



## Tutorial

# Automatisierte Methoden der Musikverarbeitung 47. Jahrestagung der Gesellschaft für Informatik

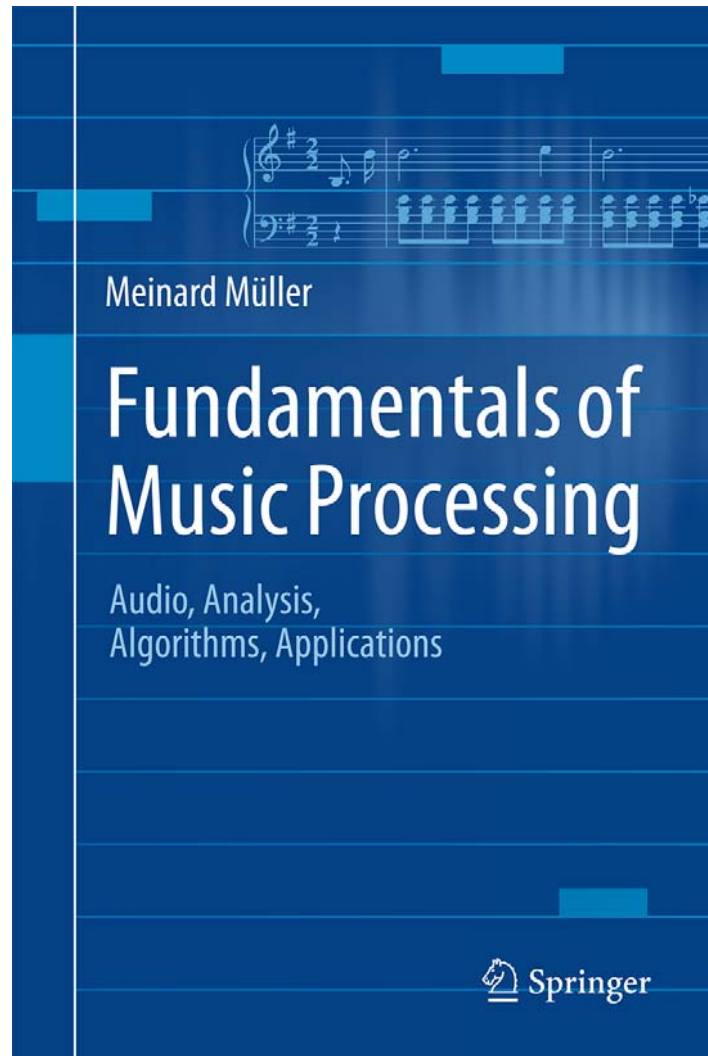
## Tempo and Beat Tracking

**Meinard Müller, Christof Weiss, Stefan Balke**

International Audio Laboratories Erlangen

{meinard.mueller, christof.weiss, stefan.balke}@audiolabs-erlangen.de

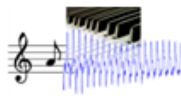

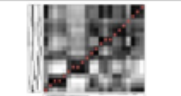


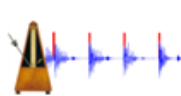
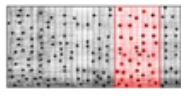
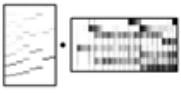
# Book: Fundamentals of Music Processing



Meinard Müller  
Fundamentals of Music Processing  
Audio, Analysis, Algorithms, Applications  
483 p., 249 illus., hardcover  
ISBN: 978-3-319-21944-8  
Springer, 2015

Accompanying website:  
[www.music-processing.de](http://www.music-processing.de)

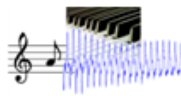

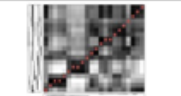
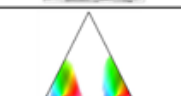

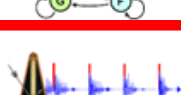


# Book: Fundamentals of Music Processing

Chapter		Music Processing Scenario
1		Music Representations
2		Fourier Analysis of Signals
3		Music Synchronization
4		Music Structure Analysis
5		Chord Recognition
6		Tempo and Beat Tracking
7		Content-Based Audio Retrieval
8		Musically Informed Audio Decomposition

Meinard Müller  
Fundamentals of Music Processing  
Audio, Analysis, Algorithms, Applications  
483 p., 249 illus., hardcover  
ISBN: 978-3-319-21944-8  
Springer, 2015

Accompanying website:  
[www.music-processing.de](http://www.music-processing.de)

# Book: Fundamentals of Music Processing

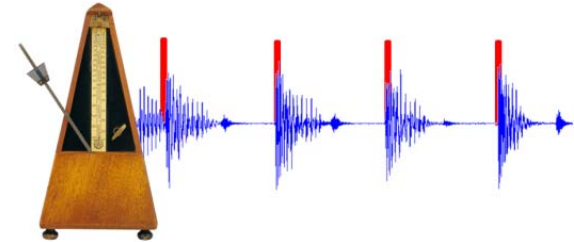
Chapter		Music Processing Scenario
1		Music Representations
2		Fourier Analysis of Signals
3		Music Synchronization
4		Music Structure Analysis
5		Chord Recognition
6		Tempo and Beat Tracking
7		Content-Based Audio Retrieval
8		Musically Informed Audio Decomposition

Meinard Müller  
Fundamentals of Music Processing  
Audio, Analysis, Algorithms, Applications  
483 p., 249 illus., hardcover  
ISBN: 978-3-319-21944-8  
Springer, 2015

Accompanying website:  
[www.music-processing.de](http://www.music-processing.de)

# Chapter 6: Tempo and Beat Tracking

- 6.1 Onset Detection
- 6.2 Tempo Analysis
- 6.3 Beat and Pulse Tracking
- 6.4 Further Notes



Tempo and beat are further fundamental properties of music. In Chapter 6, we introduce the basic ideas on how to extract tempo-related information from audio recordings. In this scenario, a first challenge is to locate note onset information—a task that requires methods for detecting changes in energy and spectral content. To derive tempo and beat information, note onset candidates are then analyzed with regard to quasiperiodic patterns. This leads us to the study of general methods for local periodicity analysis of time series.

---

# Introduction

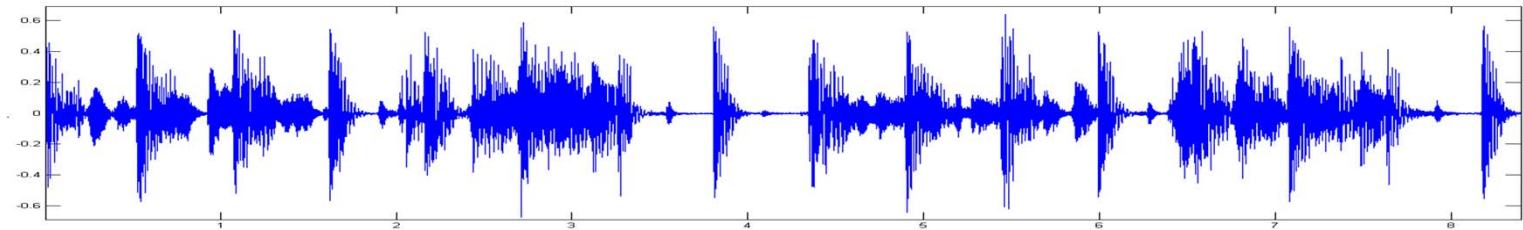
## **Basic beat tracking task:**

Given an audio recording of a piece of music,  
determine the periodic sequence of beat positions.

“Tapping the foot when listening to music”

# Introduction

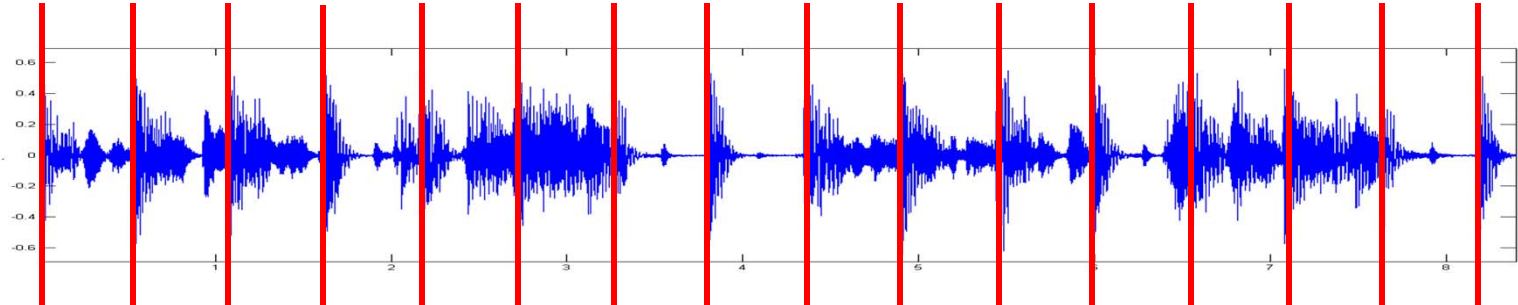
Example: Queen – Another One Bites The Dust



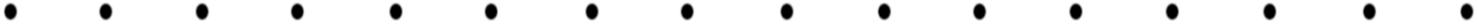
Time (seconds)

# Introduction

Example: Queen – Another One Bites The Dust



Time (seconds)





# Introduction

Example: Happy Birthday to you

Pulse level: **Measure**

The image shows two staves of musical notation for the song 'Happy Birthday to you'. The first staff contains the first two phrases: 'Hap - py Birth - day to you,' and 'Hap - py Birth - day to you, Hap - py'. The second staff contains the third phrase: 'Birth - day dear \_\_\_\_\_, Hap - py Birth - day to you!'. The music is in 3/4 time with a key signature of one sharp (F#). Four red arrows point downwards to the first note of each of the four measures in the first staff, indicating the pulse level at the measure level.

# Introduction

Example: Happy Birthday to you

Pulse level: **Tactus (beat)**

The image shows a musical score for the song "Happy Birthday to you" in 3/4 time. The first staff is marked with red arrows pointing to the downbeats of each measure, indicating the pulse level (Tactus). The lyrics are: "Hap - py Birth - day to you, Hap - py Birth - day to you, Hap - py Birth - day dear \_\_\_\_\_, Hap - py Birth - day to you!".

Hap - py Birth - day to you, Hap - py Birth - day to you, Hap - py Birth - day dear \_\_\_\_\_, Hap - py Birth - day to you!

# Introduction

Example: Happy Birthday to you

Pulse level: **Tatum (temporal atom)**

The image shows a musical score for the song "Happy Birthday to you" in 3/4 time. The score is written on two staves. The first staff contains the melody for the first two phrases: "Hap - py Birth - day to you, Hap - py Birth - day to you, Hap - py". The second staff contains the melody for the final phrase: "Birth - day dear \_\_\_\_\_, Hap - py Birth - day to you!". Above the first staff, there are 24 red arrows pointing downwards, indicating the pulse level (Tatum) for each note. The arrows are placed above the notes on the first staff, with one arrow above each of the 24 notes. The notes on the first staff are: G4 (quarter), A4 (quarter), B4 (quarter), C5 (quarter), B4 (quarter), A4 (quarter), G4 (quarter), F#4 (quarter), E4 (quarter), D4 (quarter), C4 (quarter), B3 (quarter), A3 (quarter), G3 (quarter), F#3 (quarter), E3 (quarter), D3 (quarter), C3 (quarter), B2 (quarter), A2 (quarter), G2 (quarter), F#2 (quarter), E2 (quarter), D2 (quarter), C2 (quarter). The notes on the second staff are: G4 (quarter), A4 (quarter), B4 (quarter), C5 (quarter), B4 (quarter), A4 (quarter), G4 (quarter), F#4 (quarter), E4 (quarter), D4 (quarter), C4 (quarter), B3 (quarter), A3 (quarter), G3 (quarter), F#3 (quarter), E3 (quarter), D3 (quarter), C3 (quarter), B2 (quarter), A2 (quarter), G2 (quarter), F#2 (quarter), E2 (quarter), D2 (quarter), C2 (quarter).

---

# Introduction

Example: Chopin – Mazurka Op. 68-3

Pulse level: Quarter note

Tempo: ???



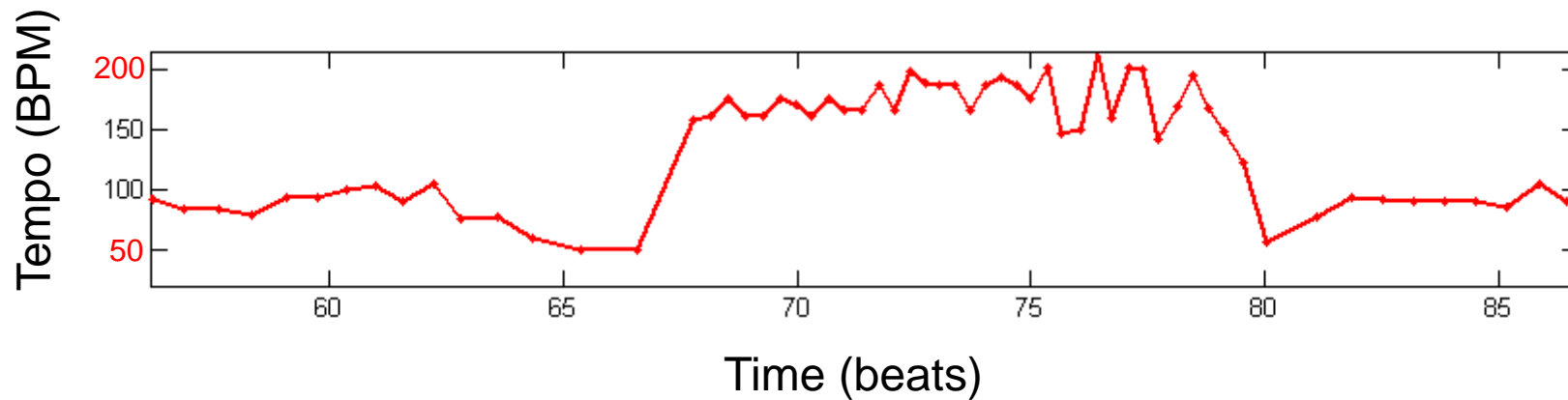
# Introduction

Example: Chopin – Mazurka Op. 68-3

Pulse level: Quarter note

Tempo: **50-200 BPM** 

Tempo curve



# Introduction

Example: Borodin – String Quartet No. 2

Pulse level: Quarter note

Tempo: 120-140 BPM (roughly)

Beat tracker without any prior knowledge



Beat tracker with prior knowledge on  
rough tempo range



# Introduction

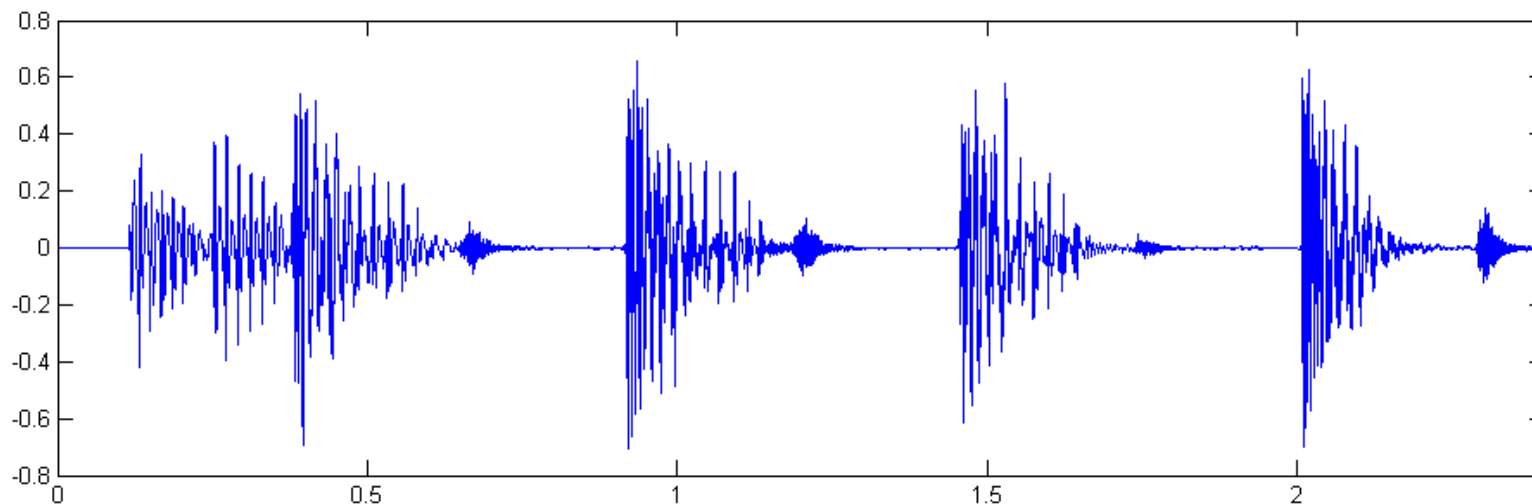
## Challenges in beat tracking

- Pulse level often unclear
- Local/sudden tempo changes (e.g. rubato)
- Vague information  
(e.g., soft onsets, extracted onsets corrupt)
- Sparse information  
(often only note onsets are used)

# Introduction

## Tasks

- Onset detection
- Beat tracking
- Tempo estimation

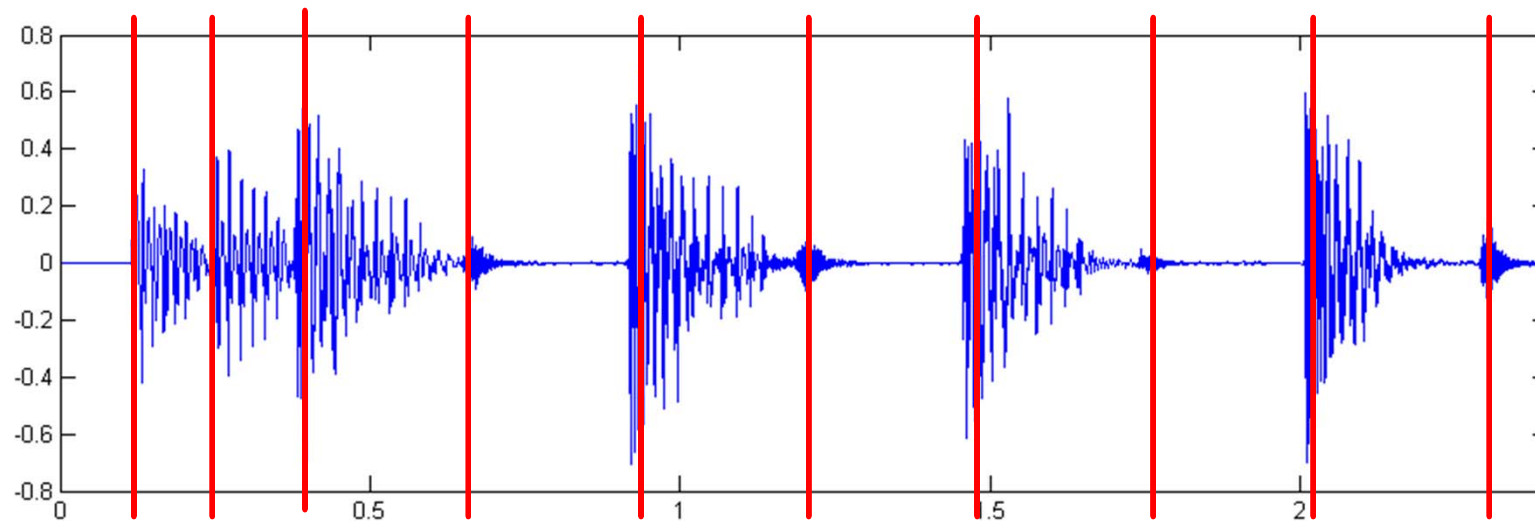




# Introduction

## Tasks

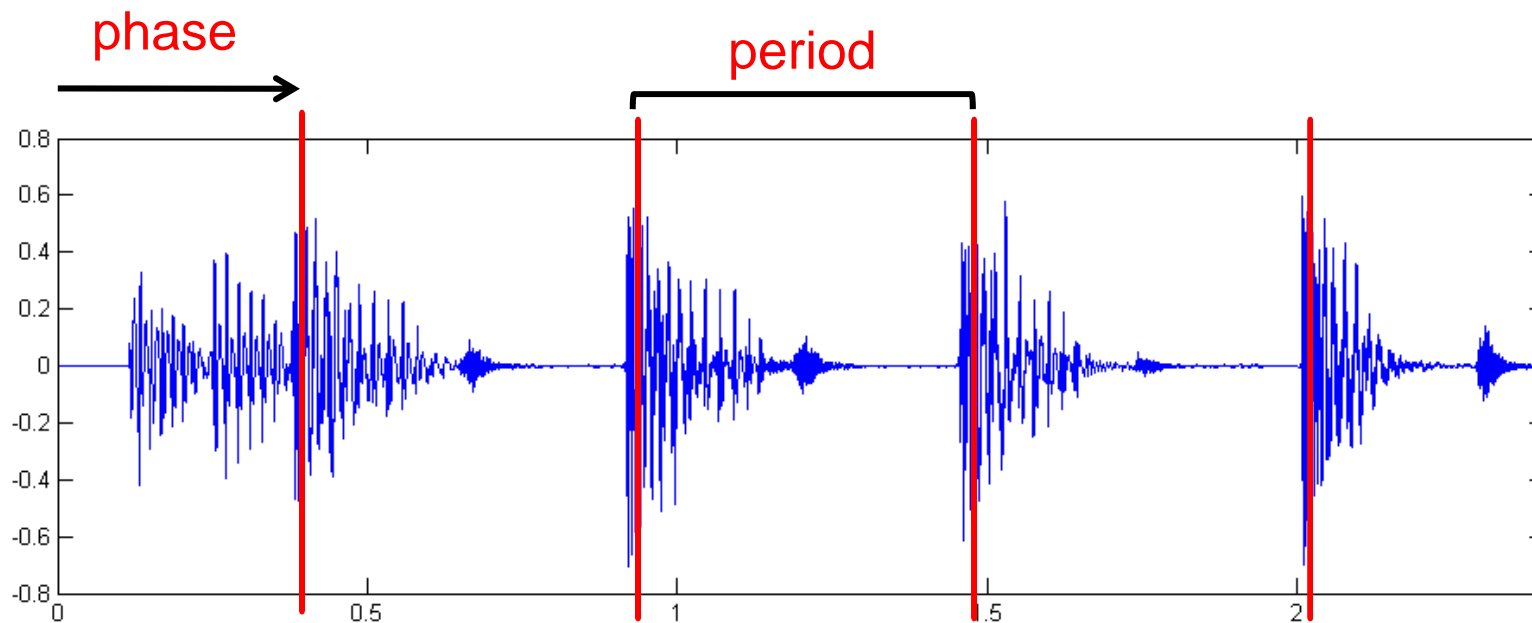
- Onset detection
- Beat tracking
- Tempo estimation



# Introduction

## Tasks

- Onset detection
- **Beat tracking**
- Tempo estimation



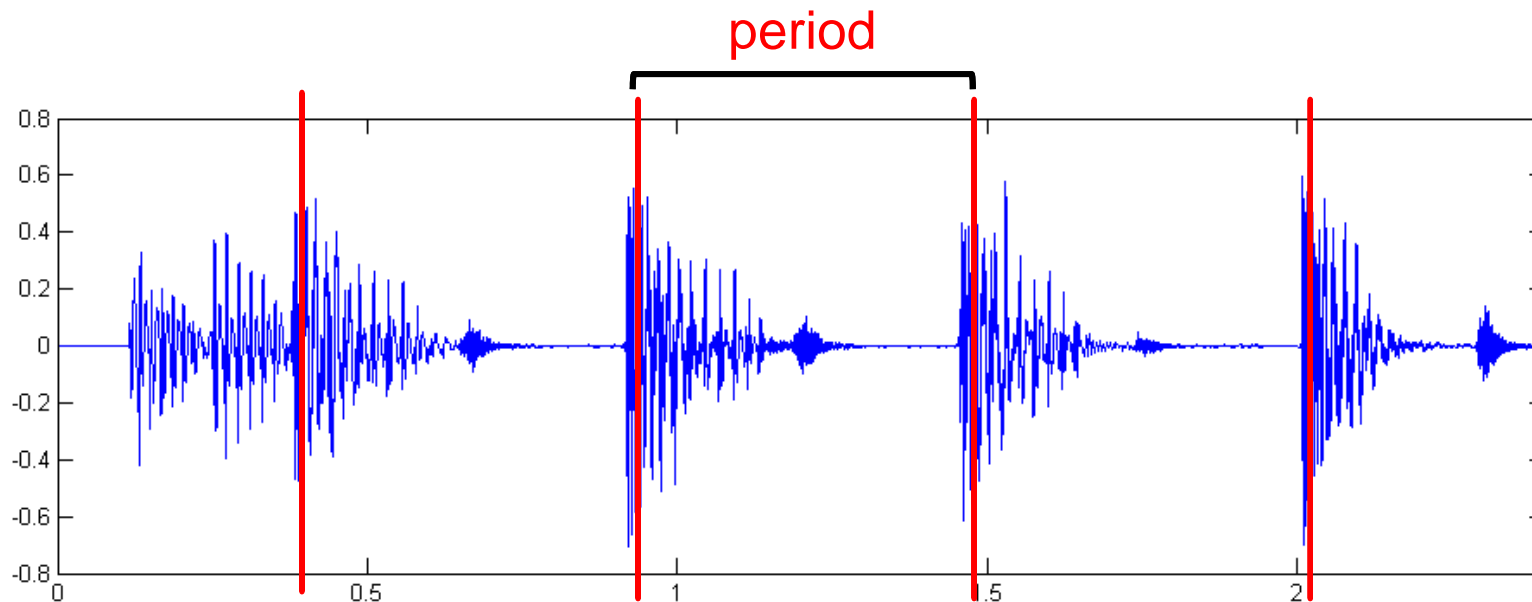
# Introduction

## Tasks

- Onset detection
- Beat tracking
- Tempo estimation

Tempo :=  $60 / \text{period}$

Beats per minute (BPM)

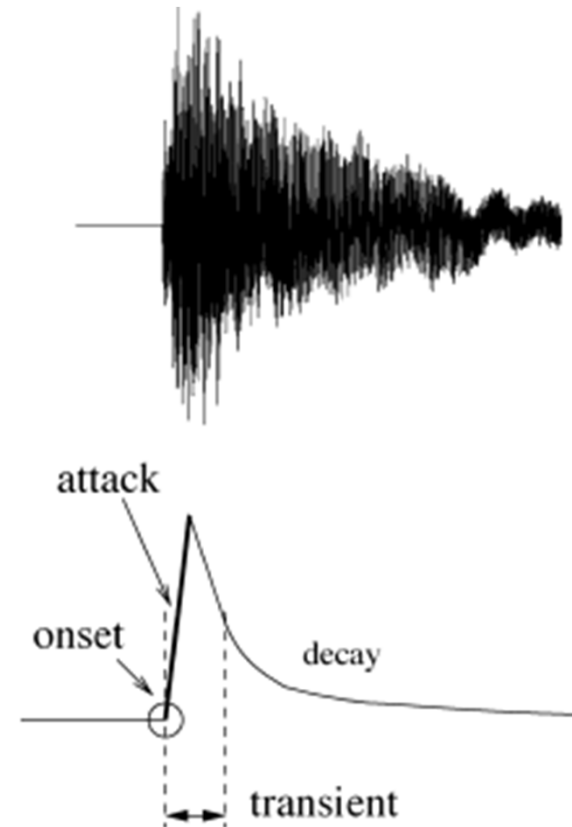


# Onset Detection

- Finding start times of perceptually relevant acoustic events in music signal
- Onset is the time position where a note is played
- Onset typically goes along with a change of the signal's properties:
  - energy or loudness
  - pitch or harmony
  - timbre

# Onset Detection

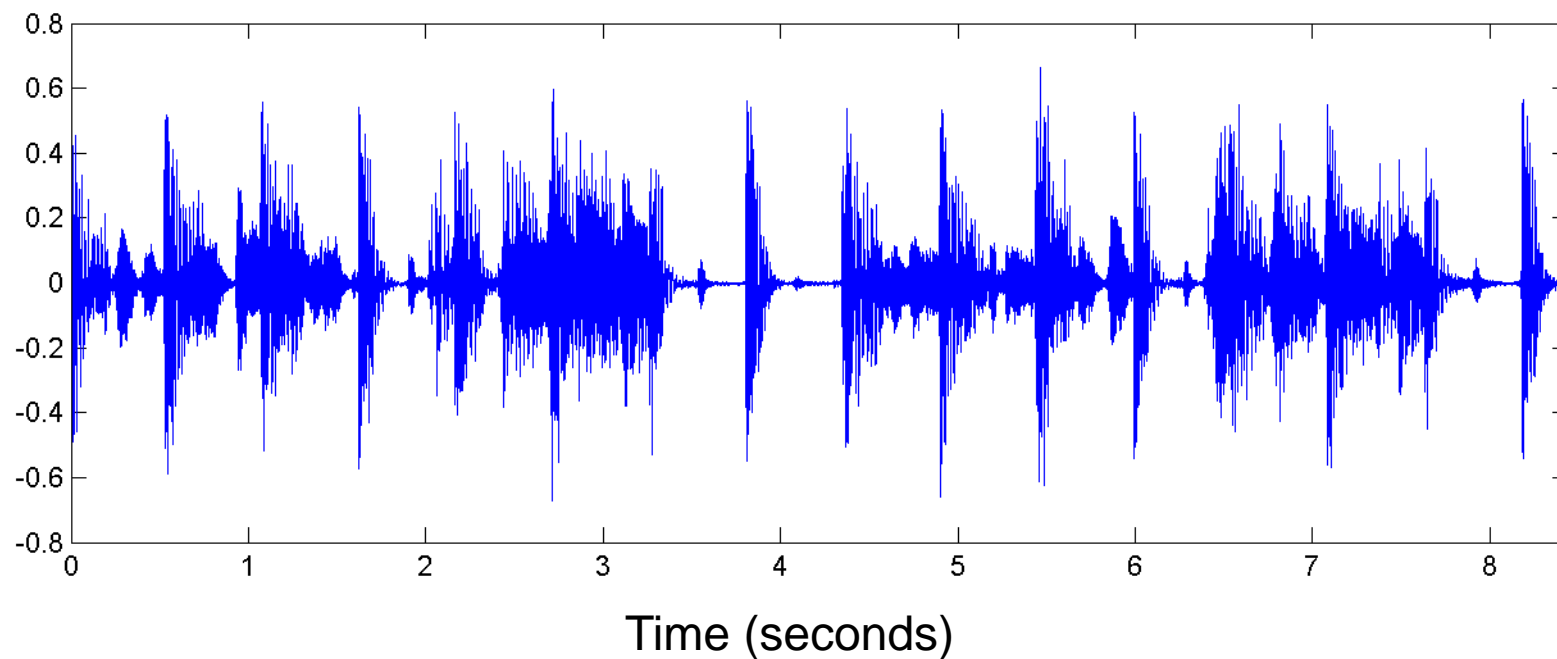
- Finding start times of perceptually relevant acoustic events in music signal
- Onset is the time position where a note is played
- Onset typically goes along with a change of the signal's properties:
  - energy or loudness
  - pitch or harmony
  - timbre



# Onset Detection (Energy-Based)

## Steps

### Waveform

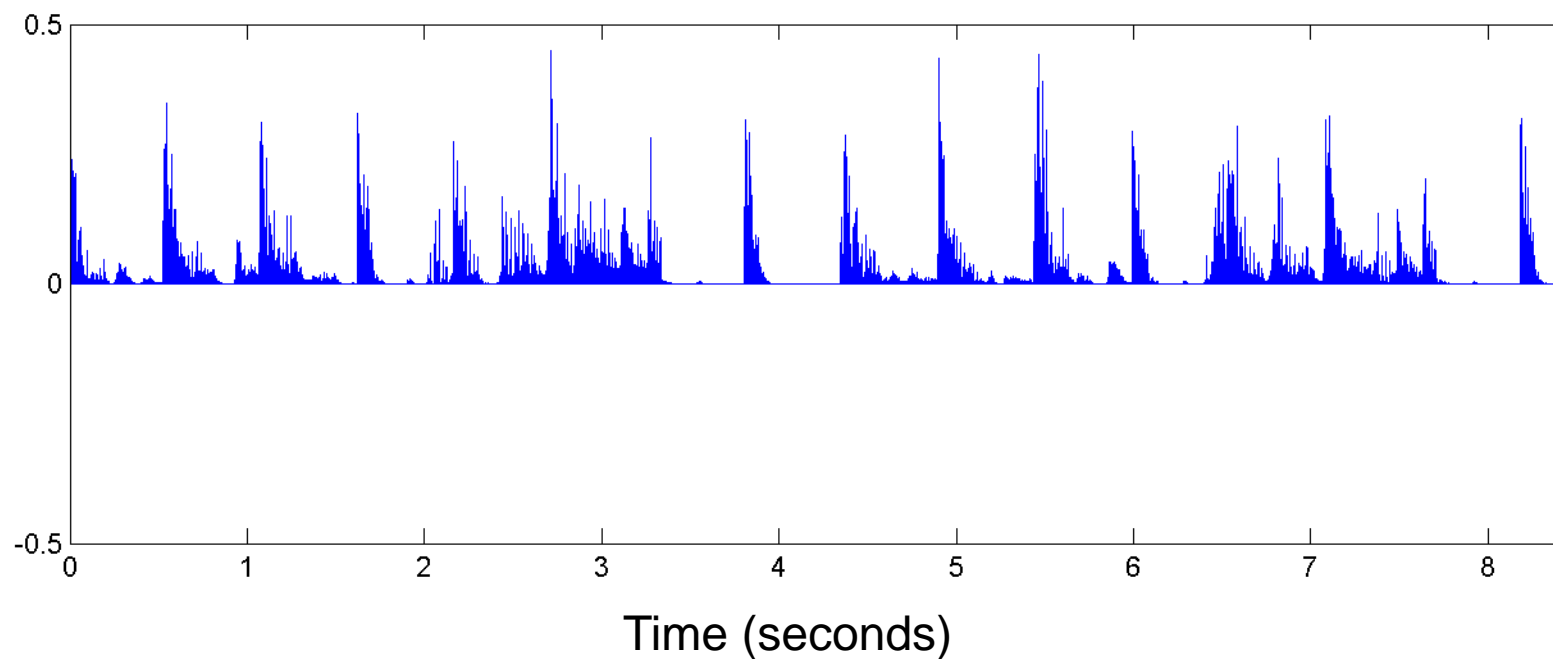


# Onset Detection (Energy-Based)

## Steps

1. Amplitude squaring

Squared waveform

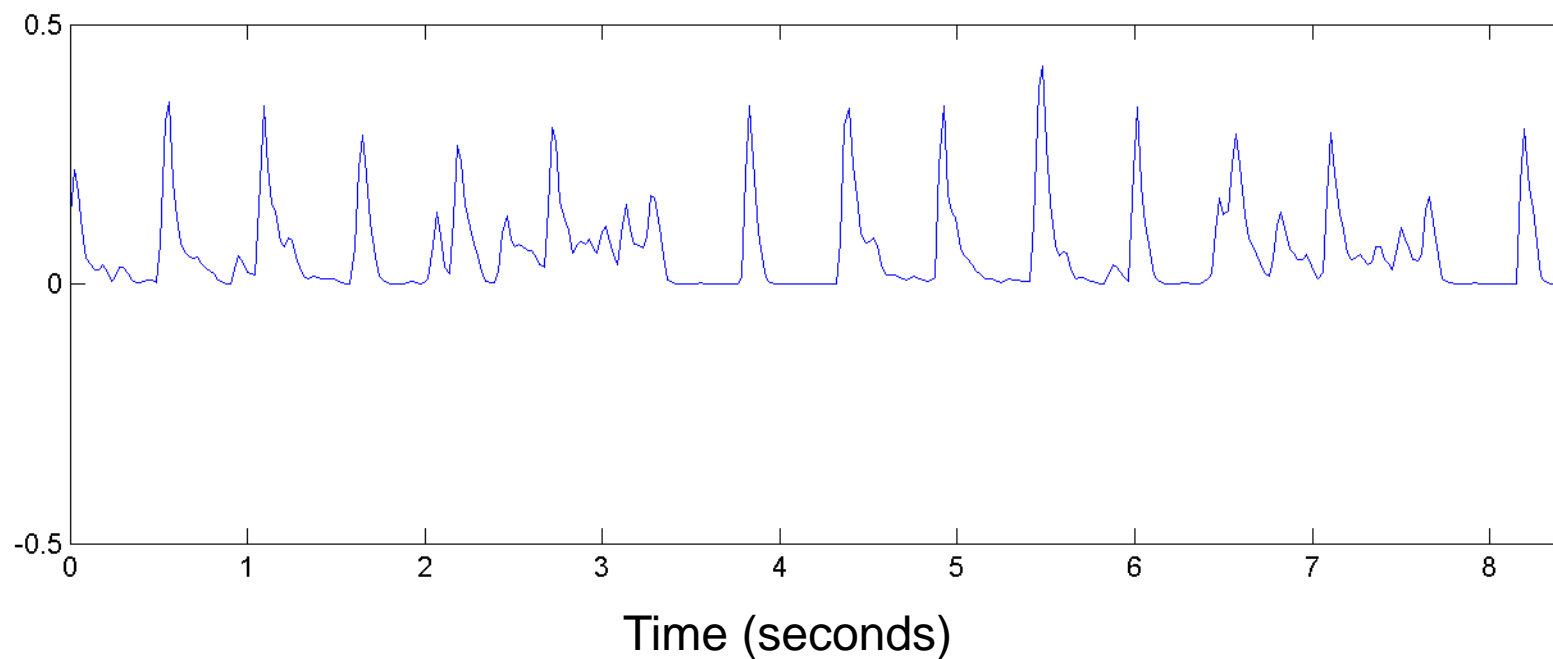


# Onset Detection (Energy-Based)

## Steps

1. Amplitude squaring
2. Windowing

## Energy envelope





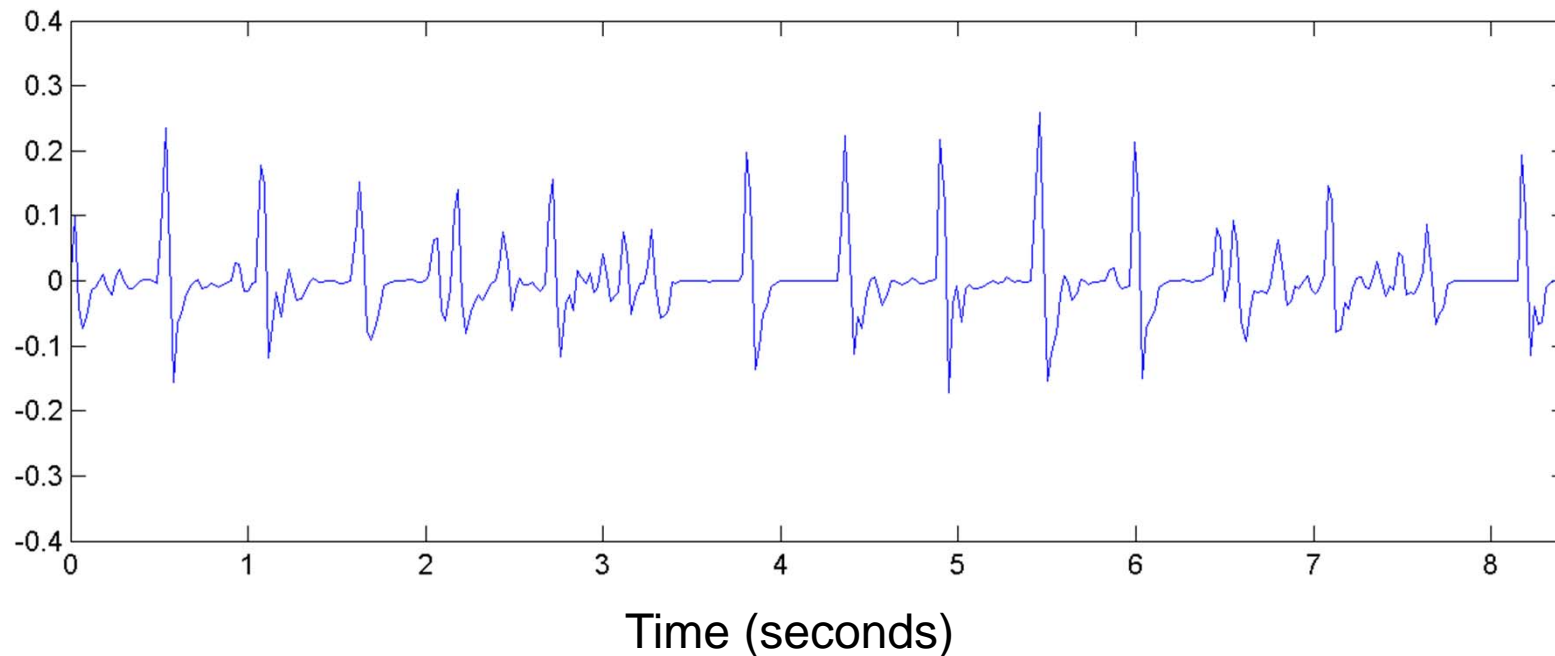
# Onset Detection (Energy-Based)

## Steps

1. Amplitude squaring
2. Windowing
3. Differentiation

Capturing energy changes

## Differentiated energy envelope



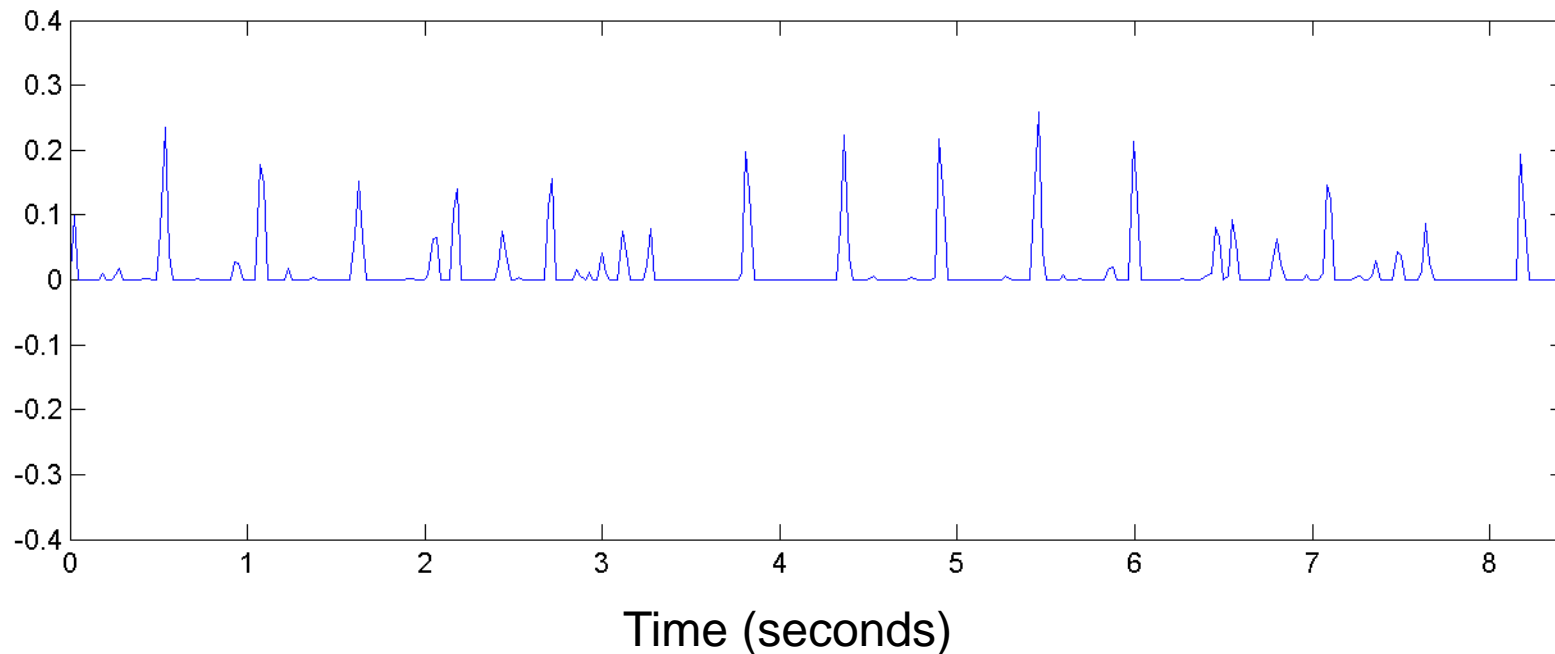
# Onset Detection (Energy-Based)

## Steps

1. Amplitude squaring
2. Windowing
3. Differentiation
4. Half wave rectification

Only energy increases are relevant for note onsets

## Novelty curve

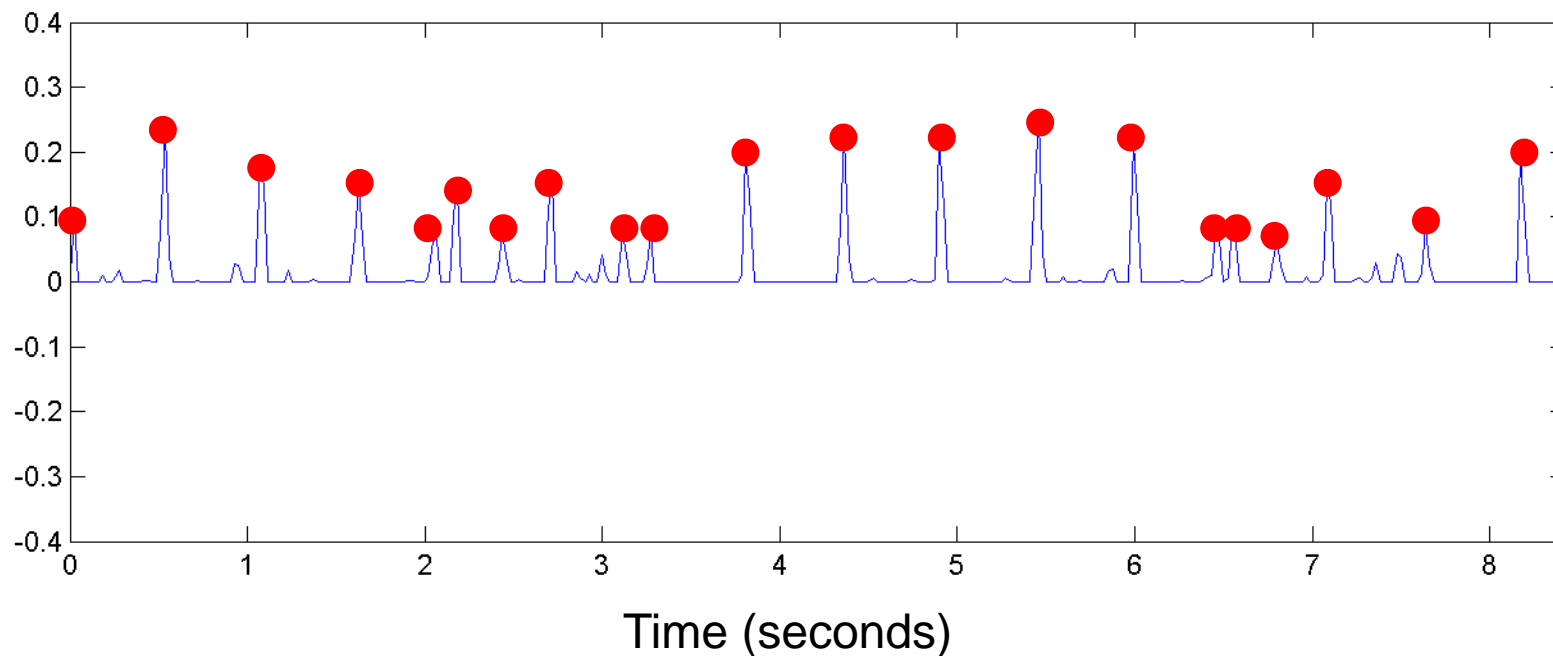


# Onset Detection (Energy-Based)

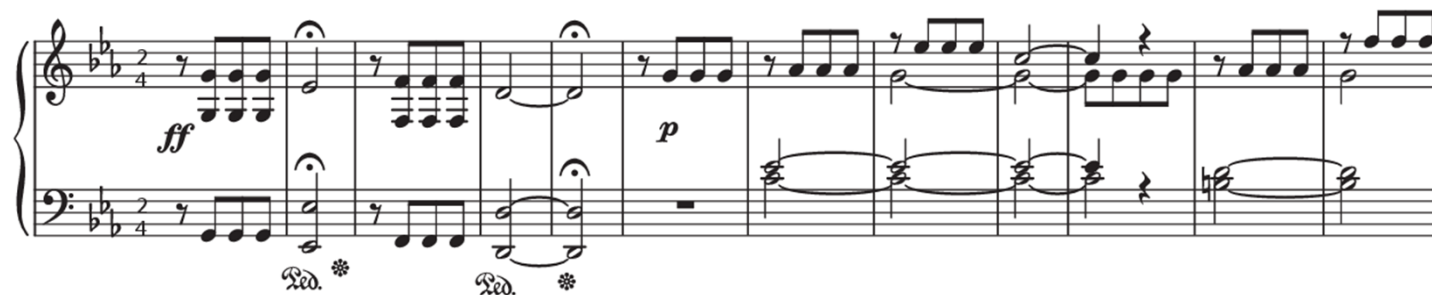
## Steps

1. Amplitude squaring
2. Windowing
3. Differentiation
4. Half wave rectification
5. Peak picking

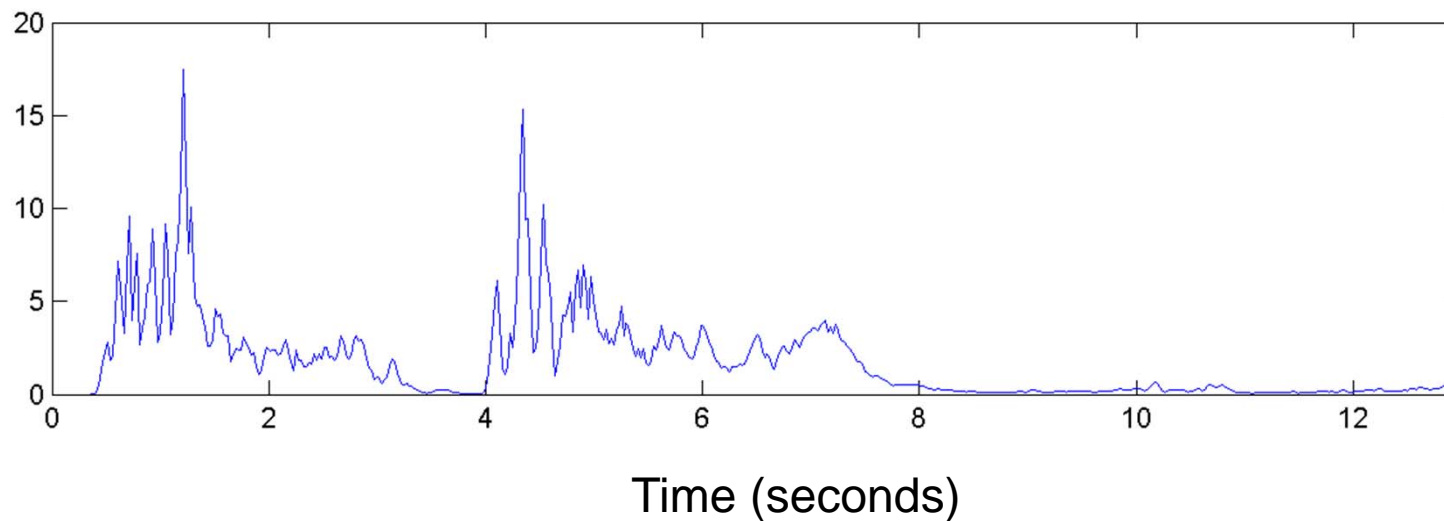
Peak positions indicate  
note onset candidates



# Onset Detection (Energy-Based)



Energy envelope

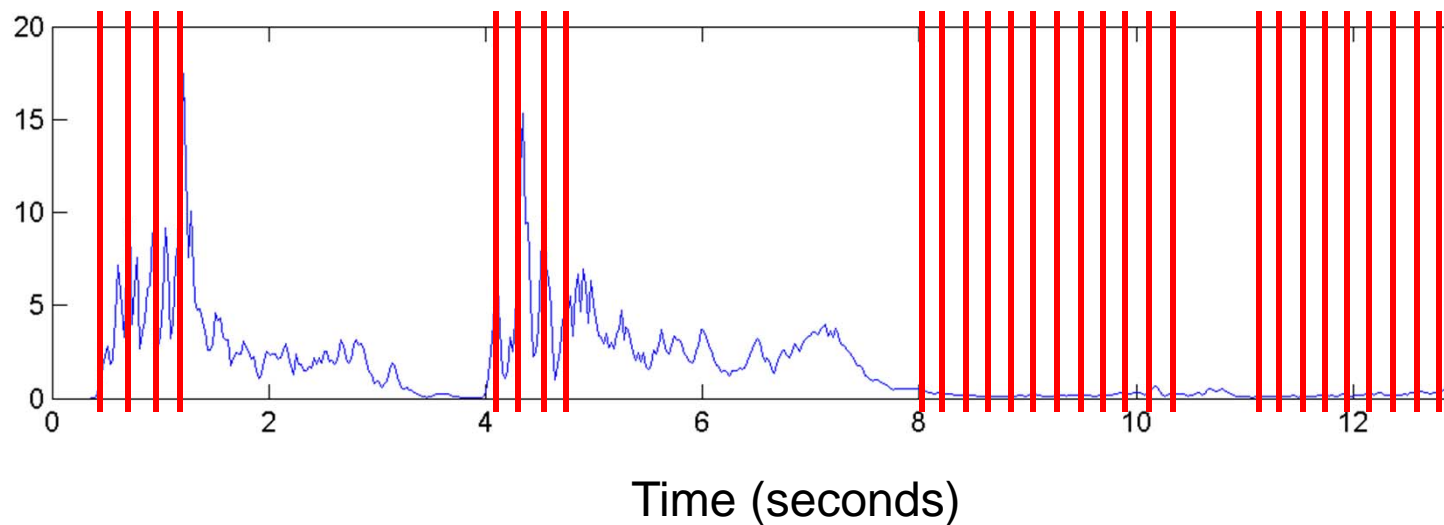


# Onset Detection (Energy-Based)

A musical score for piano in 2/4 time, featuring dynamics *ff* and *p*, and articulation marks such as accents and staccato. The score is presented in a standard musical notation format with a treble and bass clef.



Energy envelope / note onsets positions

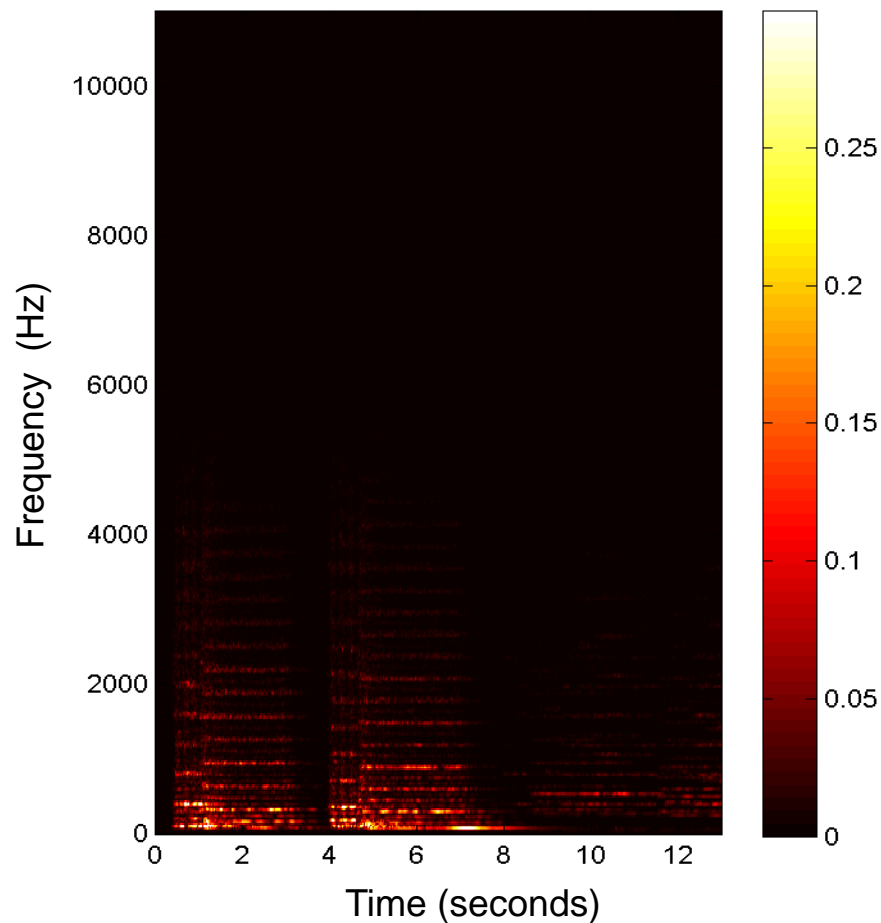


# Onset Detection

- Energy curves often only work for percussive music
- Many instruments such as strings have weak note onsets
- No energy increase may be observable in complex sound mixtures
- More refined methods needed that capture
  - changes of spectral content
  - changes of pitch
  - changes of harmony

# Onset Detection (Spectral-Based)

Magnitude spectrogram  $|X|$



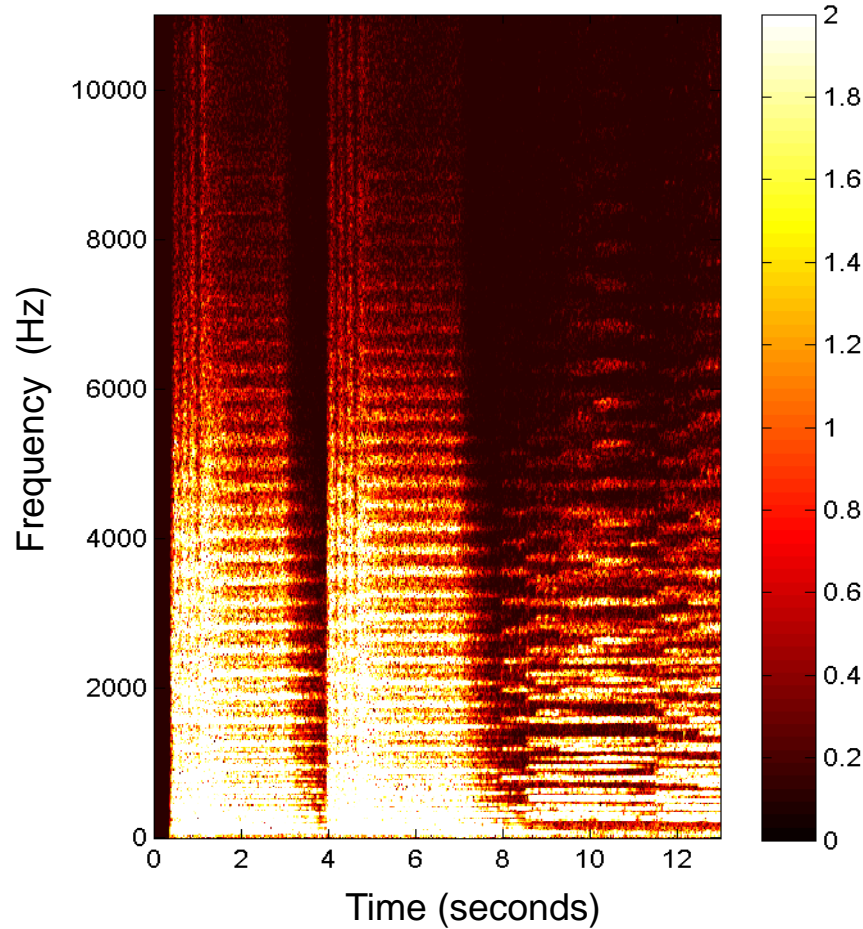
## Steps:

### 1. Spectrogram

- *Aspects concerning pitch, harmony, or timbre are captured by spectrogram*
- *Allows for detecting local energy changes in certain frequency ranges*

# Onset Detection (Spectral-Based)

Compressed spectrogram  $Y$



## Steps:

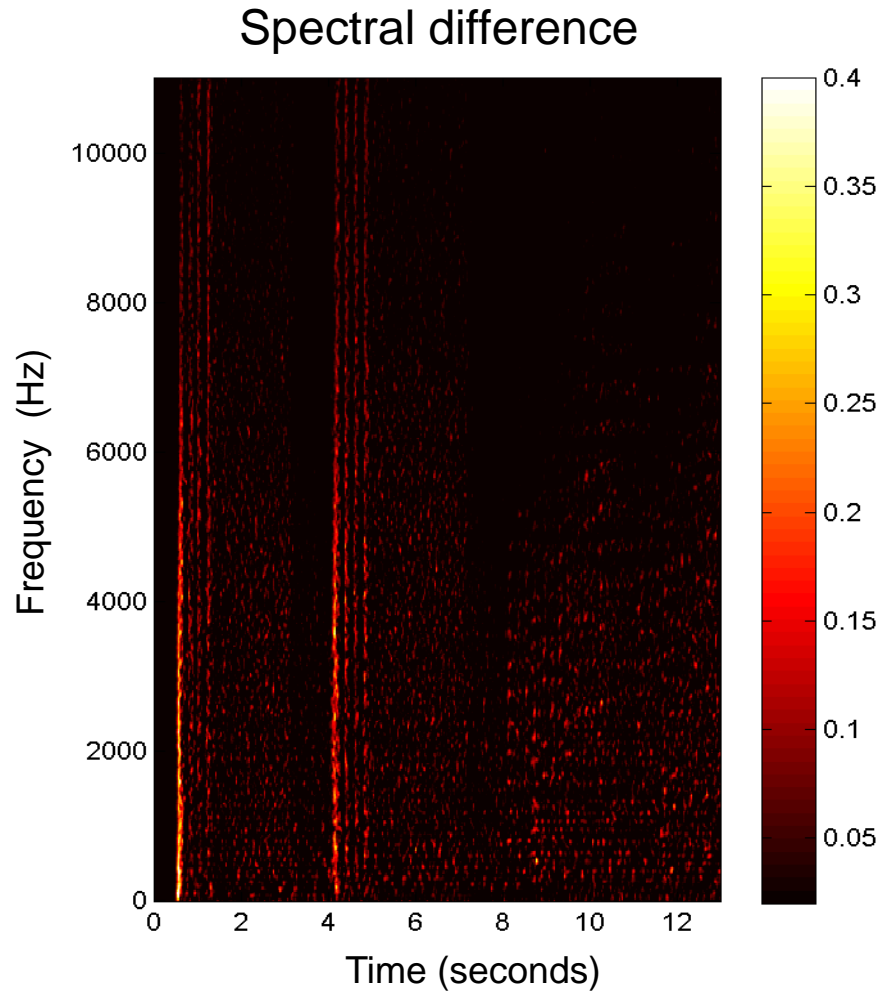
1. Spectrogram
2. Logarithmic compression

$$Y = \log(1 + C \cdot |X|)$$

- *Accounts for the logarithmic sensation of sound intensity*
- *Dynamic range compression*
- *Enhancement of low-intensity values*
- *Often leading to enhancement of high-frequency spectrum*



# Onset Detection (Spectral-Based)

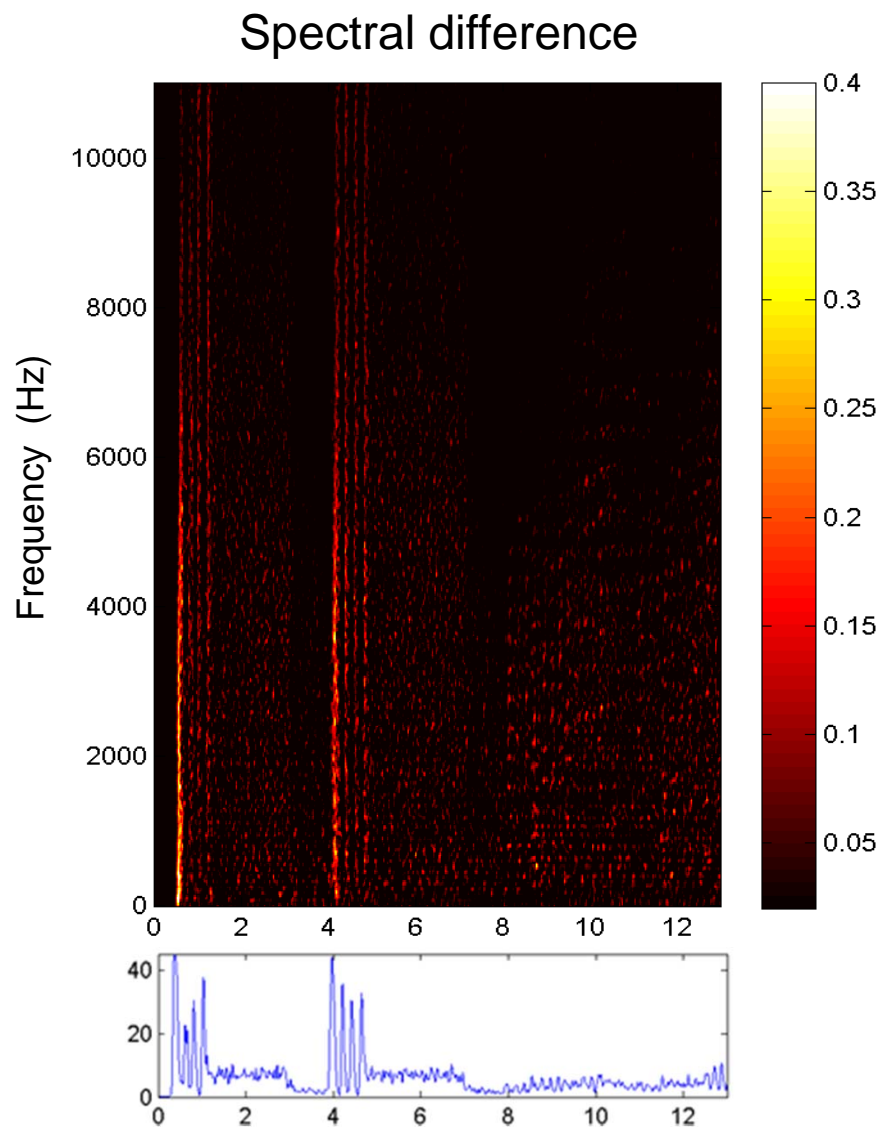


## Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation

- *First-order temporal difference*
- *Captures changes of the spectral content*
- *Only positive intensity changes considered*

# Onset Detection (Spectral-Based)



## Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation
4. Accumulation

- *Frame-wise accumulation of all positive intensity changes*
- *Encodes changes of the spectral content*

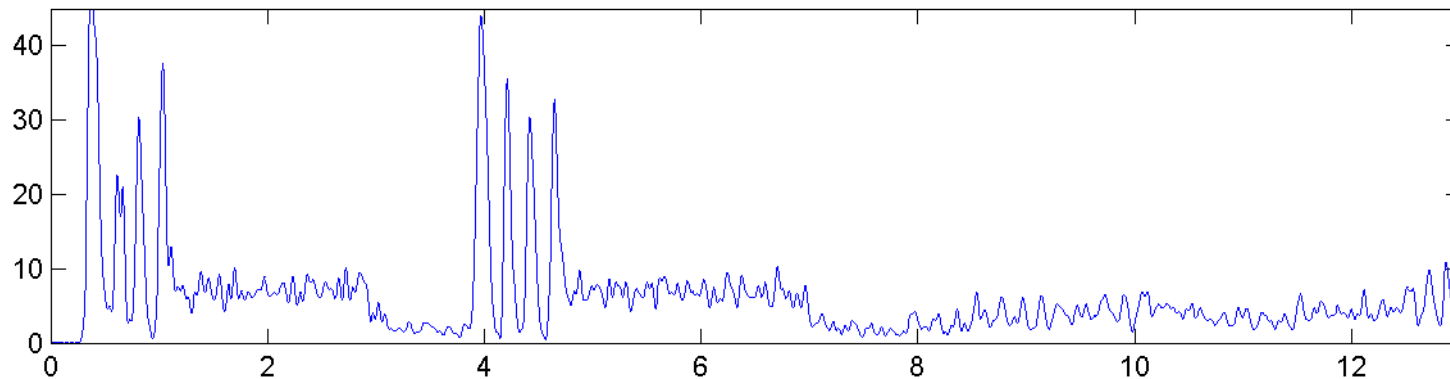
Novelty curve

# Onset Detection (Spectral-Based)

## Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation
4. Accumulation

Novelty curve



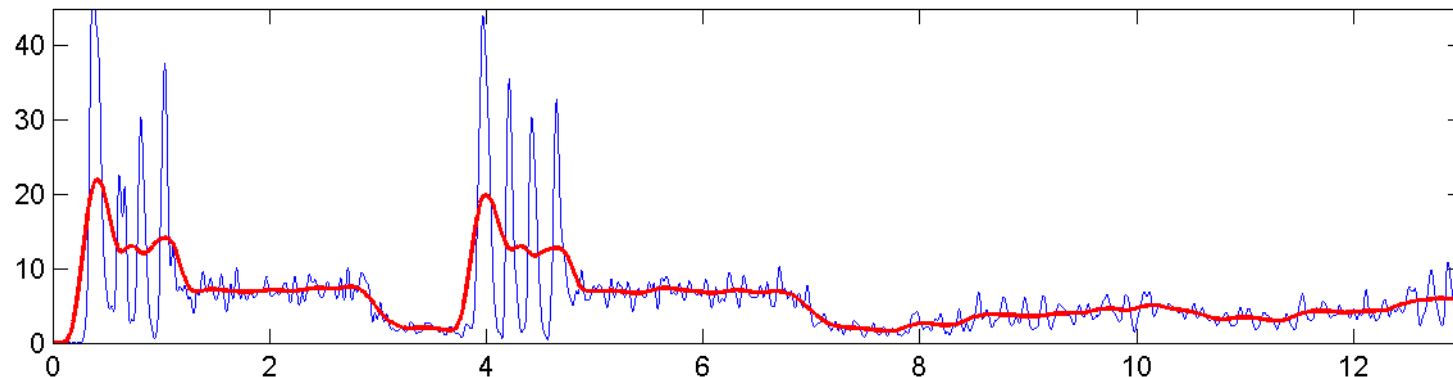
# Onset Detection (Spectral-Based)

## Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation
4. Accumulation
5. Normalization

Novelty curve

Substraction of local average

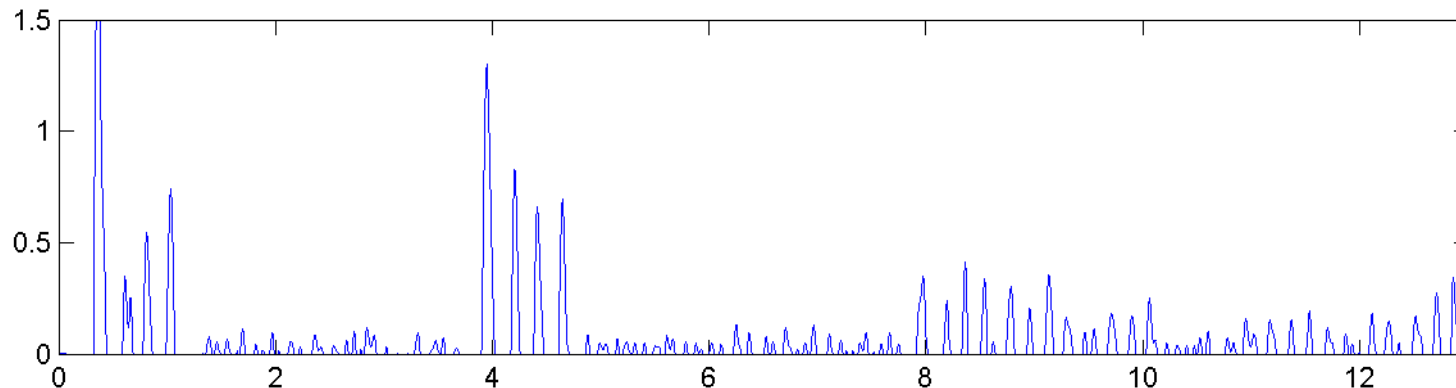


# Onset Detection (Spectral-Based)

## Steps:

1. Spectrogram
2. Logarithmic compression
3. Differentiation
4. Accumulation
5. Normalization

Normalized novelty curve

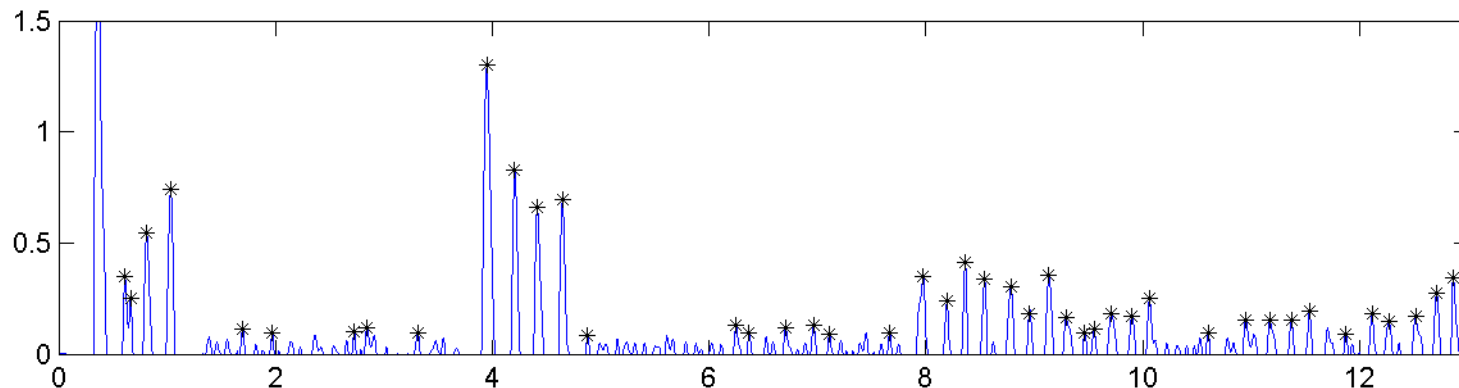


# Onset Detection (Spectral-Based)

## Steps:

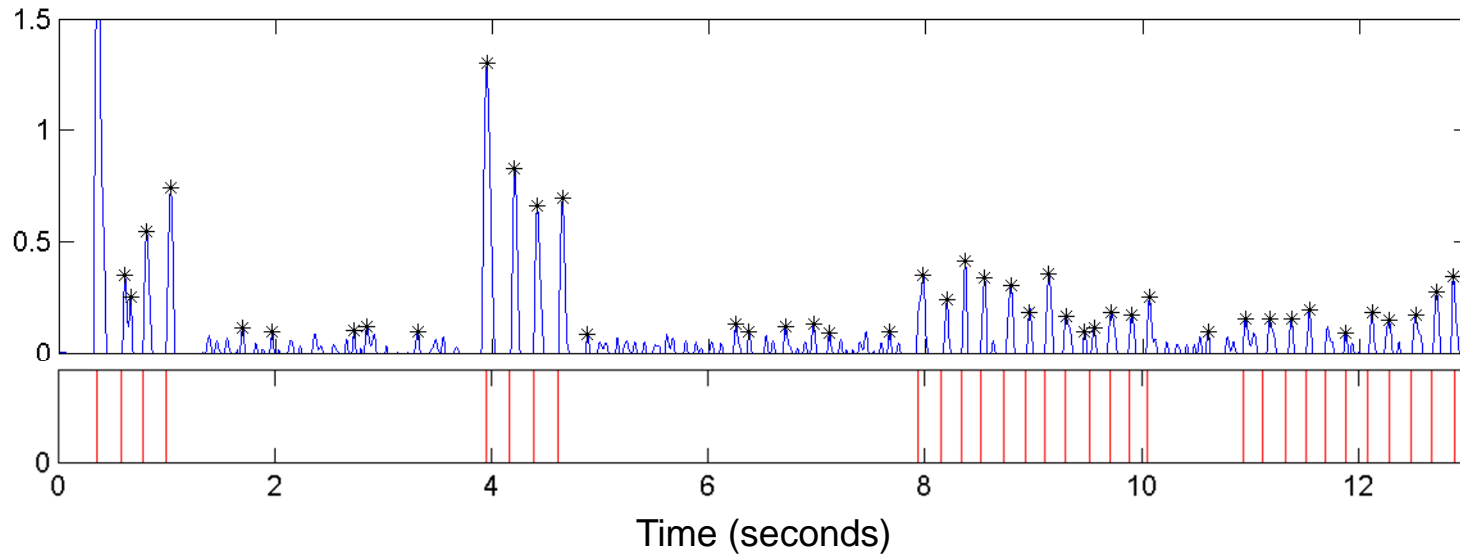
1. Spectrogram
2. Logarithmic compression
3. Differentiation
4. Accumulation
5. Normalization
6. Peak picking

Normalized novelty curve



# Onset Detection

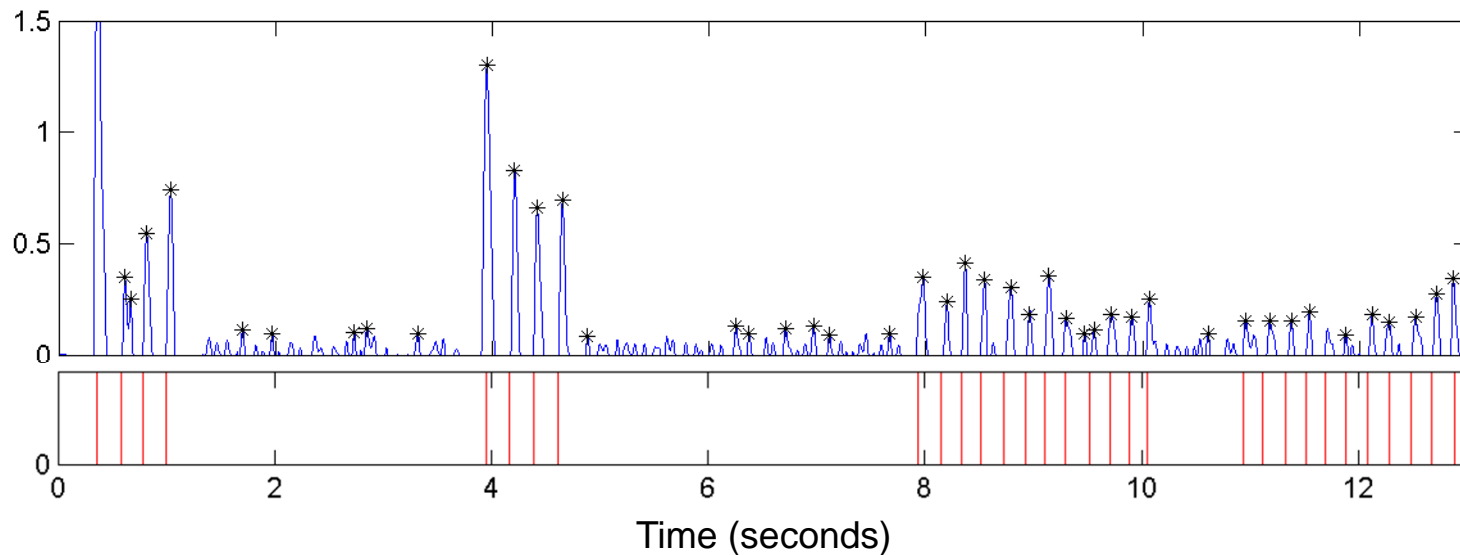
## Peak picking



- Peaks of the novelty curve indicate note onset candidates

# Onset Detection

## Peak picking

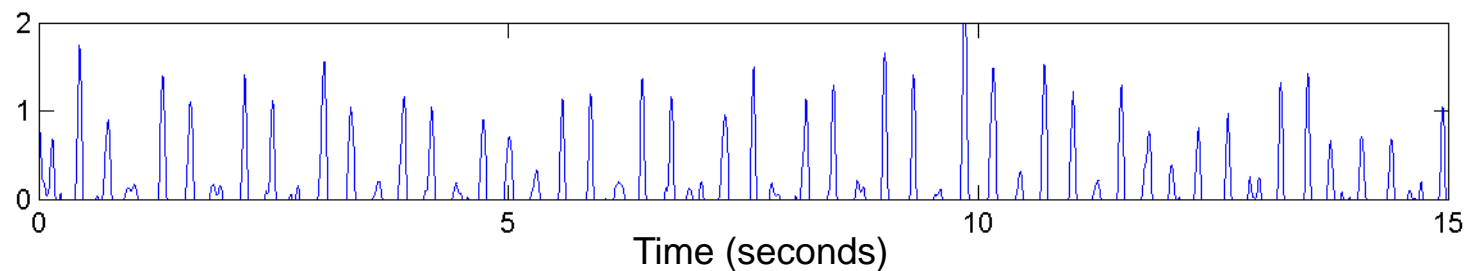


- Peaks of the novelty curve indicate note onset candidates
- In general many spurious peaks
- Usage of local thresholding techniques
- **Peak-picking very fragile step in particular for soft onsets**

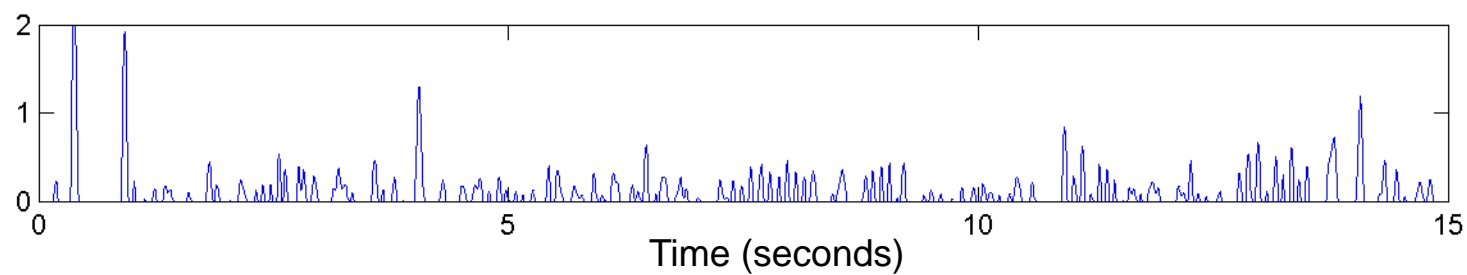


# Onset Detection

Shostakovich – 2<sup>nd</sup> Waltz



Borodin – String Quartet No. 2



---

# Onset Detection

Drumbeat



Going Home



Lyphard melodie



Por una cabeza



Donau



# Tempo Estimation and Beat Tracking

## What is a beat?

- Steady pulse that drives music forward and provides the temporal framework of a piece of music [Parncutt 1994]  
[Sethares 2007]  
[Large/Palmer 2002]
- Sequence of perceived pulses that are equally spaced in time [Lerdahl/ Jackendoff 1983]  
[Fitch/ Rosenfeld 2007]
- The pulse a human taps along when listening to the music

The term **tempo** then refers to the speed of the pulse.

# Tempo Estimation and Beat Tracking

## Strategy

- Analyze the novelty curve with respect to reoccurring or quasi-periodic patterns
- Avoid the explicit determination of note onsets (no peak picking)

# Tempo Estimation and Beat Tracking

## Strategy

- Analyze the novelty curve with respect to reoccurring or quasi-periodic patterns
- Avoid the explicit determination of note onsets (no peak picking)

[Scheirer, JASA 1998]

## Methods

- Comb-filter methods
- Autocorrelation
- Fourier transform

[Ellis, JNMR 2007]

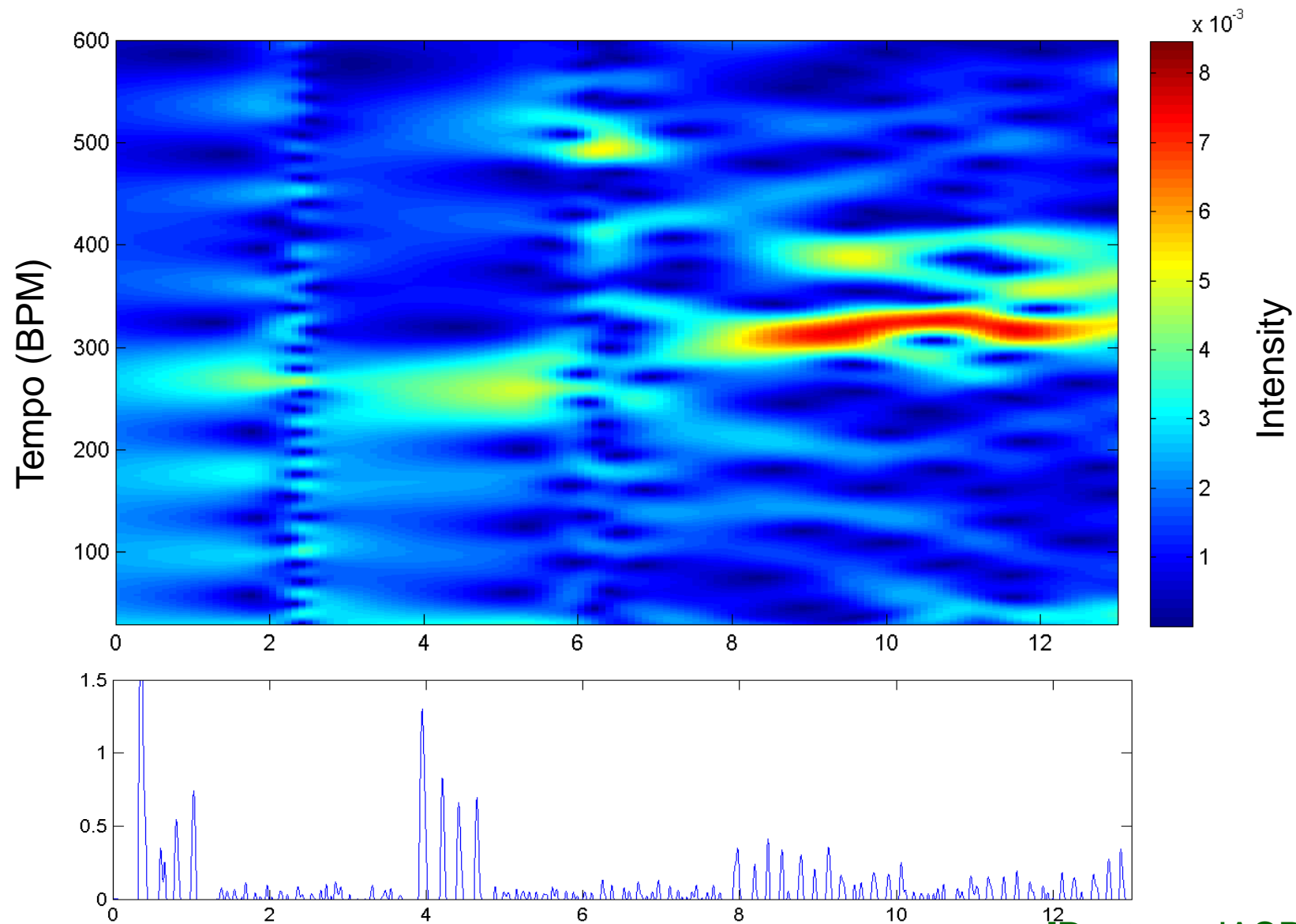
[Davies/Plumbley, IEEE-TASLP 2007]

[Peeters, JASP 2007]

[Grosche/Müller, ISMIR 2009]

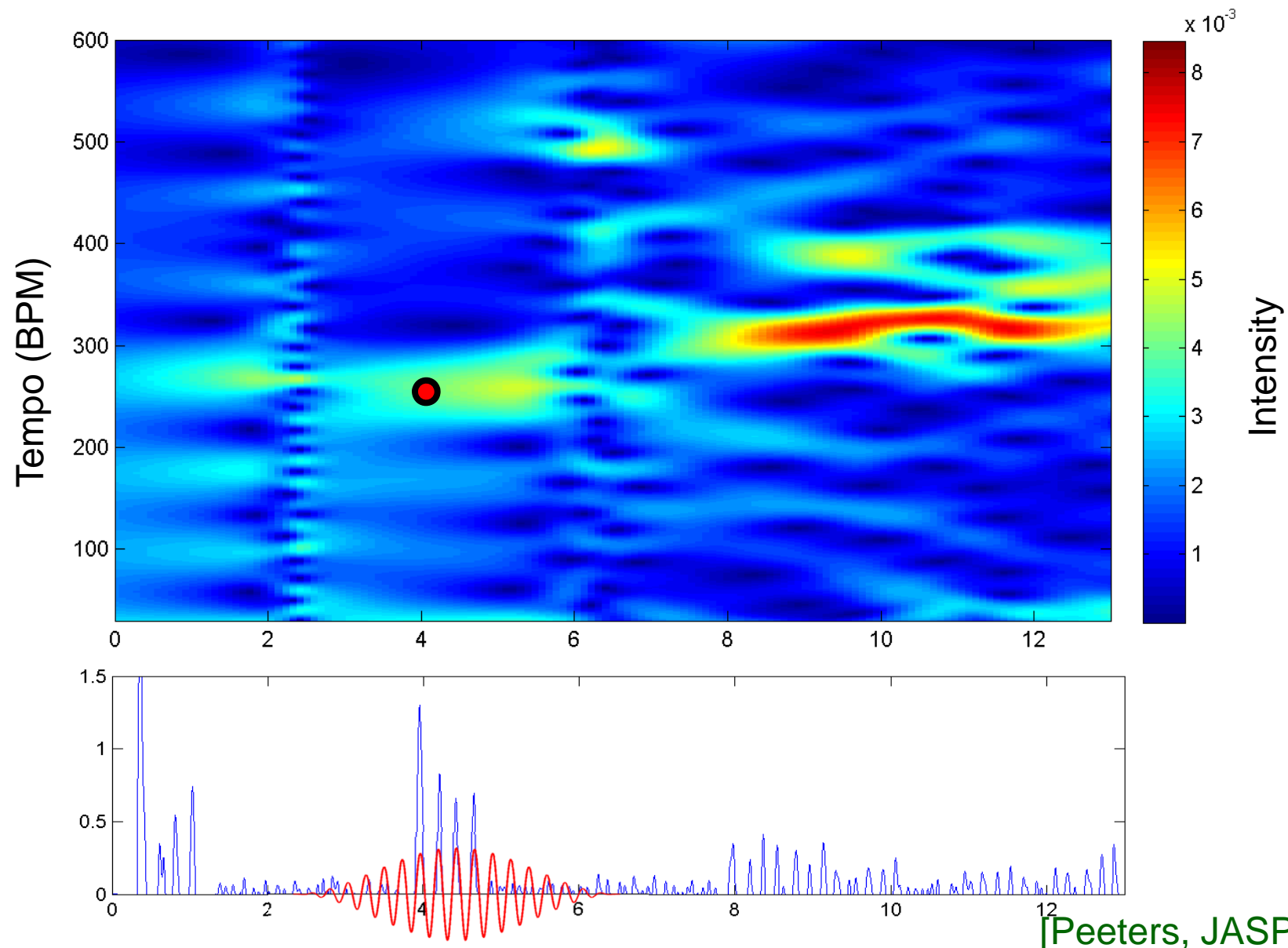
[Grosche/Müller, IEEE-TASLP 2011]

# Tempo Estimation and Beat Tracking



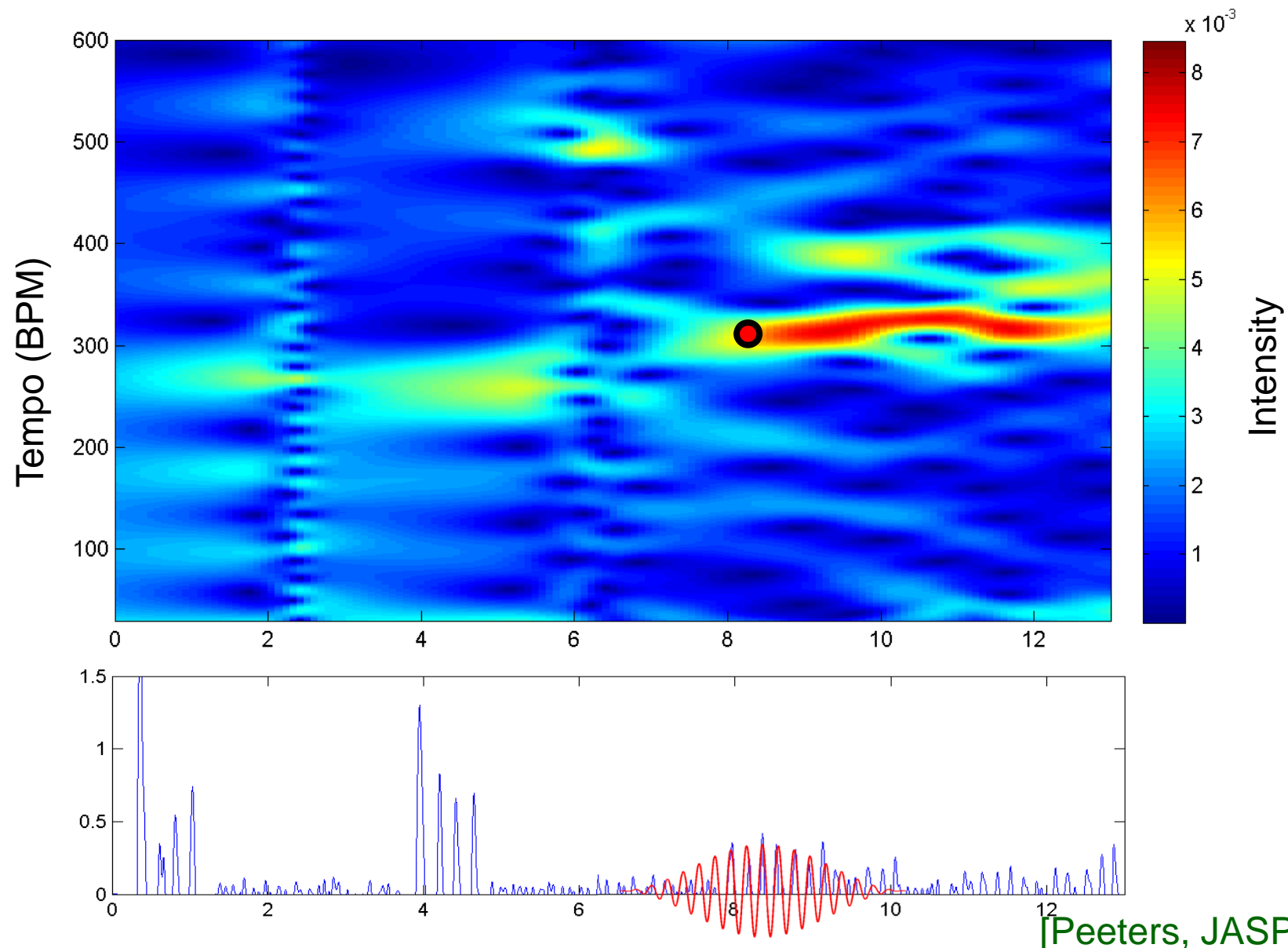
[Peeters, JASP 2007]

# Tempo Estimation and Beat Tracking



[Peeters, JASP 2007]

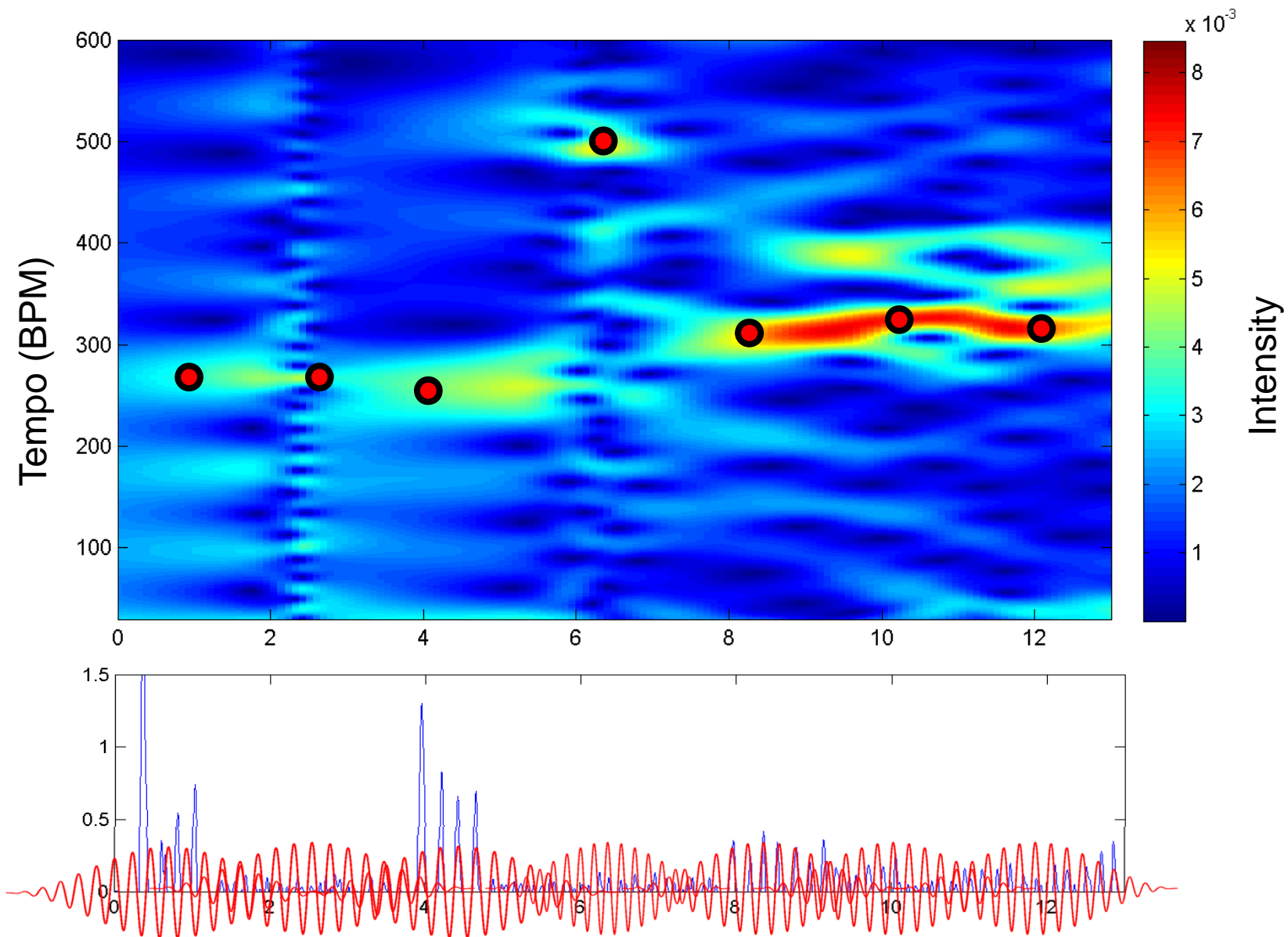
# Tempo Estimation and Beat Tracking



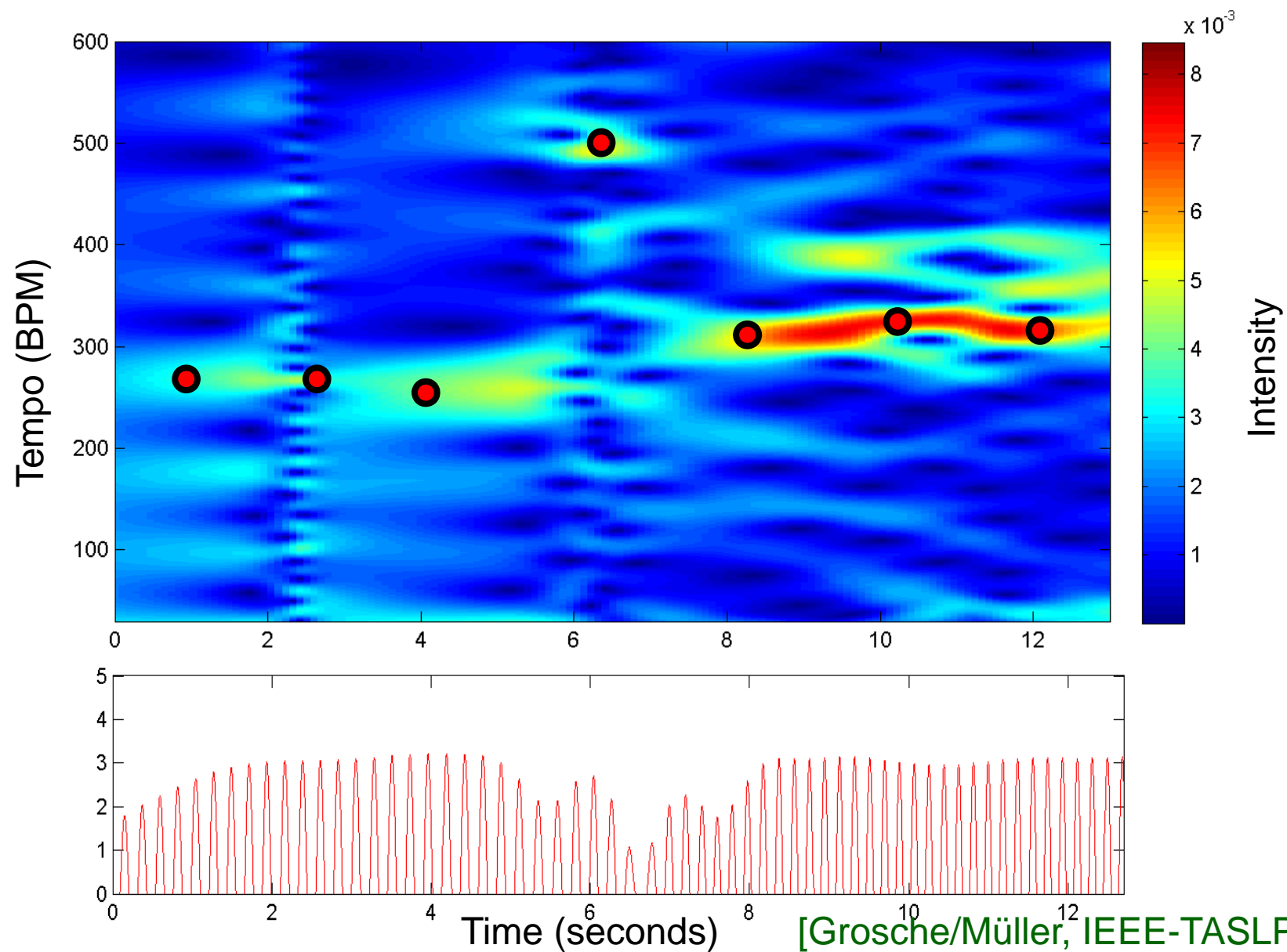
[Peeters, JASP 2007]



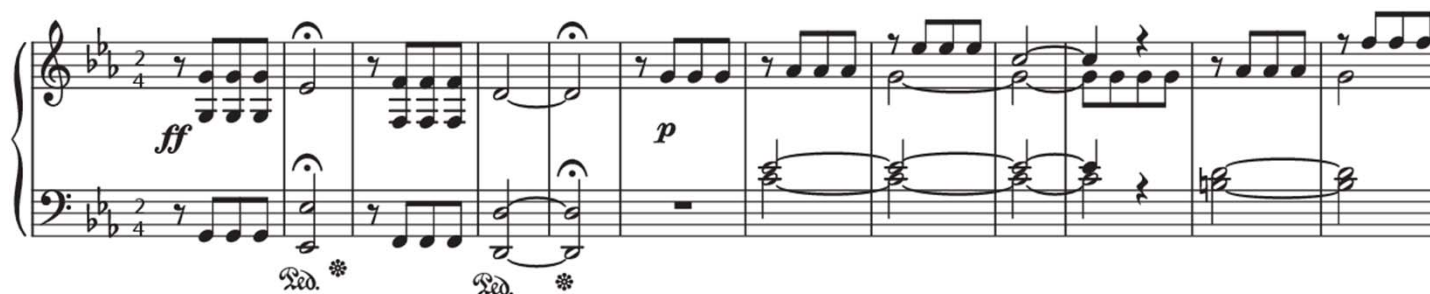
# Tempo Estimation and Beat Tracking



# Tempo Estimation and Beat Tracking

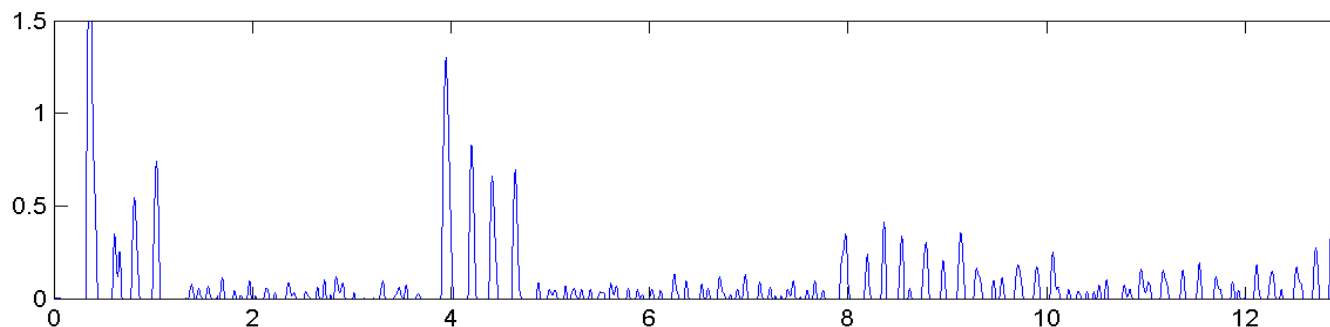


# Tempo Estimation and Beat Tracking

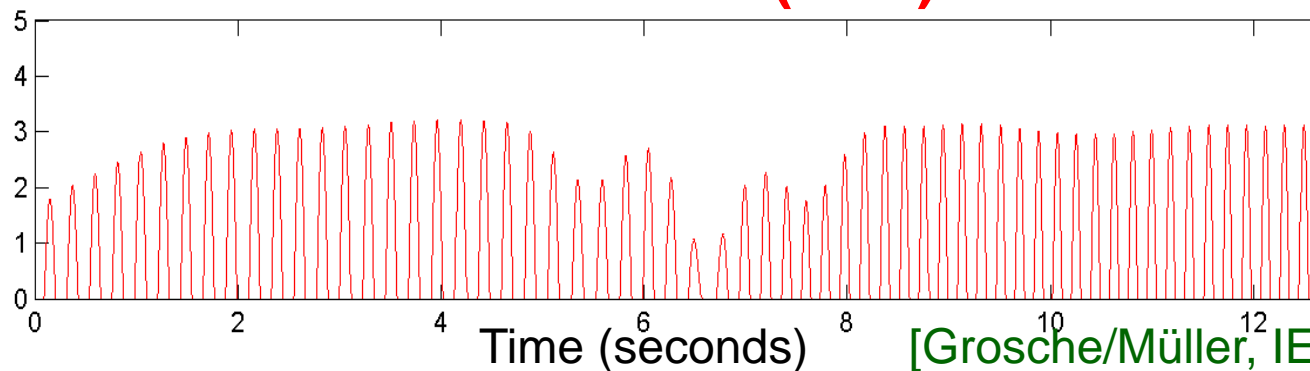


A musical score for piano in 2/4 time. The score is divided into two systems. The first system starts with a fortissimo (*ff*) dynamic and includes performance markings *rit.* and *\** under the first two measures. The second system starts with a piano (*p*) dynamic and includes *rit.* and *\** under the first two measures. The score features a melody in the right hand and a bass line in the left hand.

## Novelty Curve



## Predominant Local Pulse (PLP)



[Grosche/Müller, IEEE-TASLP 2011]

# Tempo Estimation and Beat Tracking

## Novelty Curve

- Indicates note onset candidates
- Extraction errors in particular for soft onsets
- Simple peak-picking problematic



## Predominant Local Pulse (PLP)

- Periodicity enhancement of novelty curve
- Accumulation introduces error robustness
- Locality of kernels handles tempo variations



# Tempo Estimation and Beat Tracking

- Local tempo at time  $t$  :  $\tau_t \in \Theta$   $\Theta = [60:240]$  BPM

- Phase  $\varphi_t := \frac{1}{2\pi} \arccos \left( \frac{\operatorname{Re}(\mathcal{T}(t, \tau_t))}{|\mathcal{T}(t, \tau_t)|} \right)$

- Sinusoidal kernel  $\kappa_t : \mathbb{Z} \rightarrow \mathbb{R}$

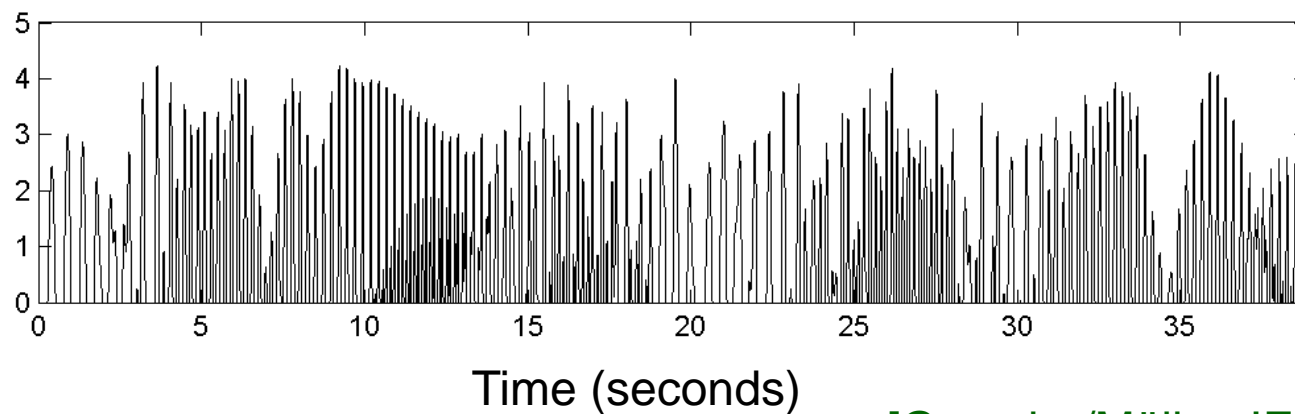
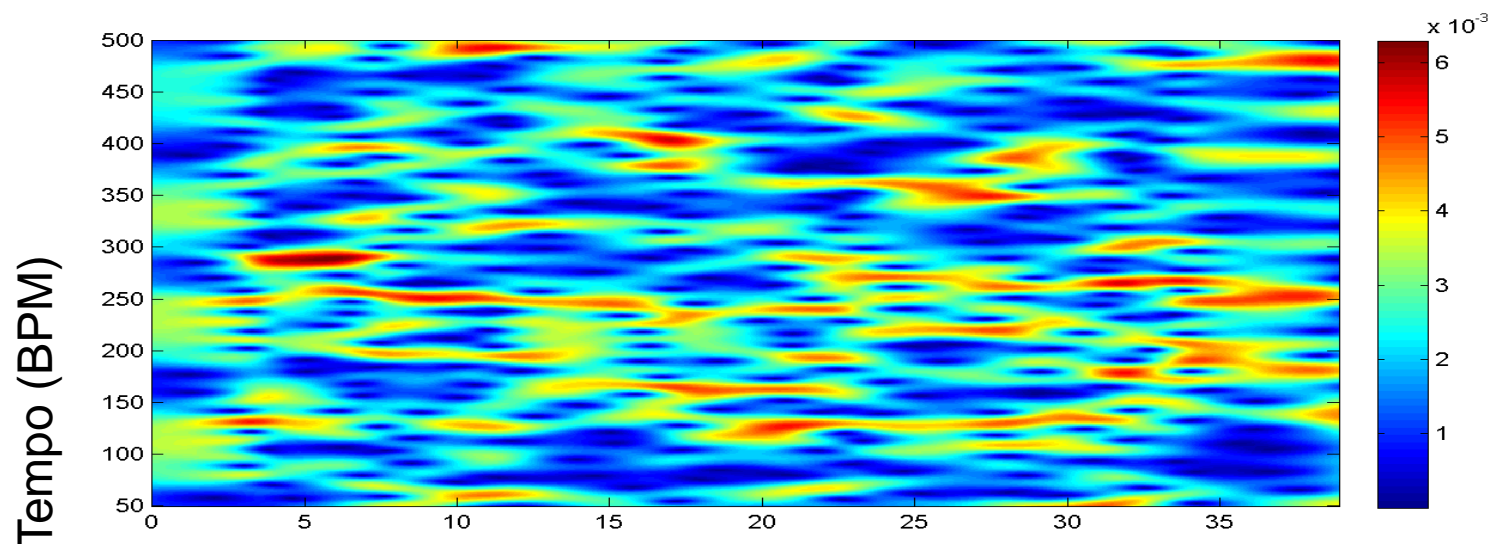
$$\kappa_t(n) := W(n - t) \cos(2\pi(\tau_t/60 \cdot n - \varphi_t)) \quad n \in \mathbb{Z}$$

- Periodicity curve  $\Gamma : [1 : T] \rightarrow \mathbb{R}_{\geq 0}$

$$\Gamma(n) = \left| \sum_{t \in [1:T]} \kappa_t(n) \right|_{\geq 0} \quad n \in [1 : T]$$

# Tempo Estimation and Beat Tracking

Borodin – String Quartet No. 2

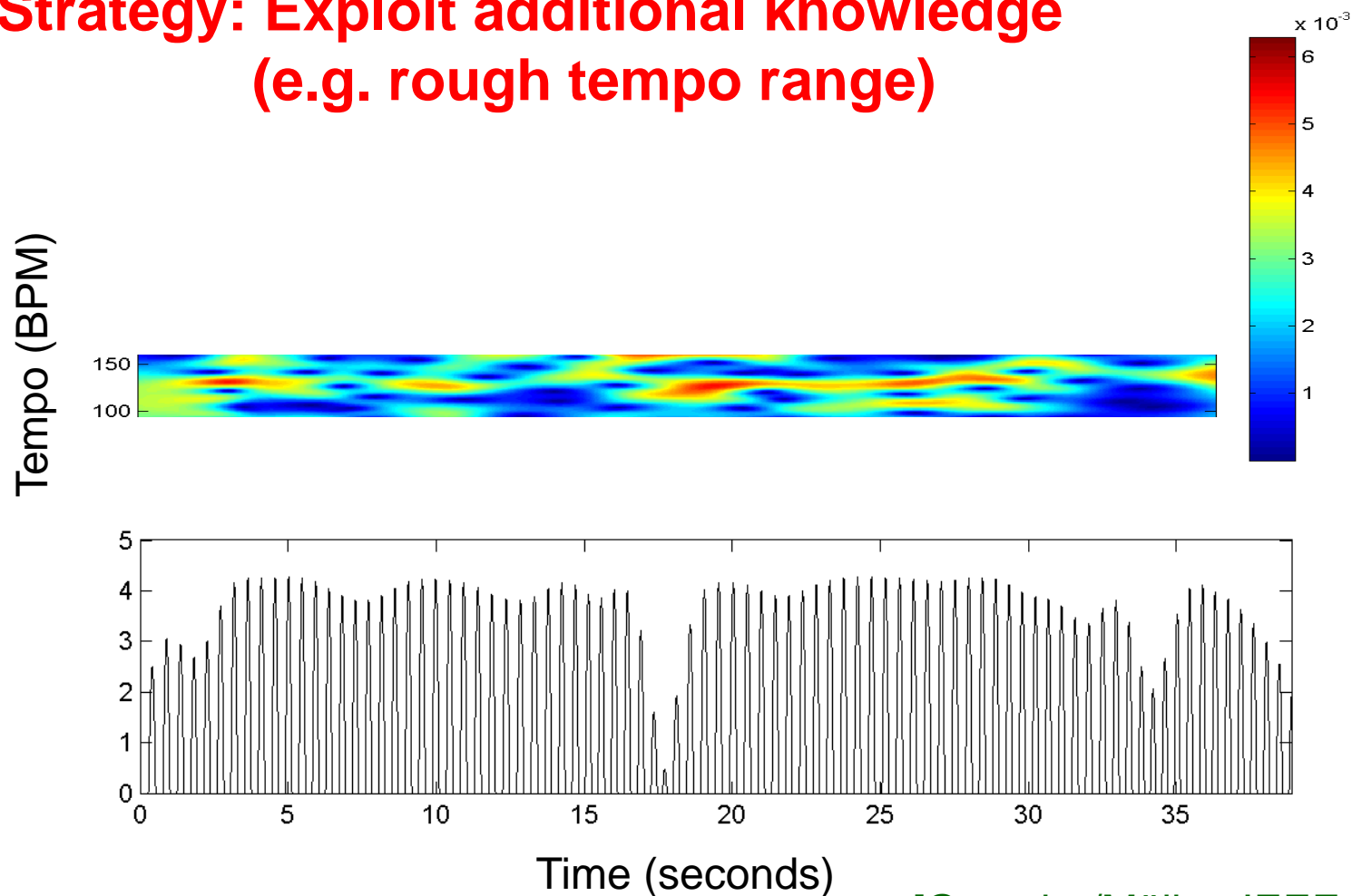


[Grosche/Müller, IEEE-TASLP 2011]

# Tempo Estimation and Beat Tracking

Borodin – String Quartet No. 2

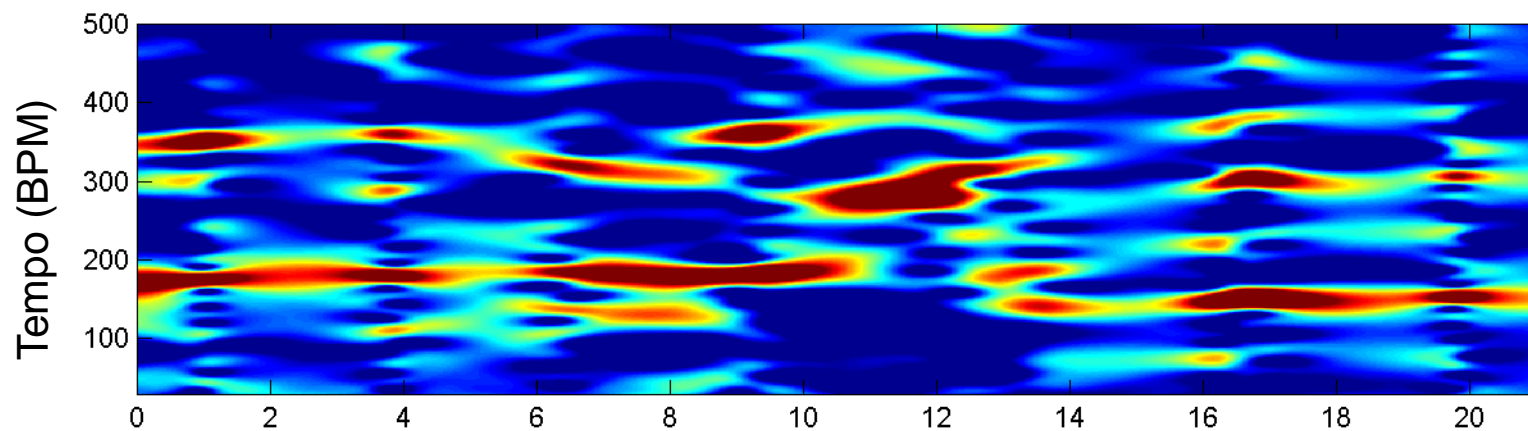
**Strategy: Exploit additional knowledge  
(e.g. rough tempo range)**



[Grosche/Müller, IEEE-TASLP 2011]

# Tempo Estimation and Beat Tracking

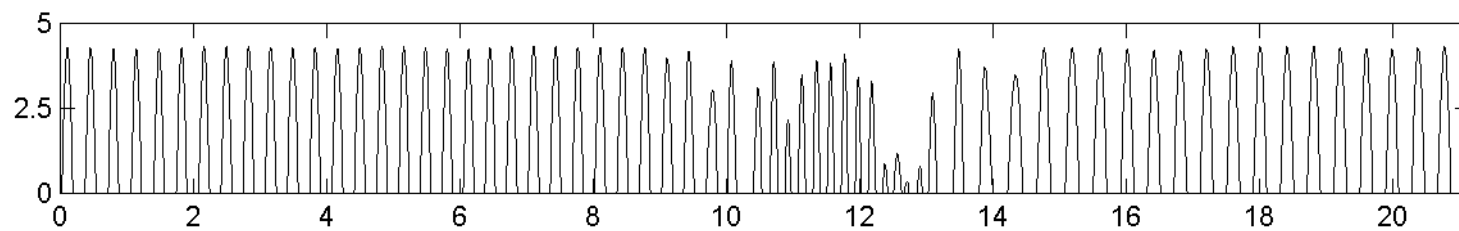
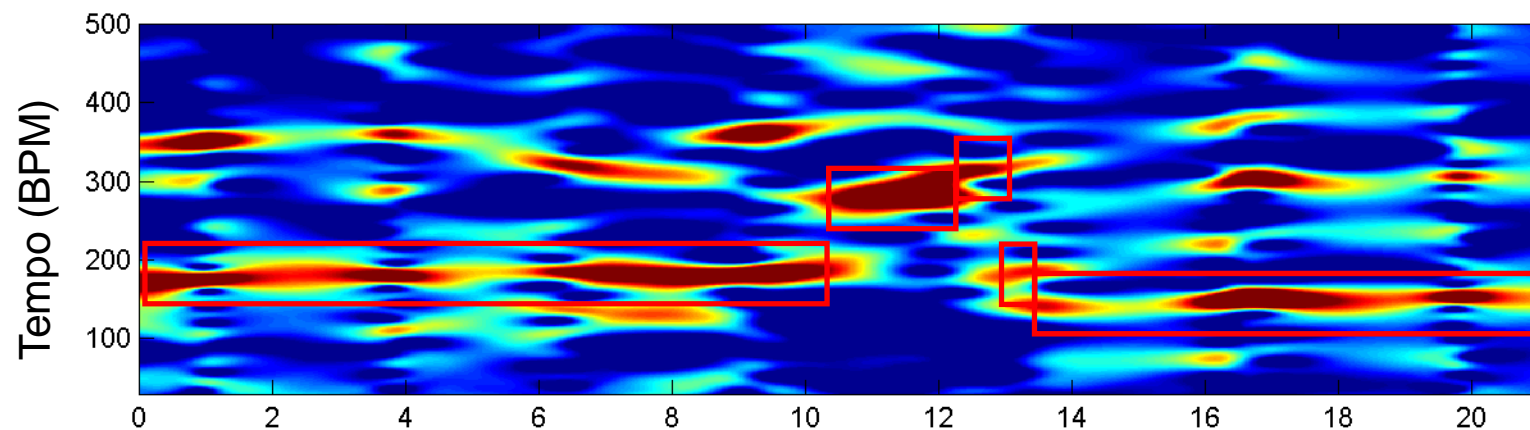
Brahms Hungarian Dance No. 5





# Tempo Estimation and Beat Tracking

Brahms Hungarian Dance No. 5



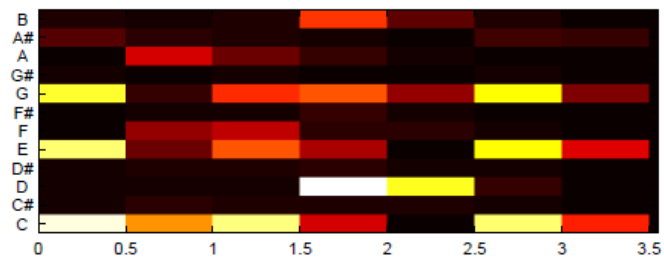
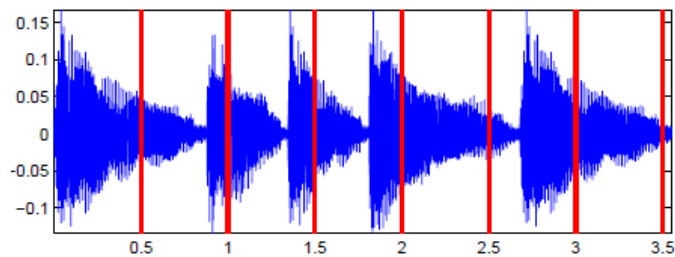
Time (seconds)



# Applications

- Feature design  
(beat-synchronous features, adaptive windowing)
- Digital DJ / audio editing  
(mixing and blending of audio material)
- Music classification
- Music recommendation
- Performance analysis  
(extraction of tempo curves)

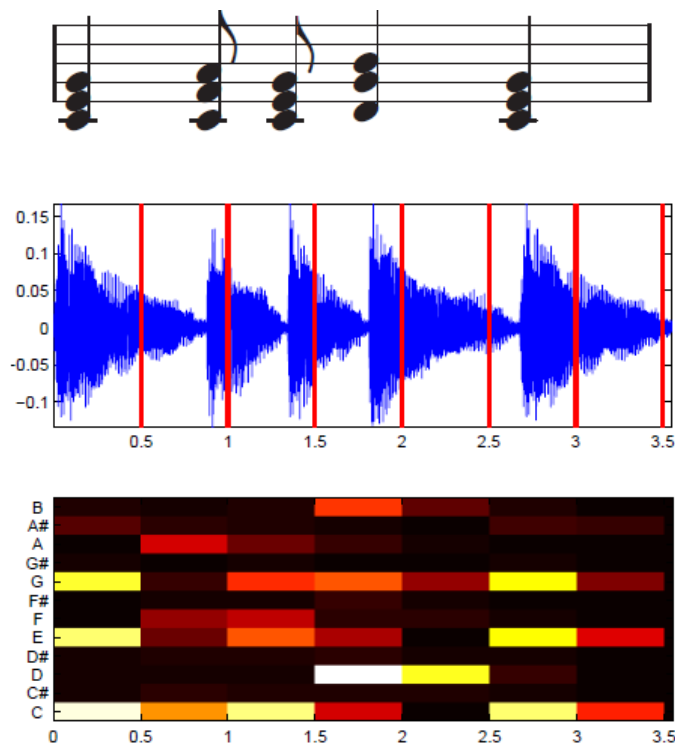
# Application: Feature Design



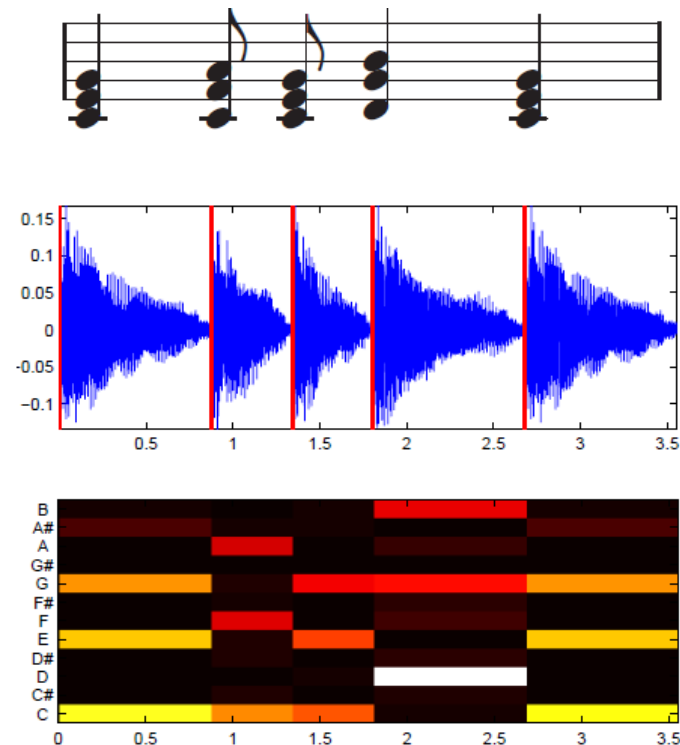
**Fixed window size**

[Ellis et al., ICASSP 2008] [Bello/Pickens, ISMIR 2005]

# Application: Feature Design

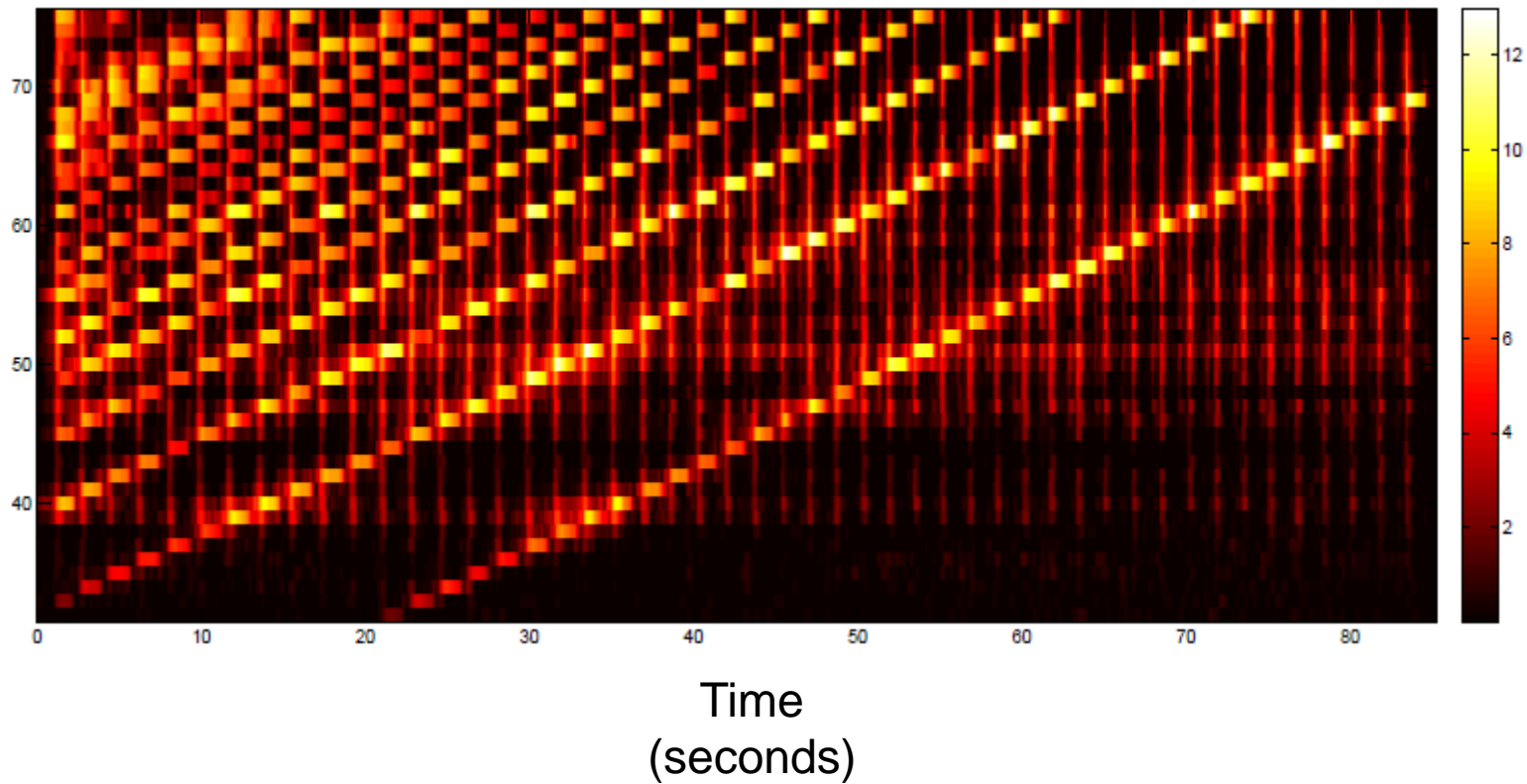


**Fixed window size**



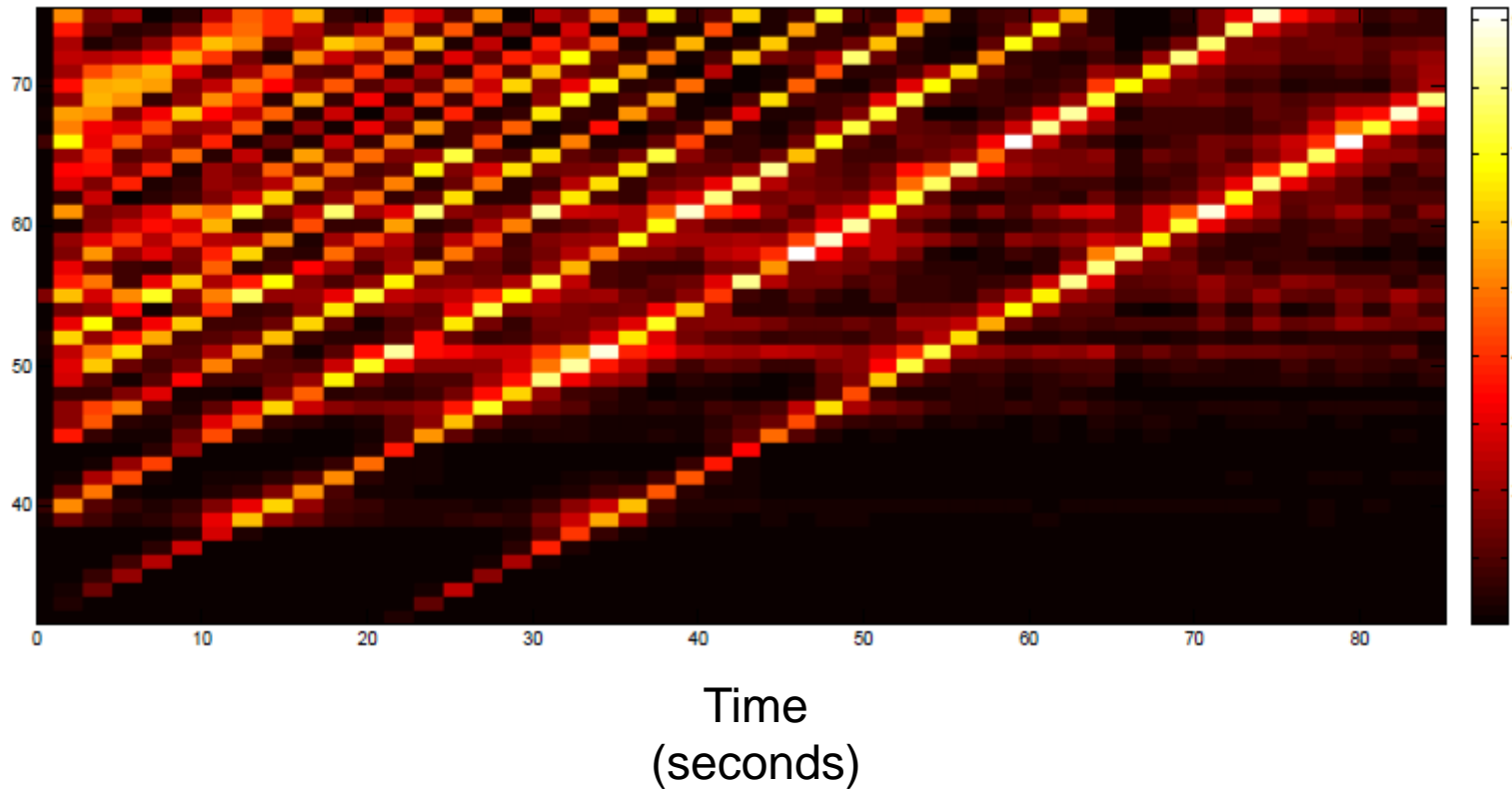
**Adaptive window size**

# Application: Feature Design



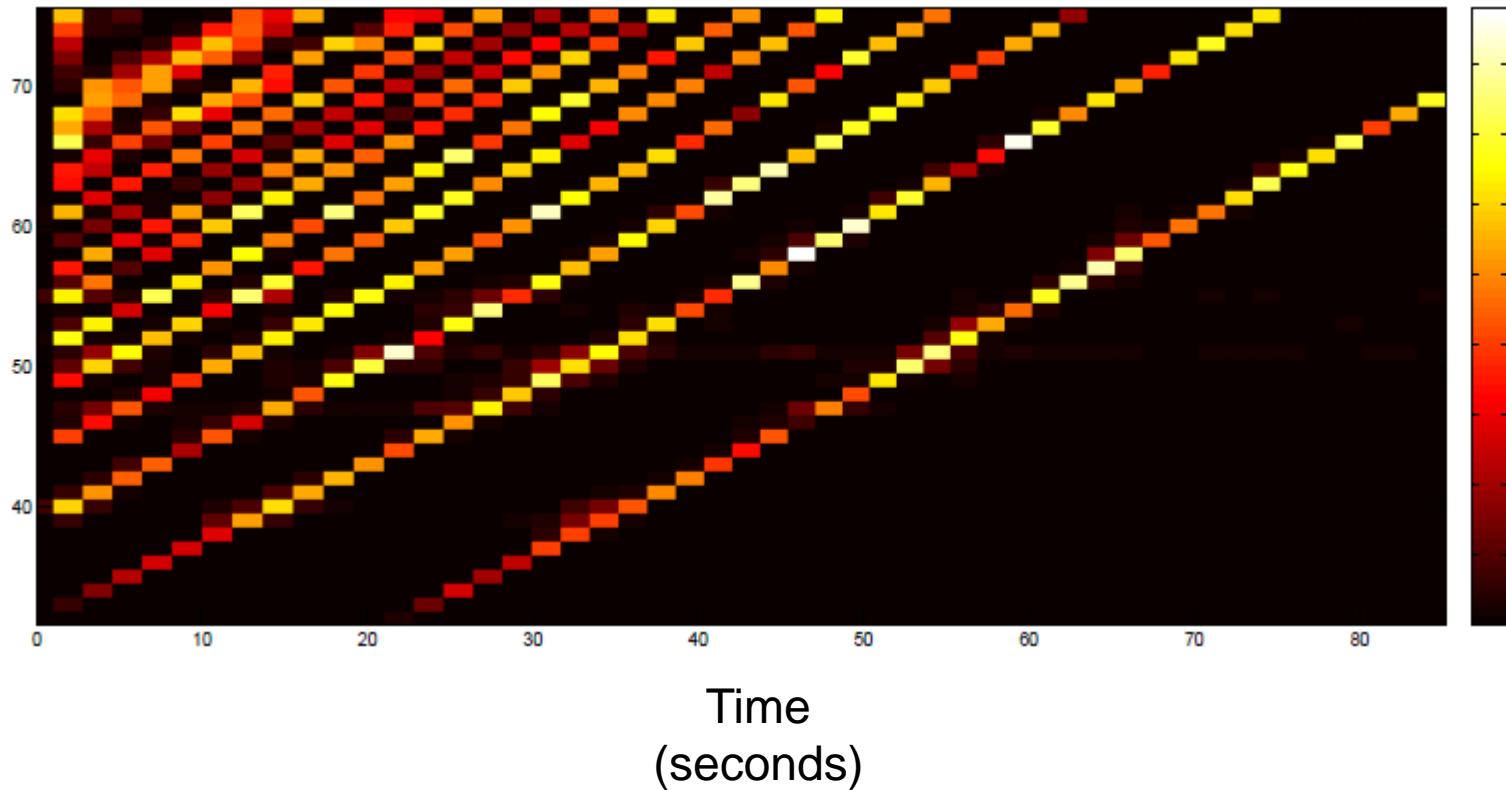
**Fixed window size (100 ms)**

# Application: Feature Design



**Adaptive window size (roughly 1200 ms)**  
**Note onset positions define boundaries**

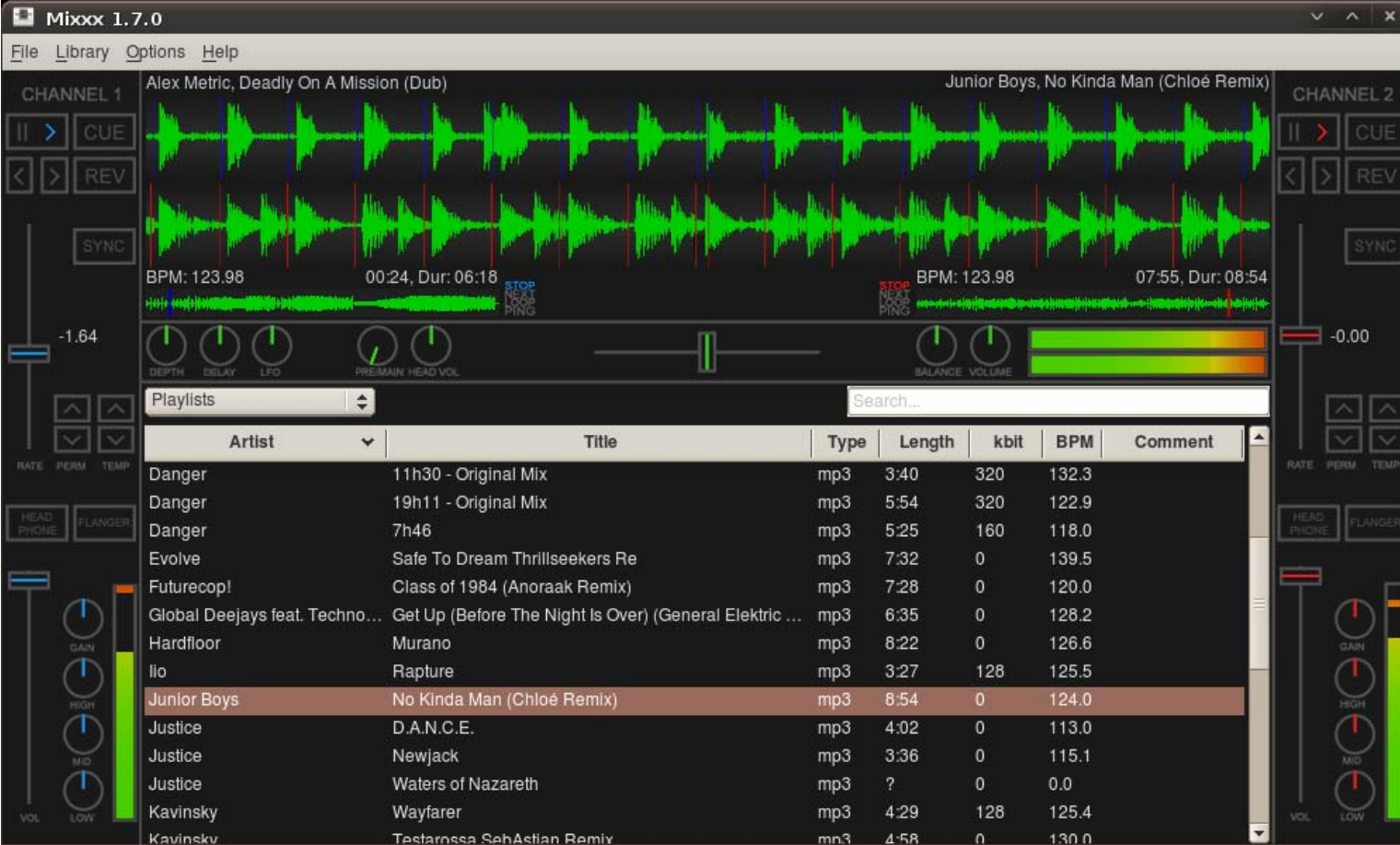
# Application: Feature Design



**Adaptive window size (roughly 1200 ms)**  
**Note onset positions define boundaries**

**Denoising by excluding boundary neighborhoods**

# Application: Audio Editing (Digital DJ)



The screenshot displays the Mixxx 1.7.0 software interface. At the top, the title bar reads "Mixxx 1.7.0" and the menu bar includes "File", "Library", "Options", and "Help".

**CHANNEL 1:** Alex Metric, Deadly On A Mission (Dub). BPM: 123.98. 00:24, Dur: 06:18. Waveform is green.

**CHANNEL 2:** Junior Boys, No Kinda Man (Chloé Remix). BPM: 123.98. 07:55, Dur: 08:54. Waveform is green.

Controls include: CUE, REV, SYNC, -1.64, RATE, PERM, TEMP, HEAD PHONE, FLANGER, GAIN, HIGH, MID, LOW, VOL, DEPTH, DELAY, LFO, PREMAIN, HEAD VOL, STOP, NEXT, LOOP, PING, BALANCE, VOLUME, -0.00, RATE, PERM, TEMP, HEAD PHONE, FLANGER, GAIN, HIGH, MID, LOW, VOL.

**Playlists Table:**

Artist	Title	Type	Length	kbit	BPM	Comment
Danger	11h30 - Original Mix	mp3	3:40	320	132.3	
Danger	19h11 - Original Mix	mp3	5:54	320	122.9	
Danger	7h46	mp3	5:25	160	118.0	
Evolve	Safe To Dream Thrillseekers Re	mp3	7:32	0	139.5	
Futurecop!	Class of 1984 (Anoraak Remix)	mp3	7:28	0	120.0	
Global Deejays feat. Techno...	Get Up (Before The Night Is Over) (General Electric ...	mp3	6:35	0	128.2	
Hardfloor	Murano	mp3	8:22	0	126.6	
lio	Rapture	mp3	3:27	128	125.5	
<b>Junior Boys</b>	<b>No Kinda Man (Chloé Remix)</b>	<b>mp3</b>	<b>8:54</b>	<b>0</b>	<b>124.0</b>	
Justice	D.A.N.C.E.	mp3	4:02	0	113.0	
Justice	Newjack	mp3	3:36	0	115.1	
Justice	Waters of Nazareth	mp3	?	0	0.0	
Kavinsky	Wayfarer	mp3	4:29	128	125.4	
Kavinsky	Testarossa SebAstian Remix	mp3	4:58	0	130.0	



# Application: Beat-Synchronous Light Effects

