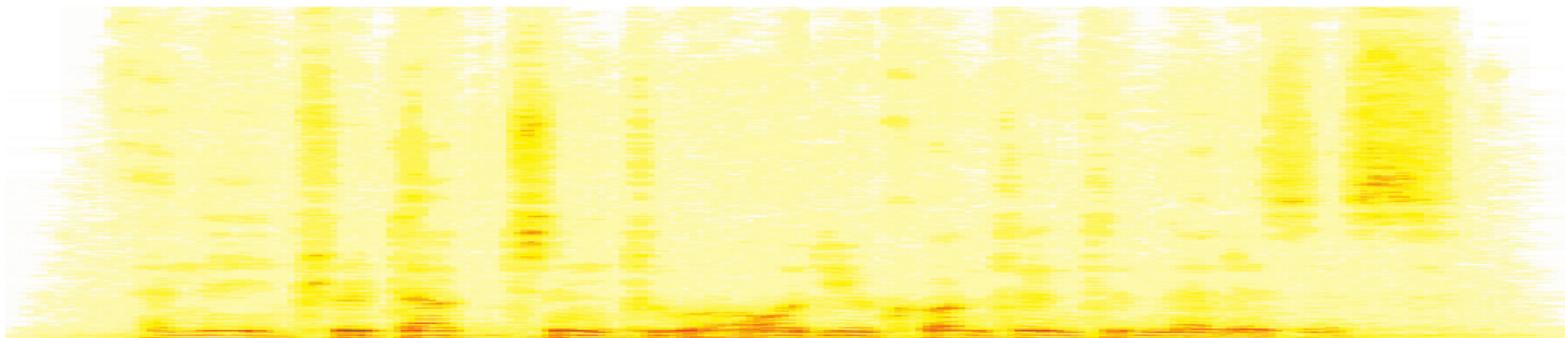


# Introduction to Audio Content Analysis

## Module 5.6: Tuning Frequency Estimation

alexander lerch



# introduction

## overview

corresponding textbook section

[Chapter 5 — Tonal Analysis](#): pp. 94–96

[Chapter 5 — Tonal Analysis](#): pp. 114–116

### ● lecture content

- definition of tuning frequency
- approaches to tuning frequency estimation

### ● learning objectives

- explain the term tuning frequency
- discuss when automatic tuning frequency estimation might be necessary
- compare different general approaches to tuning frequency estimation



# introduction

## overview

corresponding textbook section

[Chapter 5 — Tonal Analysis](#): pp. 94–96

[Chapter 5 — Tonal Analysis](#): pp. 114–116

### ● lecture content

- definition of tuning frequency
- approaches to tuning frequency estimation

### ● learning objectives

- explain the term tuning frequency
- discuss when automatic tuning frequency estimation might be necessary
- compare different general approaches to tuning frequency estimation



# tuning frequency

## introduction 1/2

### tuning frequency

frequency of the concert pitch *A4*

- used to tune groups of instruments
- standardized:<sup>1</sup> 440 Hz

---

<sup>1</sup>I. 16:1975, "Acoustics Standard tuning frequency (Standard musical pitch)," ISO, Standard, 1975.

# tuning frequency

## introduction 1/2

### tuning frequency

frequency of the concert pitch *A4*

- used to tune groups of instruments
- standardized:<sup>1</sup> 440 Hz

---

<sup>1</sup>I. 16:1975, "Acoustics Standard tuning frequency (Standard musical pitch)," ISO, Standard, 1975.

# tuning frequency

introduction 1/2

## tuning frequency

frequency of the concert pitch A4

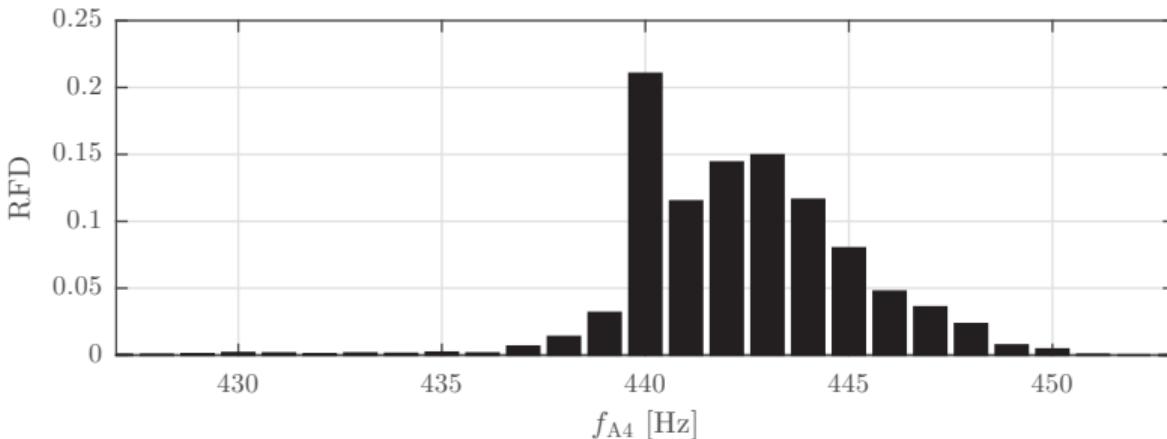
- used to tune groups of instruments
- standardized:<sup>1</sup> 440 Hz
- tuning frequency range *in history*

Year	Lower Deviation	Upper Deviation
1750	-50 Hz	+30 Hz
1850	-20 Hz	+20 Hz
1950	-5 Hz	+10 Hz

<sup>1</sup>I. 16:1975, "Acoustics Standard tuning frequency (Standard musical pitch)," ISO, Standard, 1975.

# tuning frequency

## introduction 2/2



- tuning frequencies *today*

- electronic music: often exactly 440 Hz
- “classical” music:<sup>2</sup>
  - Chicago Symphony, New York Philharmonic: 442 Hz
  - Berlin Philharmonic, Vienna Philharmonic: 443 Hz

<sup>2</sup>A. Lerch, “On the Requirement of Automatic Tuning Frequency Estimation,” in *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR)*, Victoria, 2006.

# tuning frequency estimation

## problem statement

- any pitch-based analysis system relies on tuning frequency (pre-defined or adaptive)

$$\text{recall MIDI pitch: } p(f) = 69 + 12 \cdot \log_2 \left( \frac{f}{f_{A4}} \right)$$

- tuning frequency might deviate significantly from 440 Hz
- key & temperament are unknown
- expressive intonation is unknown

⇒ a wrong tuning frequency assumption may impact the pitch reliability

# tuning frequency estimation

## problem statement

- any pitch-based analysis system relies on tuning frequency (pre-defined or adaptive)

$$\text{recall MIDI pitch: } p(f) = 69 + 12 \cdot \log_2 \left( \frac{f}{f_{A4}} \right)$$

- tuning frequency might deviate significantly from 440 Hz
  - key & temperament are unknown
  - expressive intonation is unknown
- ⇒ a wrong tuning frequency assumption may impact the pitch reliability

# tuning frequency estimation

## assumptions and limits

- ① **key is unknown** ⇒ deviations larger than half a semitone will be mapped back to semitone range
- ② **temperament/intonation is unknown** ⇒ assuming equidistant spaced pitch (equally tempered)
- ③ piece **may be polyphonic**
- ④ piece **may not contain the pitch A4**
- ⑤ if tuning frequency changes over time, it will **change slowly**

# tuning frequency estimation

overview

## • typical processing steps

- ① estimate fundamental frequencies
- ② calculate deviation  $\Delta C$  from the nearest (equally tempered) pitch frequency in Cents
- ③ average all deviations:  $\rightarrow \bar{\Delta C}$
- ④ compute the tuning frequency from the average distance:

$$\hat{f}_{A4} = 440 \text{ Hz} \cdot 2^{\frac{\bar{\Delta C}}{1200}}$$

# tuning frequency estimation

## score-based

- ① analyze score and **identify important pitches**
- ② define a set of narrow **band pass filters** per pitch
- ③ sweep **band pass filter's mid frequency** over selected audio segments
- ④ identify most likely **mid frequency** maximizing the filter's output energy
- ⑤ compute its **deviation  $\Delta C$**  from the (equally tempered) pitch frequency in Cent
- ⑥ **average all results**
- ⑦ compute overall result in Hertz

# tuning frequency estimation

## score-based

- ① analyze score and **identify important pitches**
- ② define a set of narrow **band pass filters** per pitch
- ③ sweep band pass filter's mid frequency over selected audio segments
- ④ identify most likely **mid frequency** maximizing the filter's output energy
- ⑤ compute its **deviation  $\Delta C$**  from the (equally tempered) pitch frequency in Cent
- ⑥ average all results
- ⑦ compute overall result in Hertz

# tuning frequency estimation

## score-based

- ① analyze score and **identify important pitches**
- ② define a set of narrow **band pass filters** per pitch
- ③ **sweep band pass** filter's mid frequency over selected audio segments
- ④ identify most likely **mid frequency** maximizing the filter's output energy
- ⑤ compute its **deviation  $\Delta C$**  from the (equally tempered) pitch frequency in Cent
- ⑥ **average all results**
- ⑦ compute overall result in Hertz

# tuning frequency estimation

## score-based

- ① analyze score and **identify important pitches**
- ② define a set of narrow **band pass filters** per pitch
- ③ **sweep band pass** filter's mid frequency over selected audio segments
- ④ identify most likely **mid frequency** maximizing the filter's output energy
- ⑤ compute its **deviation  $\Delta C$**  from the (equally tempered) pitch frequency in Cent
- ⑥ average all results
- ⑦ compute overall result in Hertz

# tuning frequency estimation

## score-based

- ① analyze score and **identify important pitches**
- ② define a set of narrow **band pass filters** per pitch
- ③ **sweep band pass** filter's mid frequency over selected audio segments
- ④ identify most likely **mid frequency** maximizing the filter's output energy
- ⑤ compute its **deviation**  $\Delta C$  from the (equally tempered) pitch frequency in Cent
- ⑥ **average all results**
- ⑦ compute overall result in Hertz

# tuning frequency estimation

## score-based

- ① analyze score and **identify important pitches**
- ② define a set of narrow **band pass filters** per pitch
- ③ **sweep band pass** filter's mid frequency over selected audio segments
- ④ identify most likely **mid frequency** maximizing the filter's output energy
- ⑤ compute its **deviation**  $\Delta C$  from the (equally tempered) pitch frequency in Cent
- ⑥ **average** all results
- ⑦ compute overall result in Hertz

# tuning frequency estimation

## score-based

- ① analyze score and **identify important pitches**
- ② define a set of narrow **band pass filters** per pitch
- ③ **sweep band pass** filter's mid frequency over selected audio segments
- ④ identify most likely **mid frequency** maximizing the filter's output energy
- ⑤ compute its **deviation**  $\Delta C$  from the (equally tempered) pitch frequency in Cent
- ⑥ **average** all results
- ⑦ compute overall result in Hertz

# tuning frequency estimation

## STFT-based

- ① identify **spectral peaks**
- ② compute **instantaneous frequencies**  $f_I(k)$
- ③ compute **distance** to nearest equally tempered semitone frequency  $\hat{f}_{A4}(n - 1)$  (in Cents)
- ④ iteratively **adjust tuning frequency** estimate to minimize distance

# tuning frequency estimation

## STFT-based

- ① identify **spectral peaks**
- ② compute **instantaneous frequencies**  $f_I(k)$
- ③ compute **distance** to nearest equally tempered semitone frequency  $\hat{f}_{A4}(n - 1)$  (in Cents)
- ④ iteratively **adjust tuning frequency** estimate to minimize distance

# tuning frequency estimation

## STFT-based

- ① identify **spectral peaks**
- ② compute **instantaneous frequencies**  $f_I(k)$
- ③ compute **distance** to nearest equally tempered semitone frequency  $\hat{f}_{A4}(n - 1)$  (in Cents)
- ④ iteratively **adjust tuning frequency** estimate to minimize distance

# tuning frequency estimation

## STFT-based

- ① identify **spectral peaks**
- ② compute **instantaneous frequencies**  $f_I(k)$
- ③ compute **distance** to nearest equally tempered semitone frequency  $\hat{f}_{A4}(n - 1)$  (in Cents)
- ④ iteratively **adjust tuning frequency** estimate to minimize distance

# tuning frequency estimation

## histogram-based

- ① compute **CQT**: bin resolution 10 Cents
- ② compute **distance** of peaks to nearest equally tempered semitone frequency @  $f_{A4} = 440 \text{ Hz}$  (in Cents)
- ③ compute **histogram of deviations**
- ④ choose **position of histogram maximum** as tuning frequency deviation from  $f_{A4}$

# tuning frequency estimation

## histogram-based

- ① compute **CQT**: bin resolution 10 Cents
- ② compute **distance** of peaks to nearest equally tempered semitone frequency @  $f_{A4} = 440 \text{ Hz}$  (in Cents)
- ③ compute **histogram of deviations**
- ④ choose **position of histogram maximum** as tuning frequency deviation from  $f_{A4}$

# tuning frequency estimation

## histogram-based

- ① compute **CQT**: bin resolution 10 Cents
- ② compute **distance** of peaks to nearest equally tempered semitone frequency @  $f_{A4} = 440 \text{ Hz}$  (in Cents)
- ③ compute **histogram of deviations**
- ④ choose position of histogram maximum as tuning frequency deviation from  $f_{A4}$

# tuning frequency estimation

## histogram-based

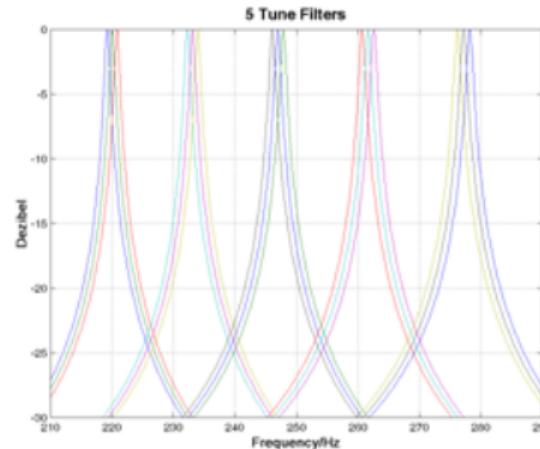
- ① compute **CQT**: bin resolution 10 Cents
- ② compute **distance** of peaks to nearest equally tempered semitone frequency @  $f_{A4} = 440 \text{ Hz}$  (in Cents)
- ③ compute **histogram of deviations**
- ④ choose **position of histogram maximum** as tuning frequency deviation from  $f_{A4}$

# tuning frequency estimation

## filterbank-based 1/2

- 1 filterbank with 3 closely spaced filters per semi-tone
- 2 compute output energy for every triplet  $E_j(0, n), E_j(1, n), E_j(2, n)$
- 3 compute overall triplet output energy

$$E(\cdot, n) = \sum_{\forall j} E_j(\cdot, n)$$



plots from<sup>3, 4</sup>

<sup>3</sup> A. Lerch, "Ein Ansatz zur automatischen Erkennung der Tonart in Musikdateien," in *Proceedings of the VDT International Audio Convention (23. Tonmeistertagung)*, Leipzig, Nov. 2004.

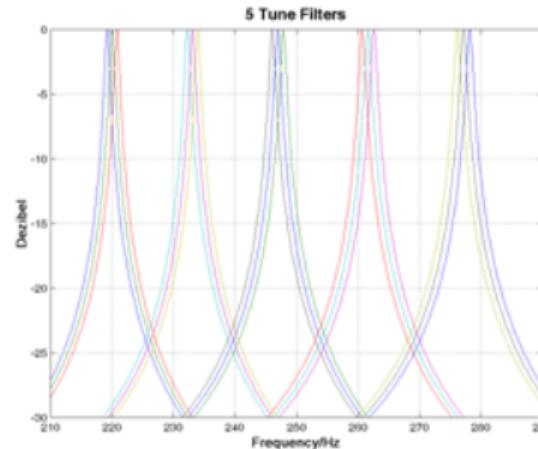
<sup>4</sup> A. Lerch, "On the Requirement of Automatic Tuning Frequency Estimation," in *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR)*, Victoria, 2006.

# tuning frequency estimation

## filterbank-based 1/2

- ① filterbank with 3 closely spaced filters per semi-tone
- ② compute output energy for every triplet  $E_j(0, n), E_j(1, n), E_j(2, n)$
- ③ compute overall triplet output energy

$$E(\cdot, n) = \sum_{\forall j} E_j(\cdot, n)$$



plots from<sup>3, 4</sup>

<sup>3</sup> A. Lerch, "Ein Ansatz zur automatischen Erkennung der Tonart in Musikdateien," in *Proceedings of the VDT International Audio Convention (23. Tonmeistertagung)*, Leipzig, Nov. 2004.

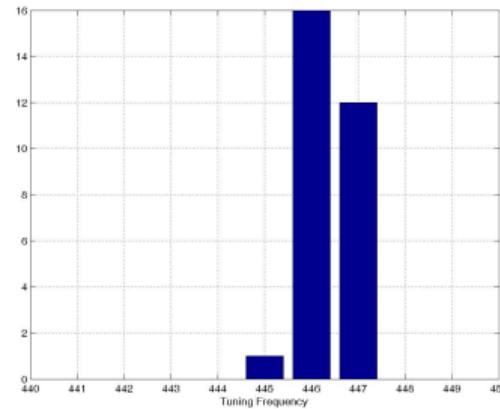
<sup>4</sup> A. Lerch, "On the Requirement of Automatic Tuning Frequency Estimation," in *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR)*, Victoria, 2006.

# tuning frequency estimation

## filterbank-based 1/2

- ① filterbank with 3 closely spaced filters per semi-tone
- ② compute output energy for every triplet  $E_j(0, n), E_j(1, n), E_j(2, n)$
- ③ compute overall triplet output energy

$$E(\cdot, n) = \sum_{\forall j} E_j(\cdot, n)$$



plots from<sup>3, 4</sup>

<sup>3</sup> A. Lerch, "Ein Ansatz zur automatischen Erkennung der Tonart in Musikdateien," in *Proceedings of the VDT International Audio Convention (23. Tonmeistertagung)*, Leipzig, Nov. 2004.

<sup>4</sup> A. Lerch, "On the Requirement of Automatic Tuning Frequency Estimation," in *Proceedings of the 7th International Conference on Music Information Retrieval (ISMIR)*, Victoria, 2006.

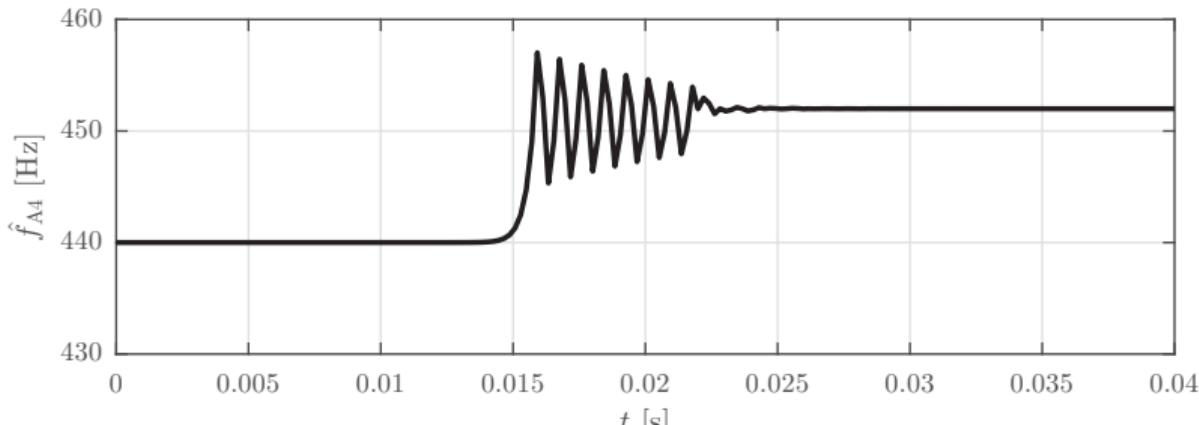
# tuning frequency estimation

## filterbank-based 2/2

### ④ adapt current tuning frequency estimate

$$\begin{aligned}\hat{f}_{A4}(n+1) &= (1 + \text{sign}(E(2) - E(0)) \cdot \lambda(n)) \cdot \hat{f}_{A4}(n) \\ \lambda(n+1) &= c \cdot \lambda(n)\end{aligned}$$

- $c > 1$  if constant direction/sign
- $c < 1$  if changed direction



# tuning frequency estimation

## filterbank-based 2/2

### ④ adapt current tuning frequency estimate

$$\begin{aligned}\hat{f}_{A4}(n+1) &= (1 + \text{sign}(E(2) - E(0)) \cdot \lambda(n)) \cdot \hat{f}_{A4}(n) \\ \lambda(n+1) &= c \cdot \lambda(n)\end{aligned}$$

- $c > 1$  if constant direction/sign
- $c < 1$  if changed direction



## ● tuning frequency

- important reference for all pitch-based algorithms
- usually around 440 Hz

## ● tuning frequency estimation

- can be online or offline, score-based or “blind”
- typical processing steps
  - 1 estimate fundamental frequencies
  - 2 calculate deviation from the nearest (equally tempered) pitch frequency
  - 3 average deviations of all fundamental frequencies
  - 4 compute the tuning frequency from the average distance

