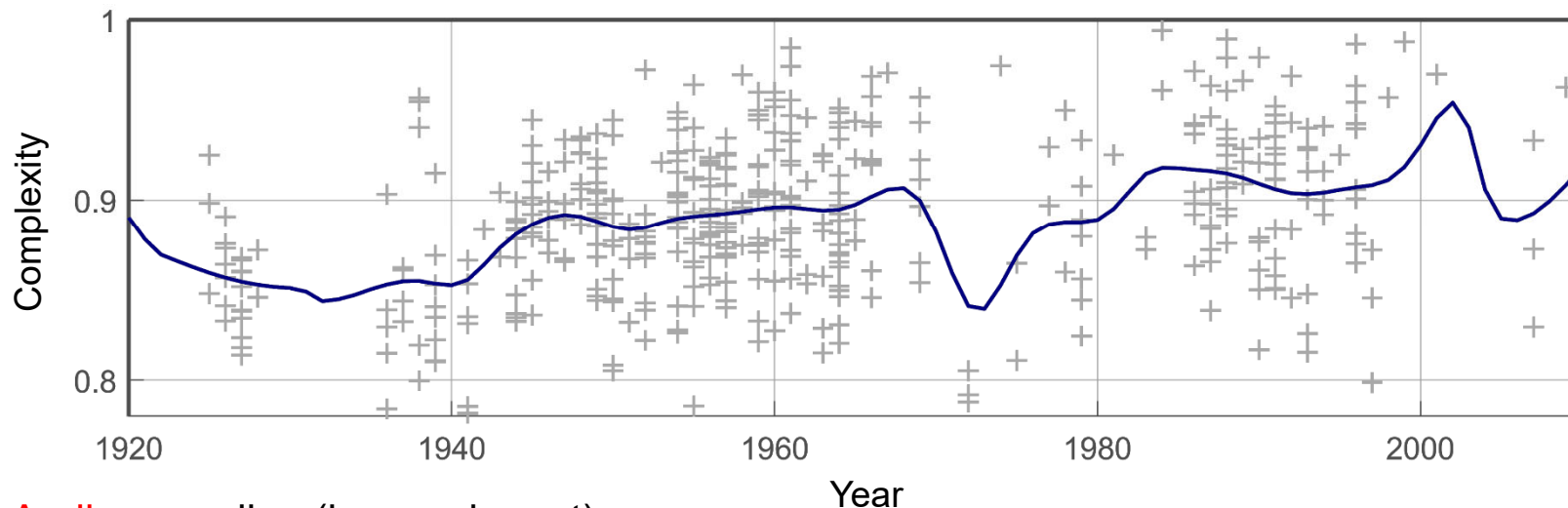


- **Audio** recording (harmonic part)

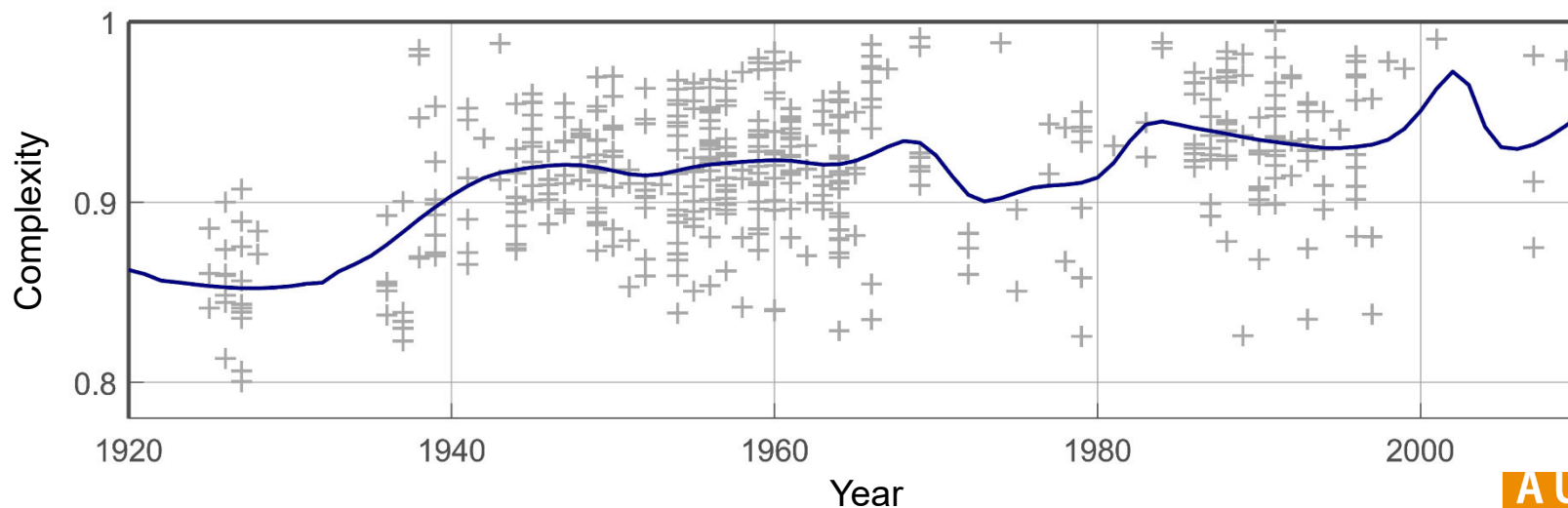


# Tonal Complexity: Jazz Solos

- **Symbolic** transcription

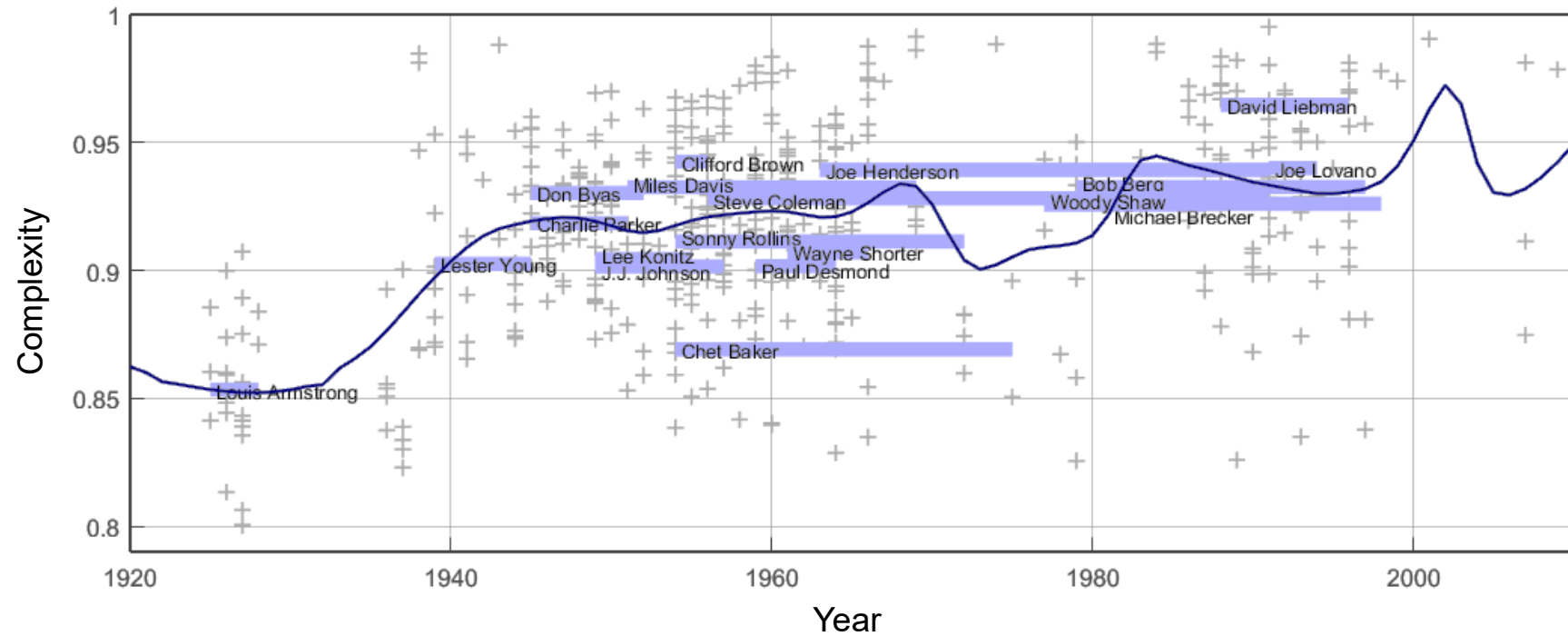


- **Audio** recording (harmonic part)



# Tonal Complexity: Jazz Solos

- Audio recording (harmonic part)



# Music Genre Classification

world music *JAZZ*  
*HipHop* **pop** **Rock**  
"classical"

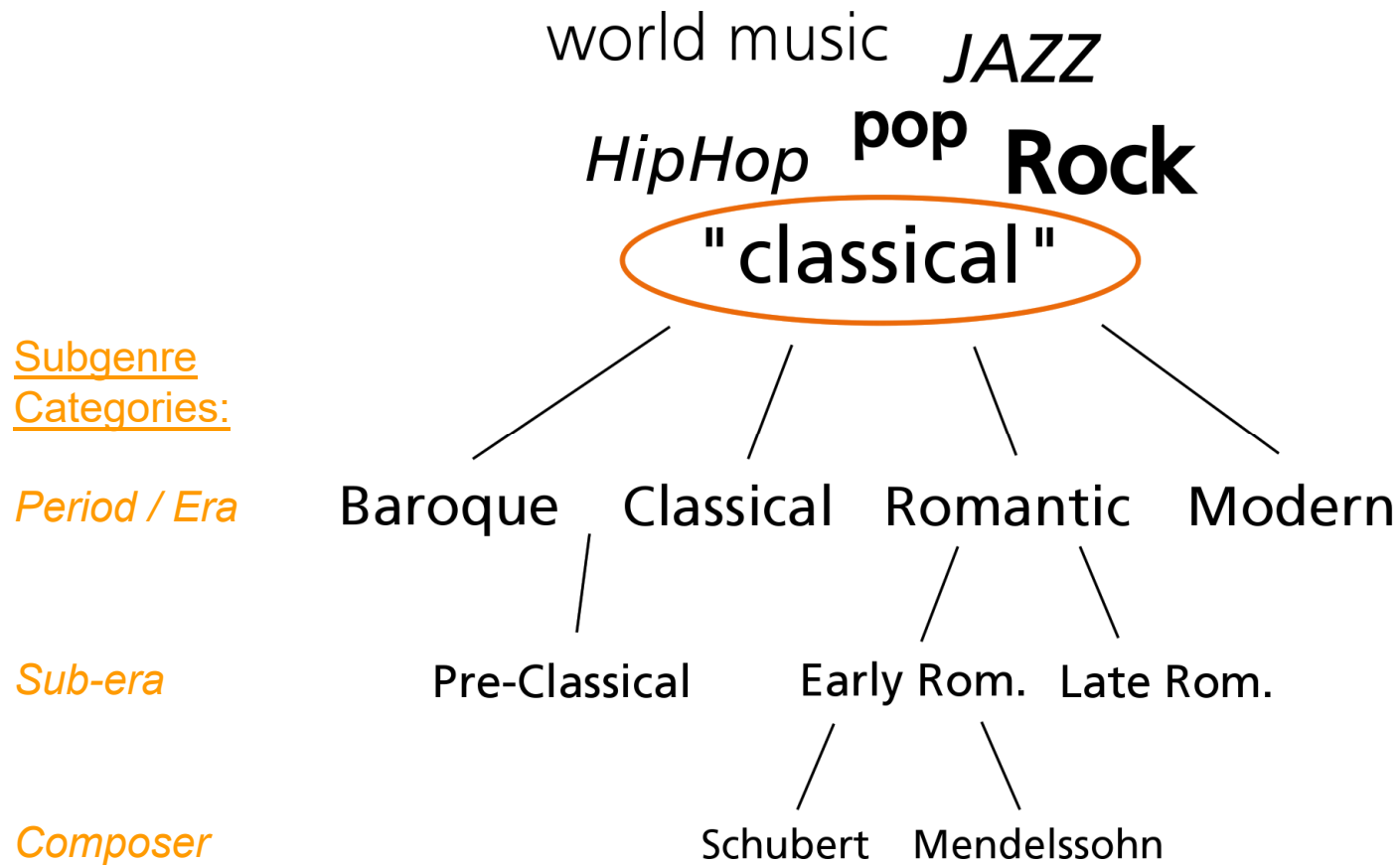
J. S. Bach,  
Brandenburg Concerto  
No. 2 in F major, I. Allegro,  
Cologne Chamber Orch.

L. van Beethoven,  
*Fidelio*, Overture,  
Slovak Philharm.

R. Schumann,  
Sonata No. 2 op. 22,  
II. Andantino  
B. Glemser, Piano

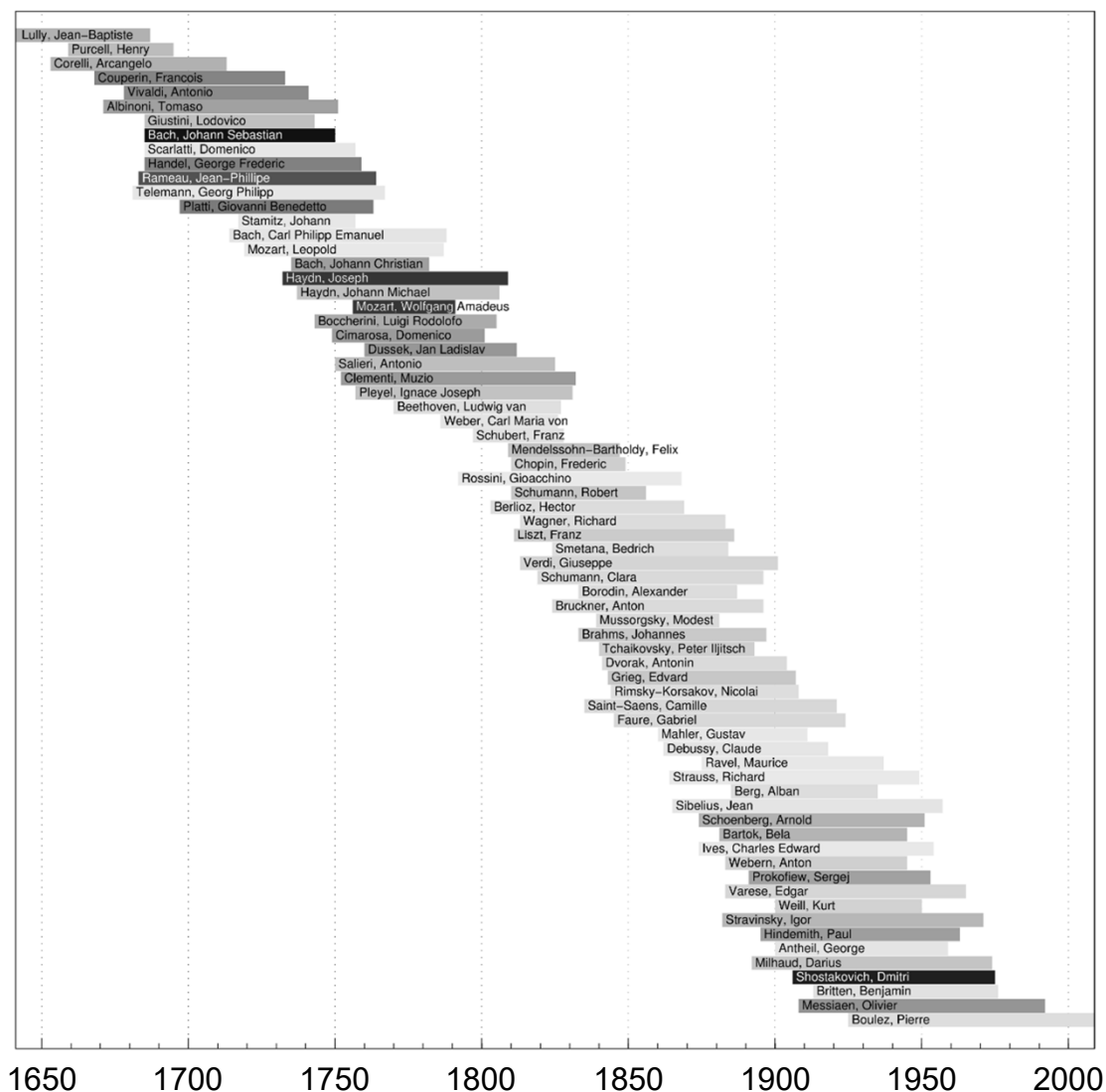
A. Webern,  
Variations for Orchestra op. 30  
Ulster Orchestra

# Music Genre Classification





# Music Genre Classification: Dataset

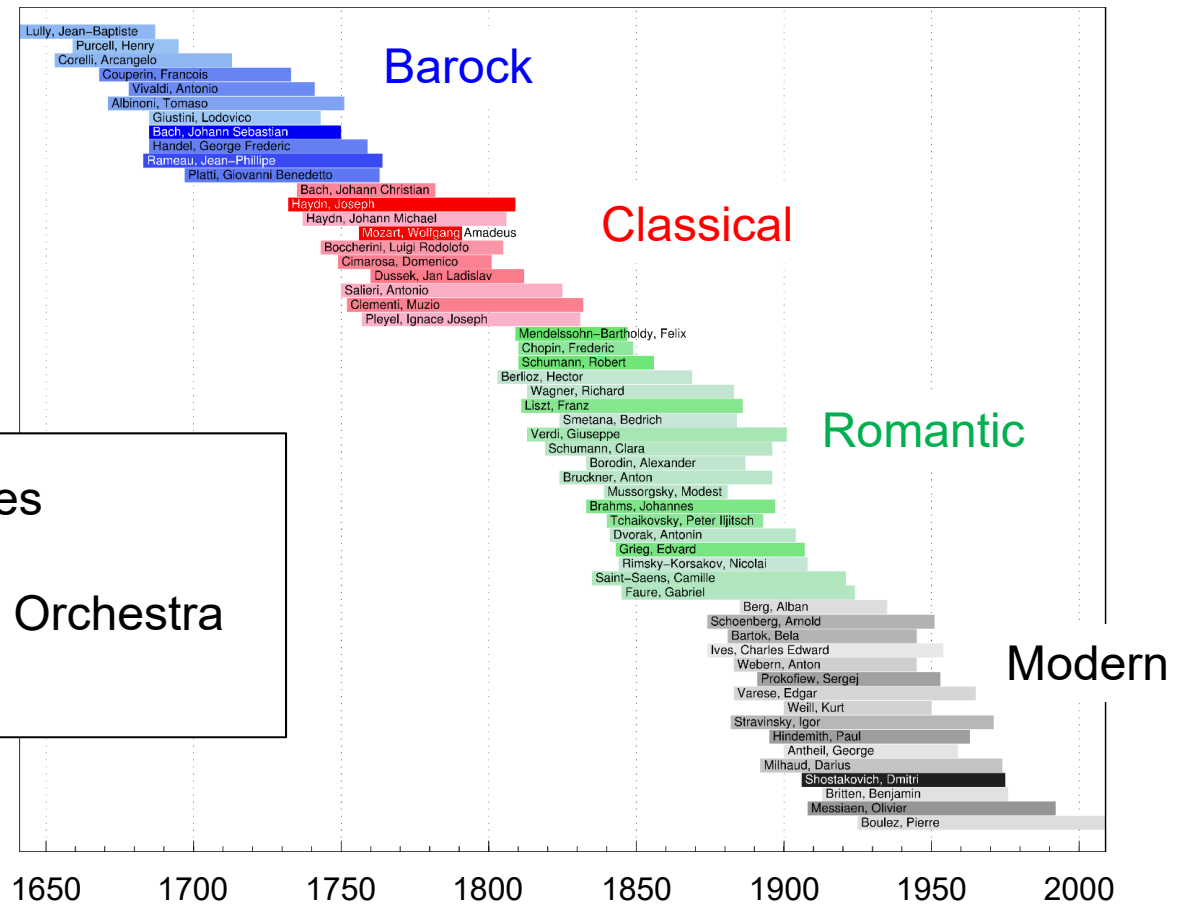


# Music Genre Classification: Dataset



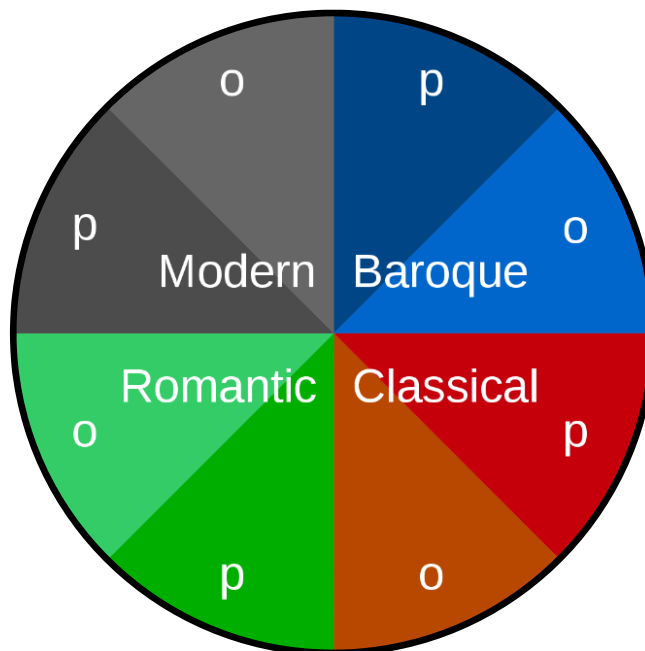
# Music Genre Classification: Dataset

- 1600 pieces
- Piano and Orchestra music



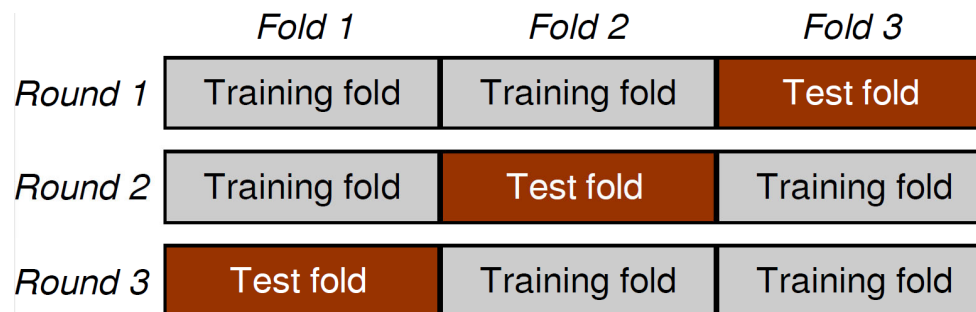
# Music Genre Classification: Dataset

- Dataset: *CrossEraDB* (Historical Periods)
  - Balanced Piano (p) – Orchestra (o)
  - Each 200 pieces → 1600 in total

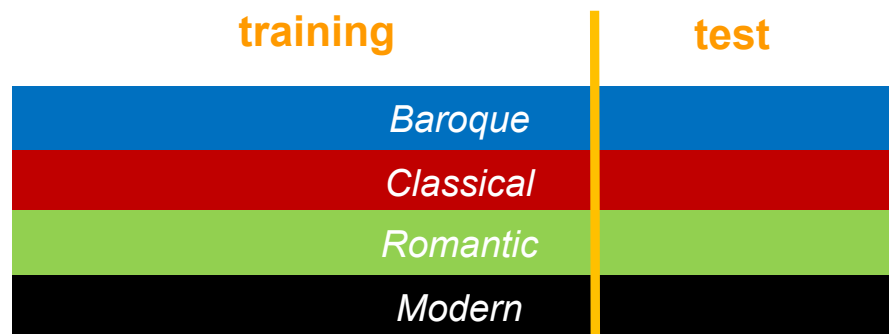


# Music Genre Classification: Dataset

- Experimental design: Evaluation with Cross Validation (CV)
- Separate data into different parts (*folders*)

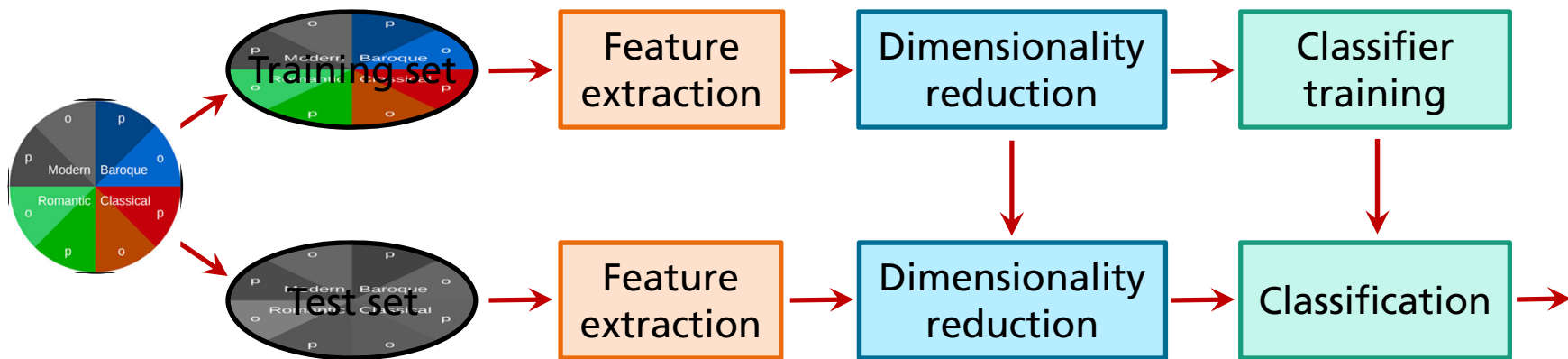


- Distribution of classes balanced for all folds

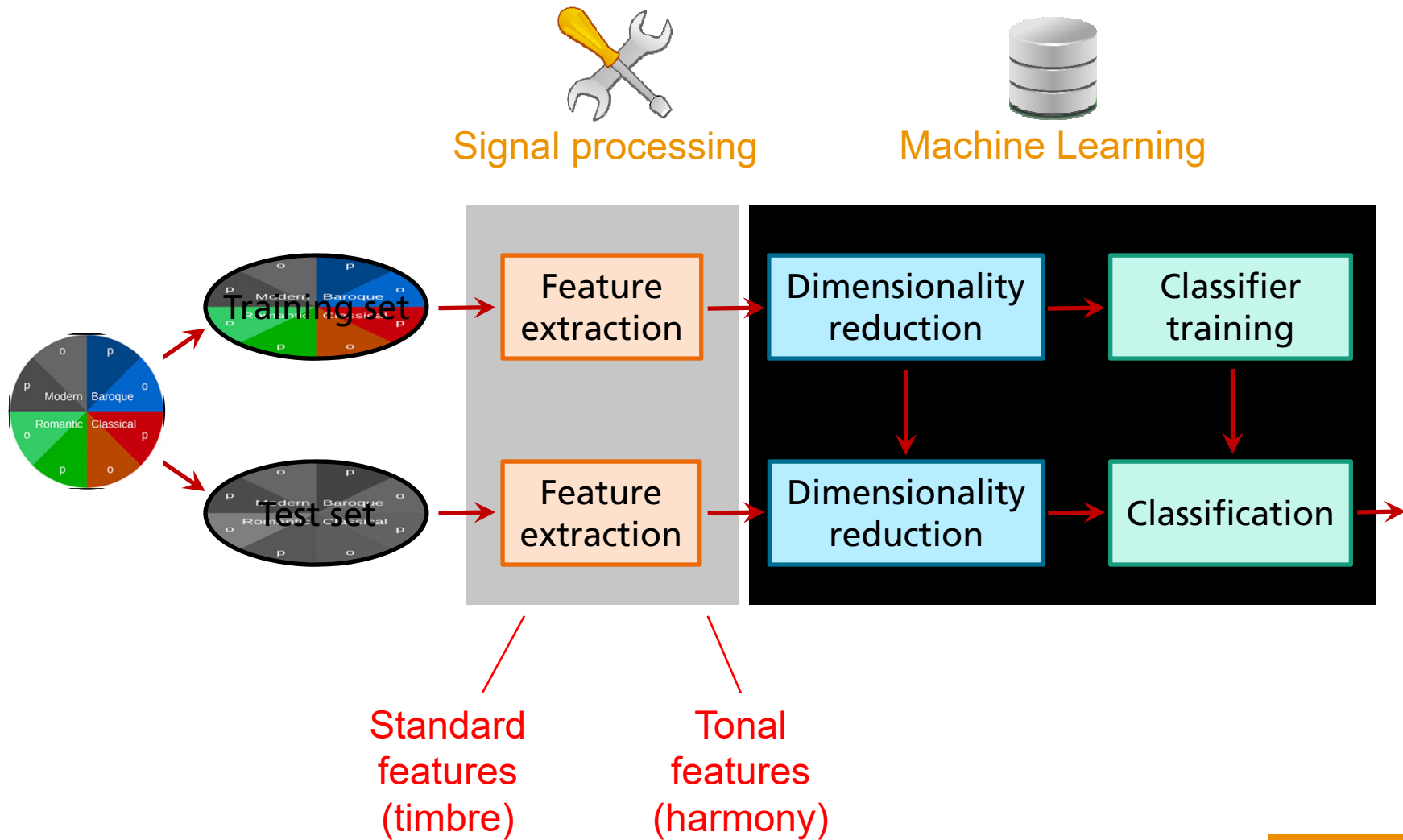


# Classification Pipeline

- Typical approach: Supervised machine learning



# Classification Pipeline



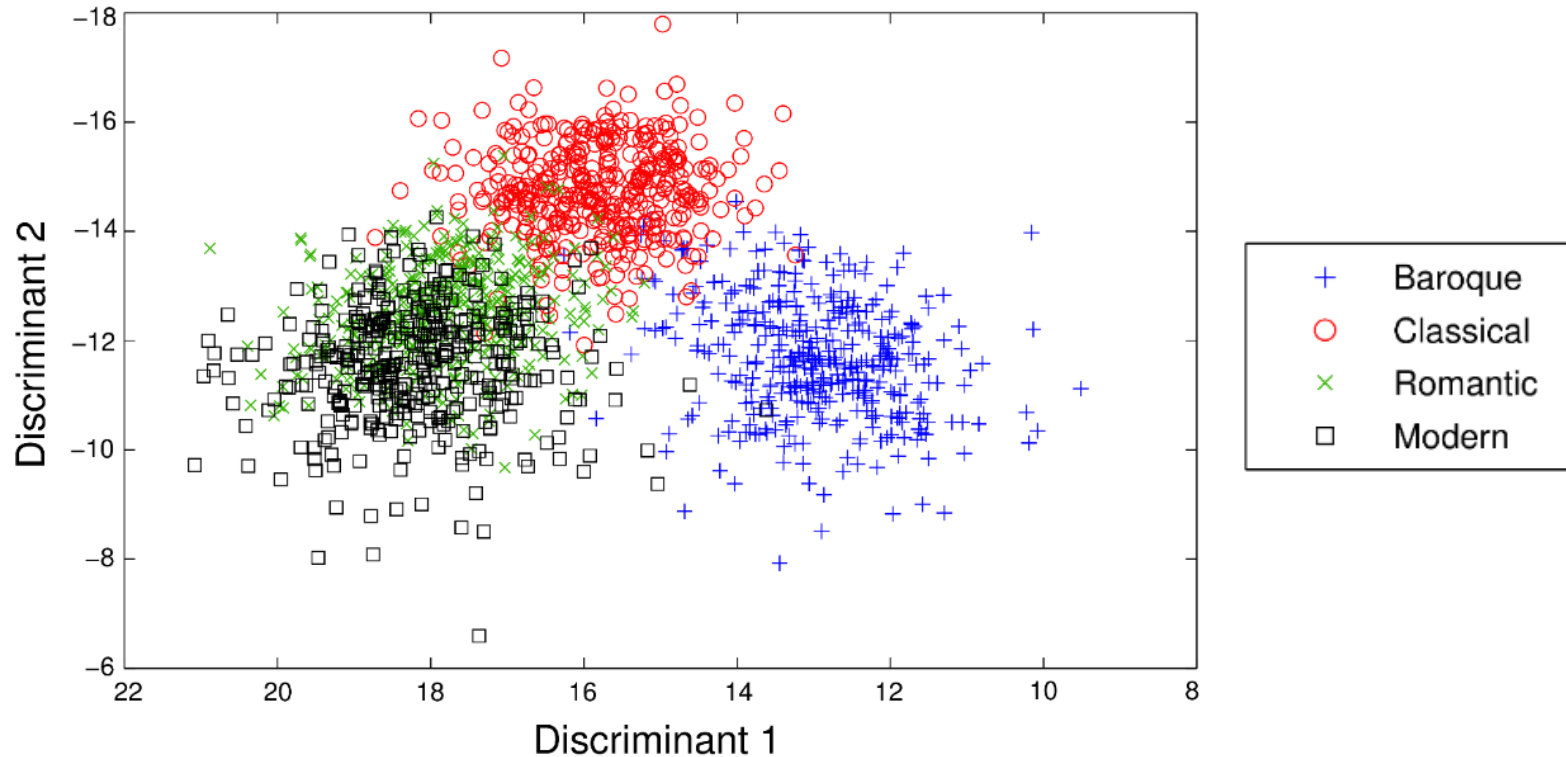
# Classification Features

Standard	Dim.	Tonal	Dim.
MFCC	16	Interval cat.	6 x 4
OSC	14	Triad types	4 x 4
ZCR	1	Complexity	7 x 4
ASE	16	Chord progr.	11 x 5
SFM	16		
SCF	16		
SC	16		
LogLoud	12		
NormLoud	12		
Sum	119	Sum	123
Mean & Std	x 2	Mean & Std	x 2
<b>Total</b>	<b>238</b>	<b>Total</b>	<b>246</b>



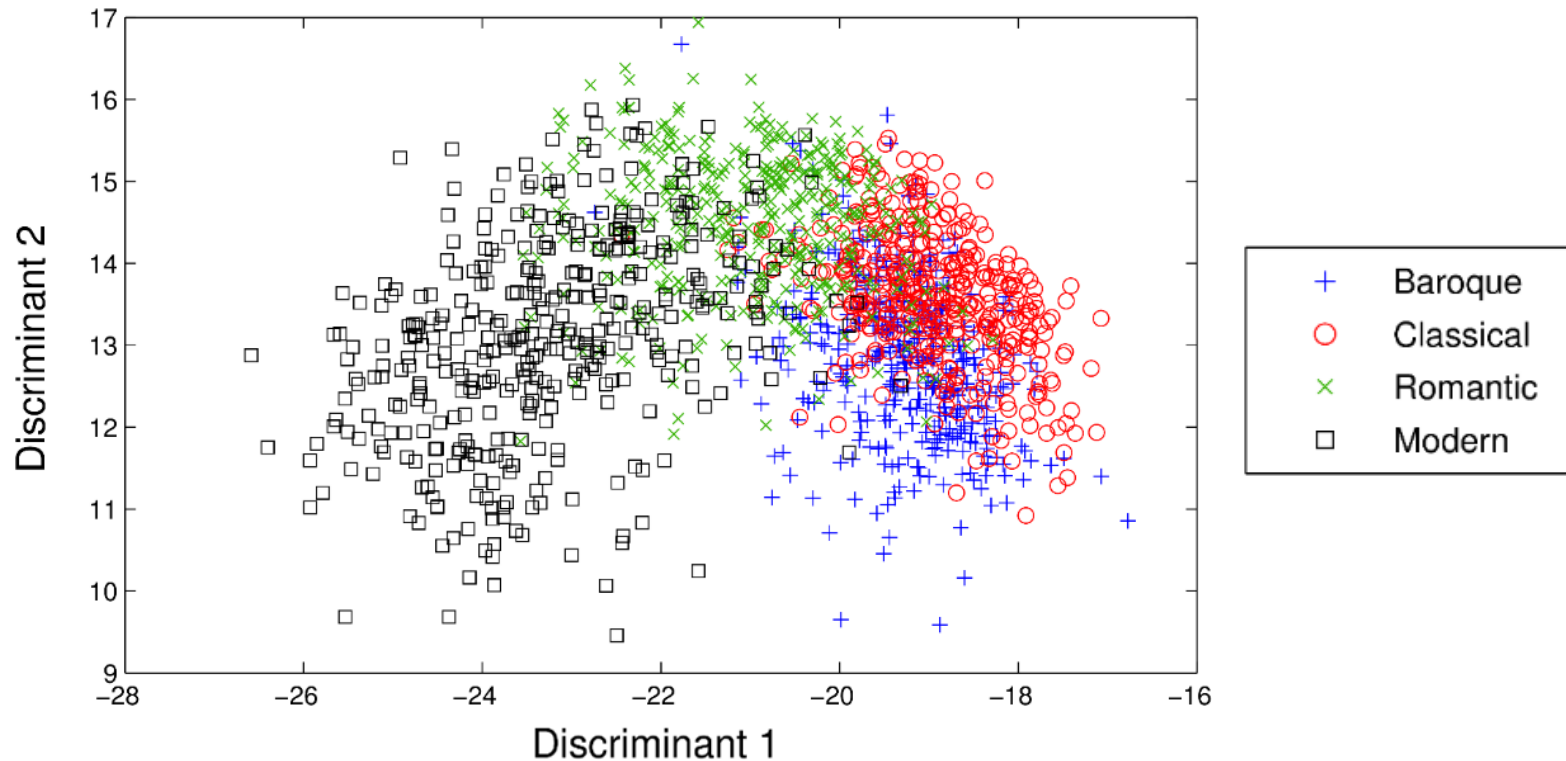
# Dimensionality Reduction

- Reduce feature space to few dimensions (prevent **curse of dimensionality**)
- Maximize separation of classes with **Linear Discriminant Analysis (LDA)**
- Using **standard features** (MFCC, spectral envelope, ...)



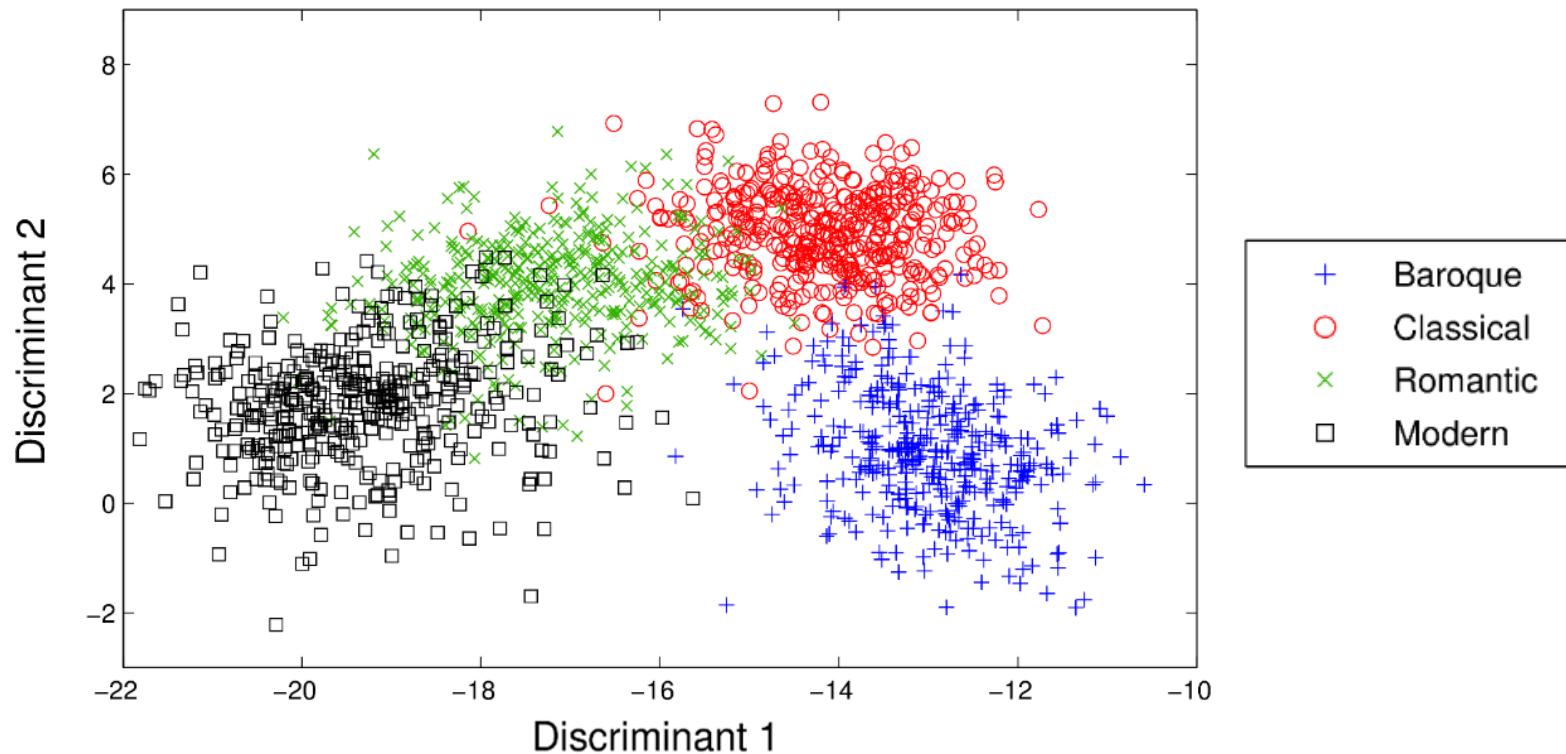
# Dimensionality Reduction

- Reduce feature space to few dimensions
- Maximize separation of classes with **Linear Discriminant Analysis (LDA)**
- Using **tonal features** (interval, triad types, tonal complexity, ... 4 time scales)



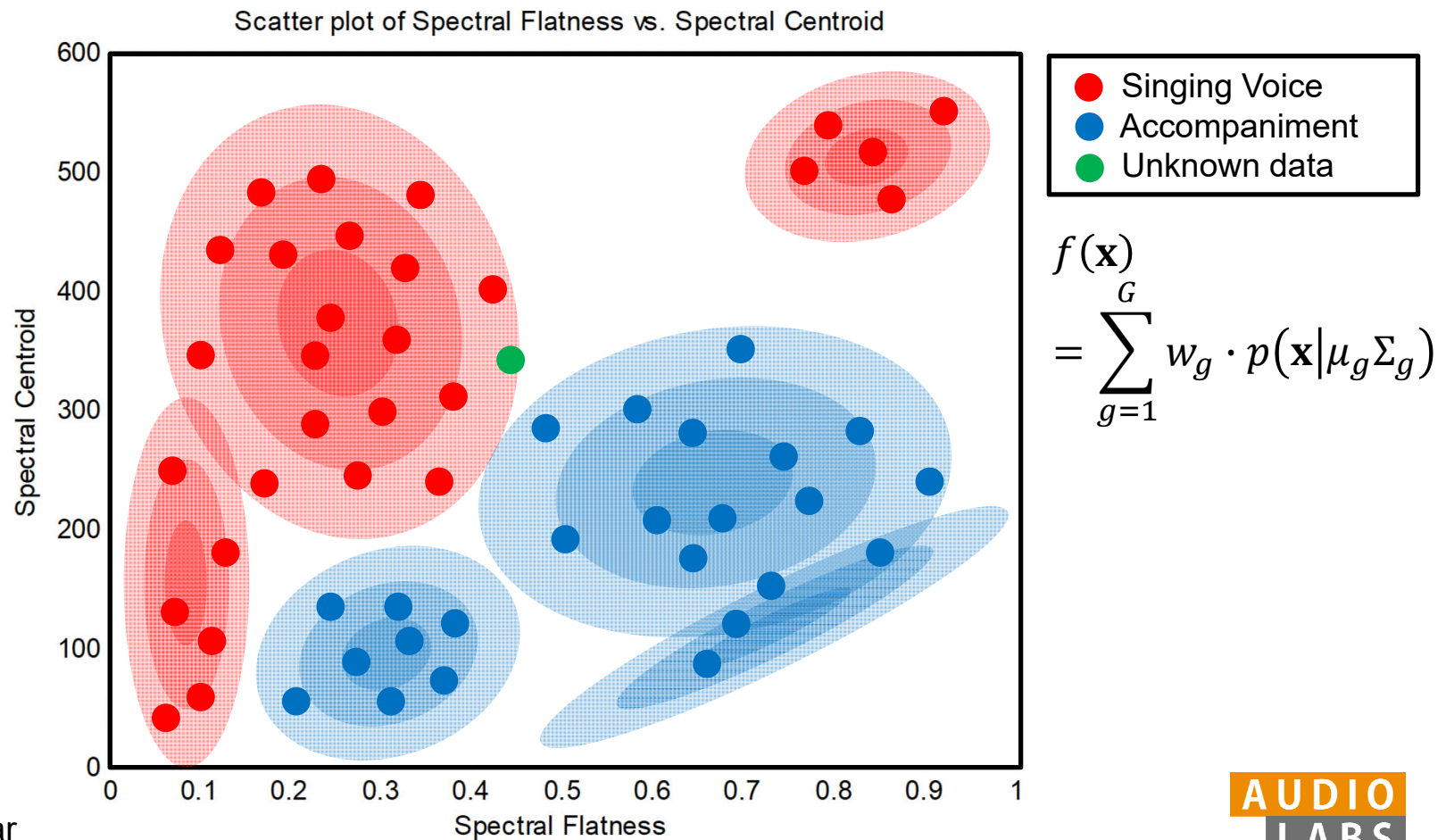
# Dimensionality Reduction

- Reduce feature space to few dimensions
- Maximize separation of classes with **Linear Discriminant Analysis (LDA)**
- Using **tonal & standard features**



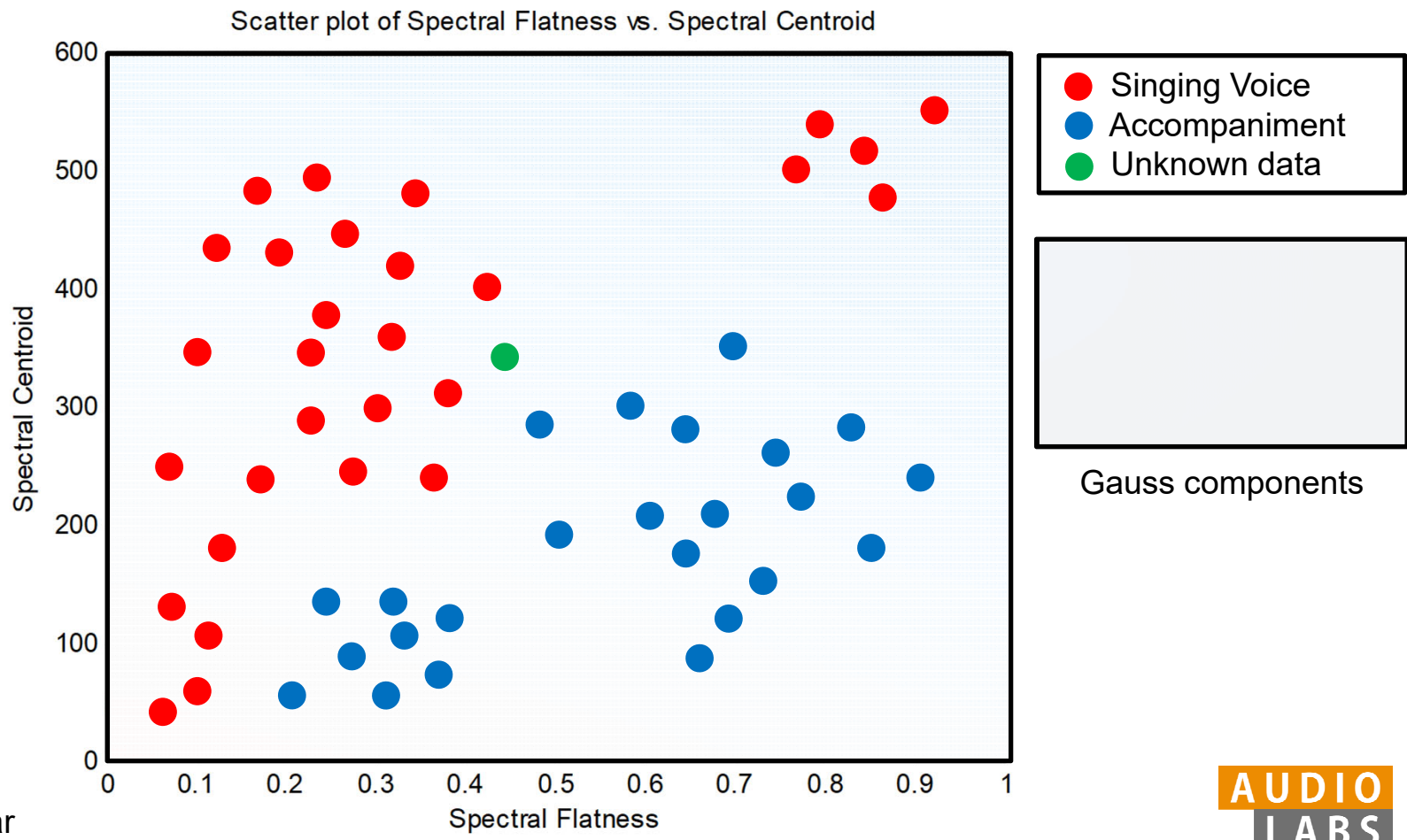
# Classifier

- Gaussian Mixture Models (GMM)



# Classifier

- Gaussian Mixture Models (GMM)



# Classification Results

- Gaussian Mixture Model (GMM) classifier, LDA reduction, 3-fold cross validation

	Full Dataset	Piano	Orchestra
<i>Standard features</i>	87 %	88 %	85 %
<i>Tonal features</i>	84 %	84 %	86 %
<b><i>Combined</i></b>	<b>92 %</b>	<b>86 %</b>	<b>80 %</b>

Weiss / Mauch / Dixon, *Timbre-Invariant Audio Features for Style Analysis of Classical Music*, ICMC / SMC 2014

# Classification Results

- Gaussian Mixture Model (GMM) classifier, LDA reduction, 3-fold cross validation

	Full Dataset	Piano	Orchestra
<i>Standard features</i>	87 %	88 %	85 %
<i>Tonal features</i>	84 %	84 %	86 %
<b>Combined</b>	<b>92 %</b>	<b>86 %</b>	<b>80 %</b>

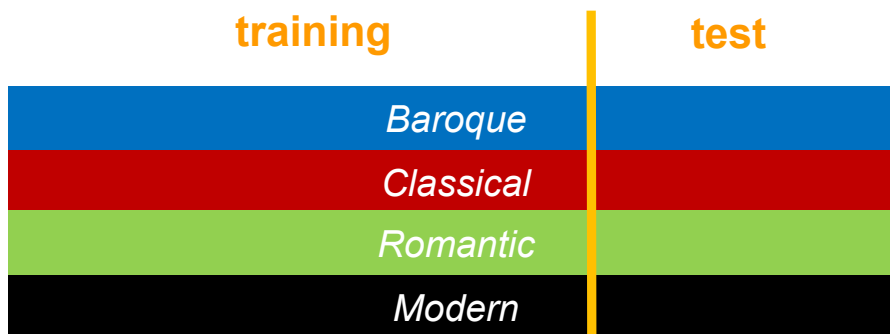
**Overfitting???**

Weiss / Mauch / Dixon, *Timbre-Invariant Audio Features for Style Analysis of Classical Music*, ICMC / SMC 2014

# Classification Results

- GMM classifier, LDA reduction, 3-fold cross validation

	Full Dataset	Piano	Orchestra
<i>Standard features</i>	87 %	88 %	85 %
<i>Tonal features</i>	84 %	84 %	86 %
<b><i>Combined</i></b>	<b>92 %</b>	<b>86 %</b>	<b>80 %</b>



Flexer, A Closer Look on Artist Filters for Musical Genre Classification, ISMIR 2007



# Classification Results

- GMM classifier, LDA reduction, 3-fold cross validation
- **No composer filter**

	<b>Full Dataset</b>	<b>Piano</b>	<b>Orchestra</b>
<i>Standard features</i>	87 %	88 %	85 %
<i>Tonal features</i>	84 %	84 %	86 %
<b>Combined</b>	<b>92 %</b>	<b>86 %</b>	<b>80 %</b>

- **Using composer filter**

	<b>Full Dataset</b>	<b>Piano</b>	<b>Orchestra</b>
<i>Standard features</i>	54 %	36 %	70 %
<i>Tonal features</i>	73 %	70 %	78 %
<b>Combined</b>	<b>68 %</b>	<b>44 %</b>	<b>68 %</b>

Weiss / Müller, *Tonal Complexity Features for Style Classification of Classical Music*, ICASSP 2015

# Classification Results

- GMM classifier, LDA reduction, 3-fold cross validation
- **No composer filter**

	Full Dataset	Piano	Orchestra
<i>Standard features</i>	87 %	88 %	85 %
<i>Tonal features</i>	84 %	84 %	86 %
<b>Combined</b>	<b>92 %</b>	<b>86 %</b>	<b>80 %</b>

- **Using composer filter**

	Full Dataset	Piano	Orchestra
<i>Standard features</i>	54 %	36 %	70 %
<i>Tonal features</i>	73 %	70 %	78 %
<b>Combined</b>	<b>68 %</b>	<b>44 %</b>	<b>68 %</b>

Weiss / Müller, *Tonal Complexity Features for Style Classification of Classical Music*, ICASSP 2015

# Classification Results

- What is actually learned?
- Pay attention to:
  - Overfitting
  - „Curse of dimensionality“ – use dimensionality reduction techniques
  - Artist / album effects
- Evaluation: „Figures of merit“:
  - Confusion matrix
  - Error examples: Consistently misclassified items
  - Listening tests
- Evaluation on unseen data (no cross validation)

Bob Sturm, *Classification Accuracy is not enough*,  
Journal of Intelligent Information Systems, 2013

# Classification Results: Confusion Matrix

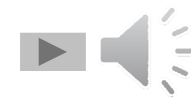
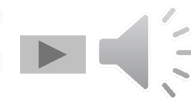
- 80 tonal features, GMM with 1 Gaussian, LDA, composer filtering
- **Full dataset**
- Mean accuracy: **75 %**
- Inter-class standard deviation: **6.7 %**

Era (correct)	Baroque	65.2	23.2	10.9	0.6
	Classical	17.0	74.9	8.1	0.0
	Romantic	6.5	5.0	77.7	10.8
	Modern	1.7	0.9	16.8	80.6
		Baroque	Classical	Romantic	Modern
		Era (classified)			

# Classification Results: Error Examples

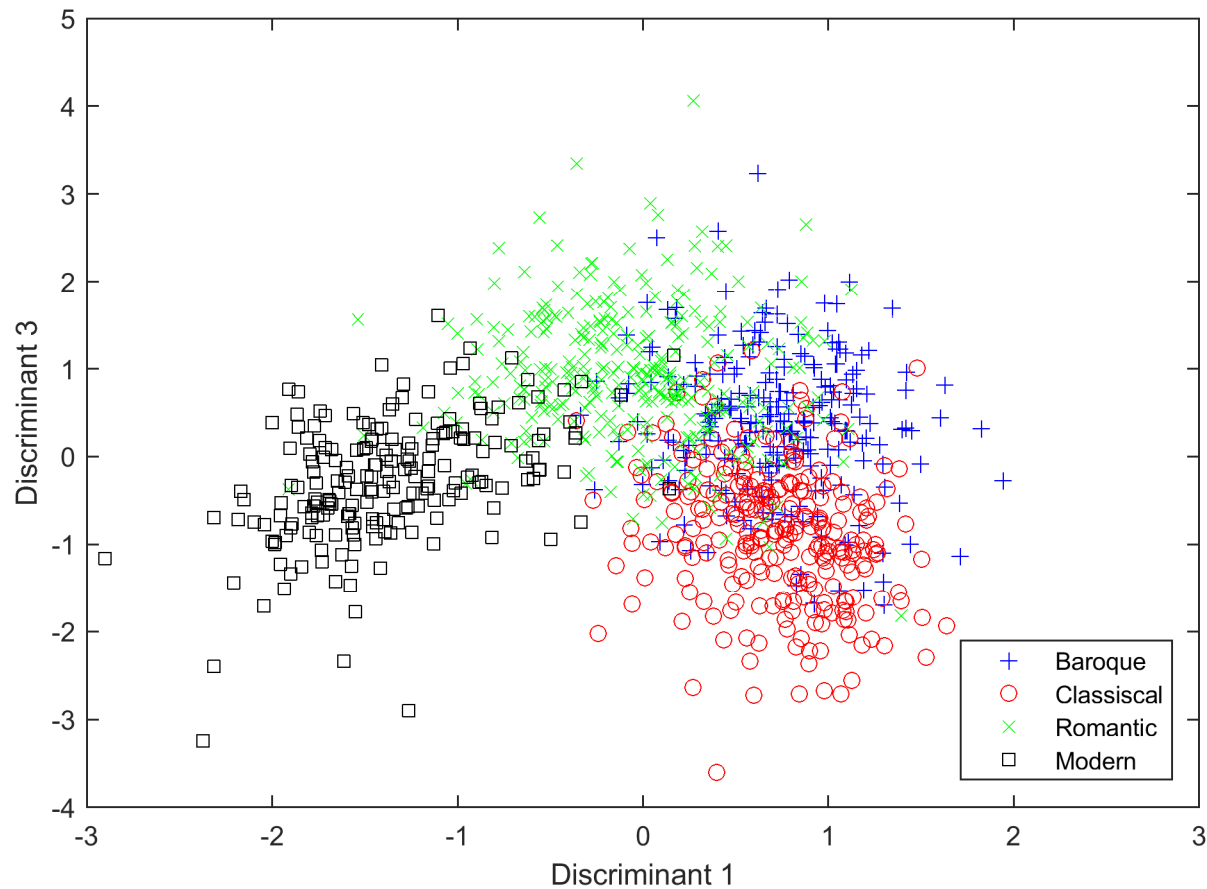
- 80 tonal features, GMM with 1 Gaussian, LDA
- Look at **consistently** and **persistently** misclassified items

<i>Class</i>	<i>Composer</i>	<i>Piece</i>	<i>Classified</i>
Baroque	Bach, J. S.	Well-Tempered Piano 1, Prelude in E $\flat$ minor BWV 853	Romantic
Baroque	Bach, J. S.	Well-Tempered Piano 1, Prelude in F major BWV 856	Romantic
Baroque	Bach, J. S.	Well-Tempered Piano 1, Prelude in A minor BWV 865	Romantic
Baroque	Bach, J. S.	Well-Tempered Piano 1, Prelude in B $\flat$ major BWV 866	Romantic
Baroque	Bach, J. S.	Well-Tempered Piano 1, Prelude in B $\flat$ minor BWV 867	Romantic
Baroque	Bach, J. S.	English Suite No. 3 in G minor BWV 808, Sarabande	Romantic
Baroque	Bach, J. S.	Brandenburg Conc. No. 1 in F major BWV 1046, Adagio	Romantic
Baroque	Bach, J. S.	Overture No. 2 in B minor BWV 1067, Badinerie	Romantic
Baroque	Bach, J. S.	Overture No. 3 in D major BWV 1068, Gigue	Romantic
Baroque	Couperin, F.	27 Ordres, Huitième ordre, IX. Rondeau passacaille	Romantic
Baroque	Corelli, A.	Concerto grosso op. 6 No. 2, III. Grave – Andante largo	Romantic
Baroque	Lully, J.-B.	Ballet de Xerces LWV 12, Gavotte en rondeau	Romantic
Baroque	Purcell, H.	Opera “Dido and Aeneas” Z. 626, Overture	Romantic
Baroque	Vivaldi, A.	“The Four Seasons,” RV 293 “Autumn,” Adagio molto	Romantic
Romantic	Schumann, R.	Kinderszenen op. 15, “Haschemann”	Baroque
Romantic	Grieg, E.	Holberg suite op. 40, Gavotte	Baroque
Romantic	Mendelssohn, F.	Symphony No. 4 in A major, IV. Saltarello, presto	Baroque
Modern	Shostakovich, D.	Preludes & Fugues op. 87 Fugue No. 1 in C major	Baroque
Modern	Shostakovich, D.	Preludes & Fugues op. 87 Fugue No. 5 in D major	Baroque



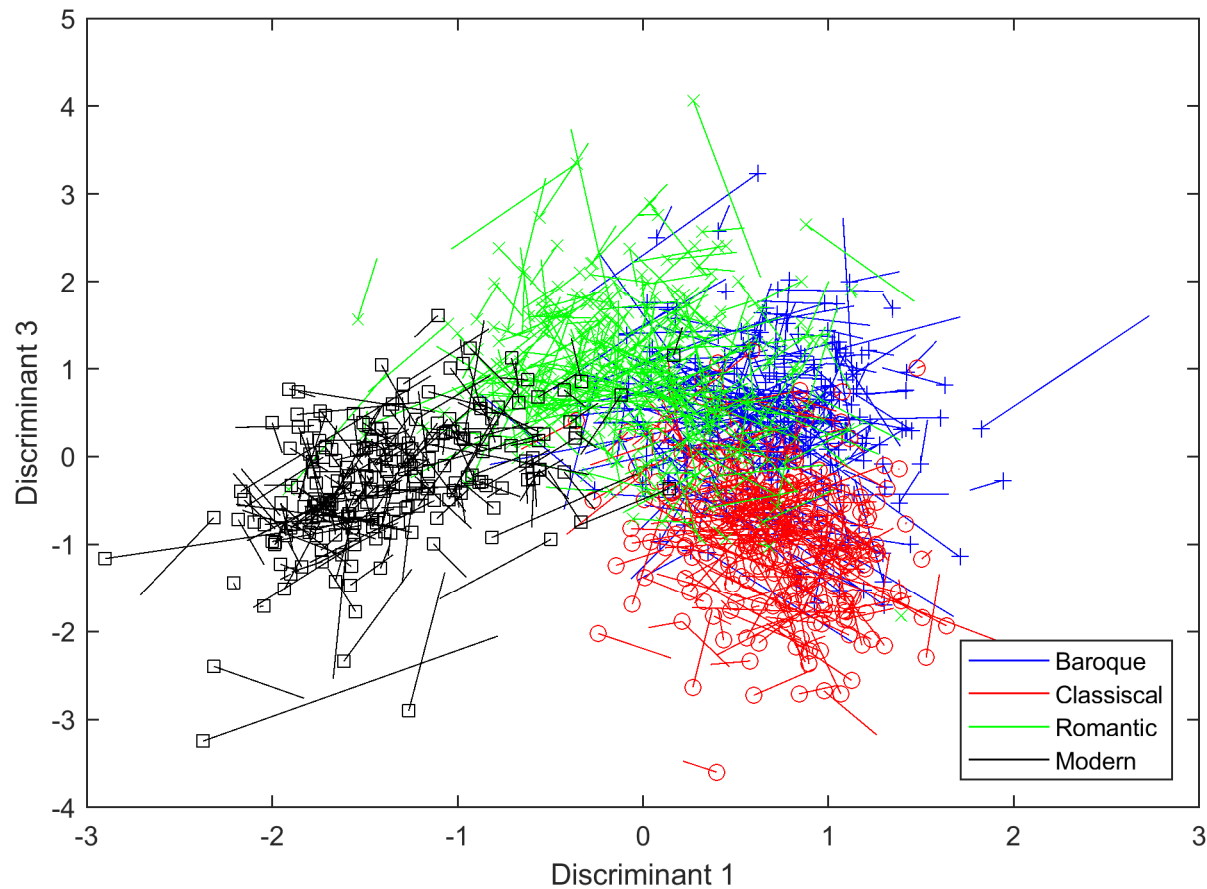
# Classification Results: Cross-Version Analysis

- Dimensionality reduction: *Cross-Era* dataset



# Classification Results: Cross-Version Analysis

- Dimensionality reduction: *Cross-Era – Cross-Era Mirror* dataset



# Conclusions

