Topic 5

Rhythm Analysis

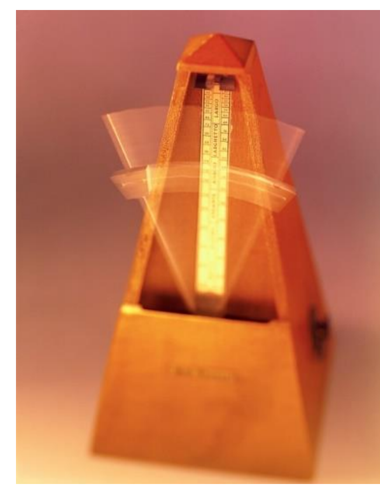
(some slides are adapted from Zafar Rafii and some figures are from Meinard Mueller)

Definitions for Rhythm Analysis

- **Rhythm:** "movement marked by the regulated succession of strong and weak elements, or of opposite or different conditions."

  ---- Oxford English Dictionary

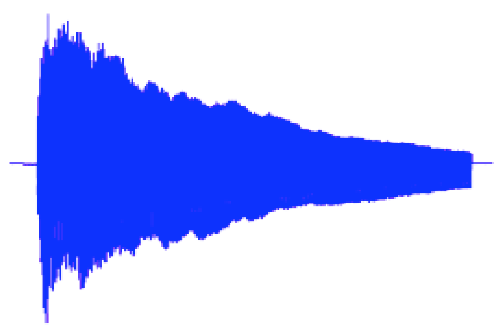
- **Beat:** basic unit of time in music
- **Tempo:** speed or pace of a given piece, typically measured in beats per minute (BPM)

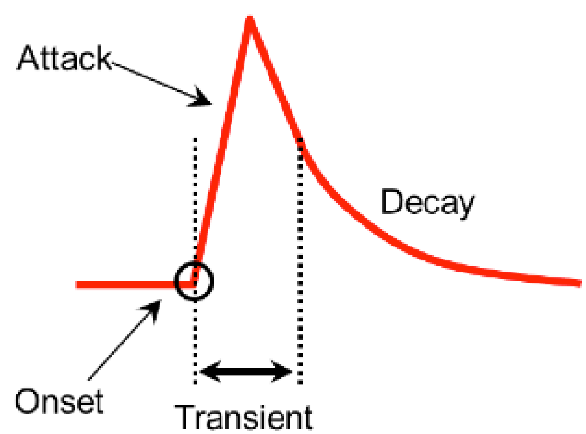


More Definitions

- **Onset**: single instant marking the beginning of transient
  - Onsets often occur on beats.
- **Attack**: sharp increase of energy
- **Transient**: a short duration with high amplitude within which signal evolves quickly

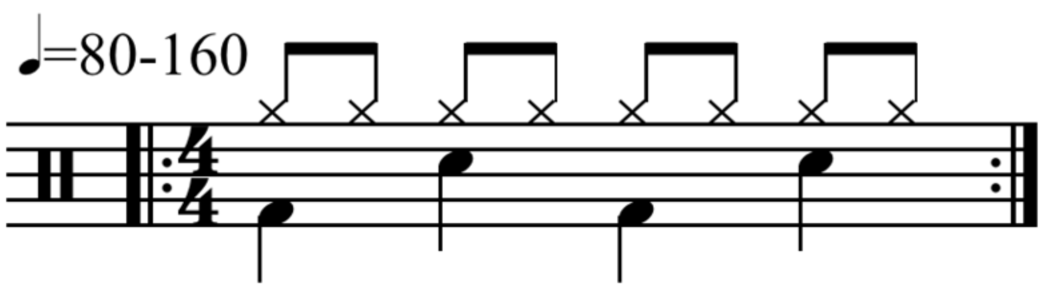
Waveform of one piano note





More Definitions

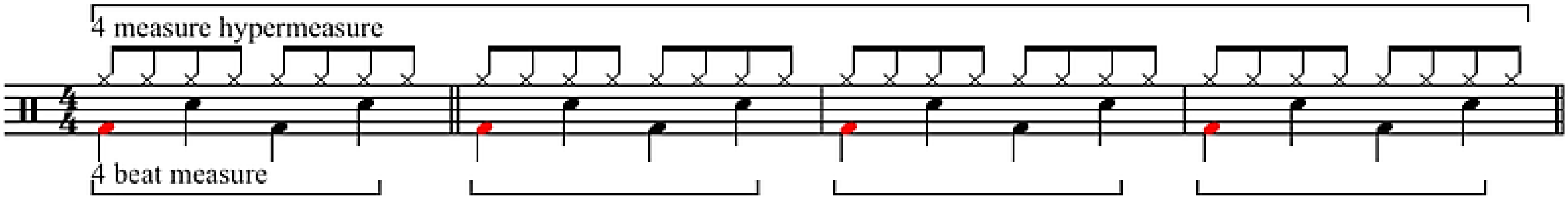
- **Measure (or bar):** segment of time defined by a given number of beats



A 4-beat measure drum pattern.
[http://en.wikipedia.org/wiki/Metre_(music)]

More Definitions

- **Meter:** organization of music into regularly recurring measures of stressed and unstressed beats



Hypermeter: 4-beat measure and 4-measure hypermeasure. Hyperbeats in red. [http://en.wikipedia.org/wiki/Metre_(music)]

Rhythm Analysis Tasks

- Onset Detection
- Beat Tracking
- Tempo Estimation
- Higher-level Structure Analysis

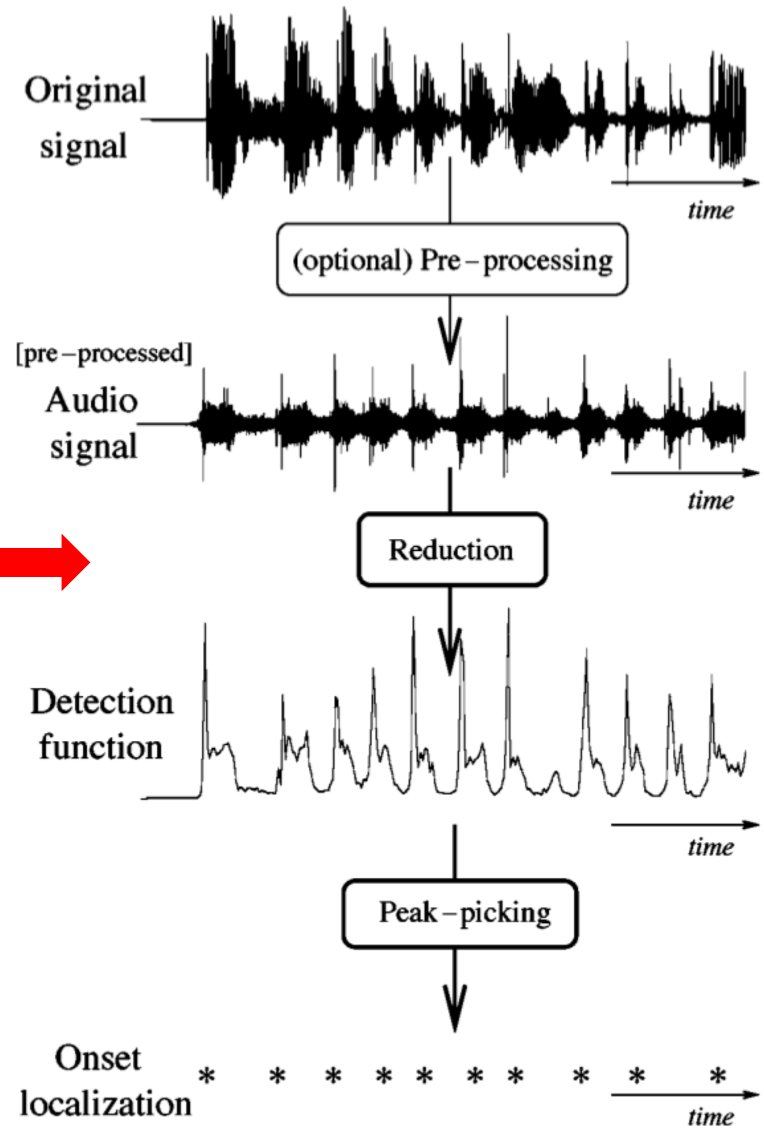
Why is it important?

- Intellectual merit
  – Important component of music understanding
  – Music cognition research

- Broad applications
  – Identify/classify/retrieve by rhythmic similarity
  – Music segmentation/summarization
  – Audio/video synchronization
  – Source separation

Onset Detection

- Signal processing: define a detection function
  - Energy-based
  - Spectral-based
  - Phase-based
- Machine Learning: learn patterns from labeled data
  - Probabilistic models
  - Neural networks



Energy-based Onset Detection

Waveform

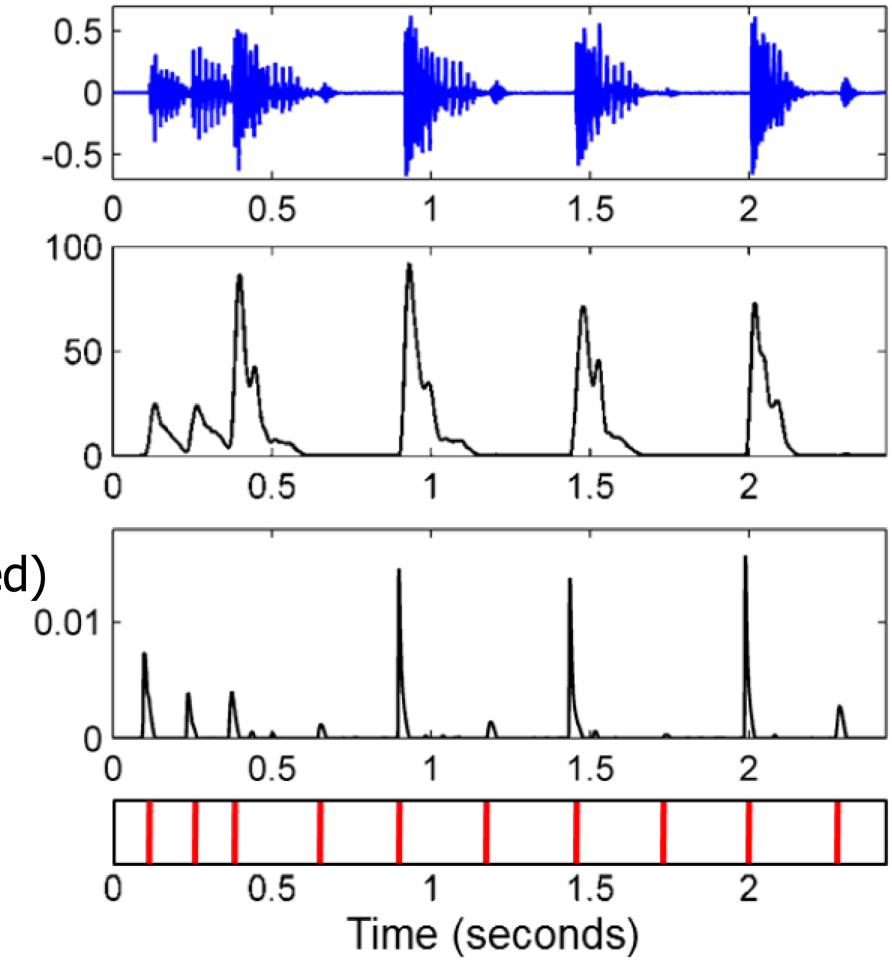
Signal Envelope (energy)

$$E_w^x(n) := \sum_{m=-M}^{M} |x(n+m)w(m)|^2$$

Envelope Derivative (half-wave rectified)

$$\Delta_{\text{Energy}}(n) := |E_w^x(n+1) - E_w^x(n)|_{\geq 0}$$

Thresholding → Onsets



Energy-based Onset Detection

- Pros and Cons
  - Simple
  - Works well for percussive sounds
  - Soft onsets by string/wind instruments are hard to detect
  - Tremolo/vibrato can cause false detections

- How to improve
  - Use logarithmic-energy to replace linear energy
  - Perform analysis in different frequency bands, then summarize

Spectral-based Onset Detection

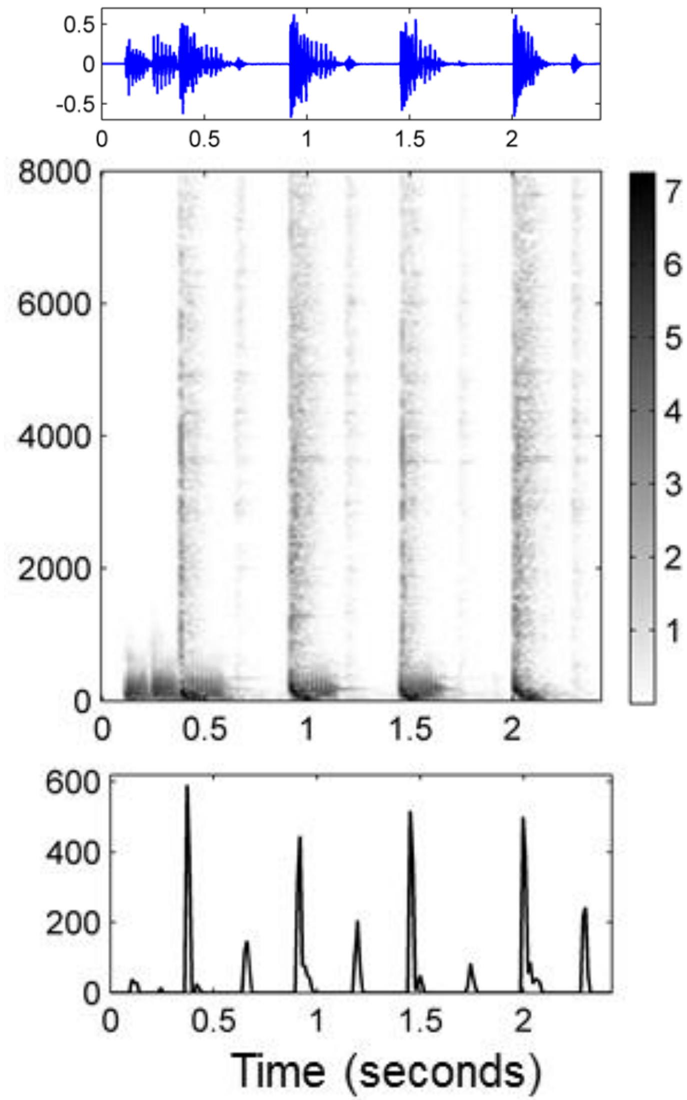
- STFT to get magnitude spectrogram $|\chi|$

- (optional) compression

$$\mathcal{Y} := \Gamma_\gamma(|\mathcal{X}|) = \log(1 + \gamma \cdot |\mathcal{X}|)$$

- Spectral flux:
  - Take derivative w.r.t. time (half-wave rectified)

$$\Delta_{\mathrm{Spectral}}(n) := \sum_{k=0}^{K} |\mathcal{Y}(n+1,k) - \mathcal{Y}(n,k)|_{\geq 0}$$



Spectral-based Onset Detection

- Pros and Cons
  - More complex than energy-based
  - Can weigh different frequencies differently
  - Works better for soft onsets (e.g., legato notes) and polyphonic music
  - Still doesn’t work very well for vibrato

Tempo Estimation

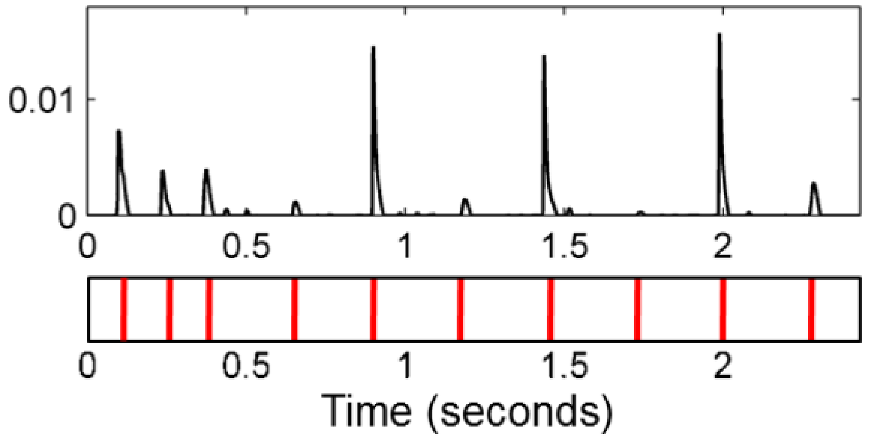
- Tempo = beats / minutes
- Beat tracking is a sufficient but not a necessary condition for tempo estimation
- How to estimate tempo without tracking beats?

- Idea: look at the regularity of onsets
- Assumptions
  - Onsets mostly occur on beats
  - Tempo is constant within a period of time

Tempo Estimation

Onset strength curve

Onsets



- Take the onset strength curve and analyze its periodicity
  - Autocorrelation
  - STFT

Tempogram

Beat Tracking

- Identify the beat times, i.e., the times to which we tap our feet
  - Detected onsets provide useful but noisy information, since not all onsets are on beats.
  - Estimated tempo tells us the space between two beats, but not the exact locations (i.e., phase).

- How to identify beats?
- To simply the problem, we assume
  - Onsets, especially strong ones, are mostly on beats.
  - Tempo is constant.

Beat Tracking

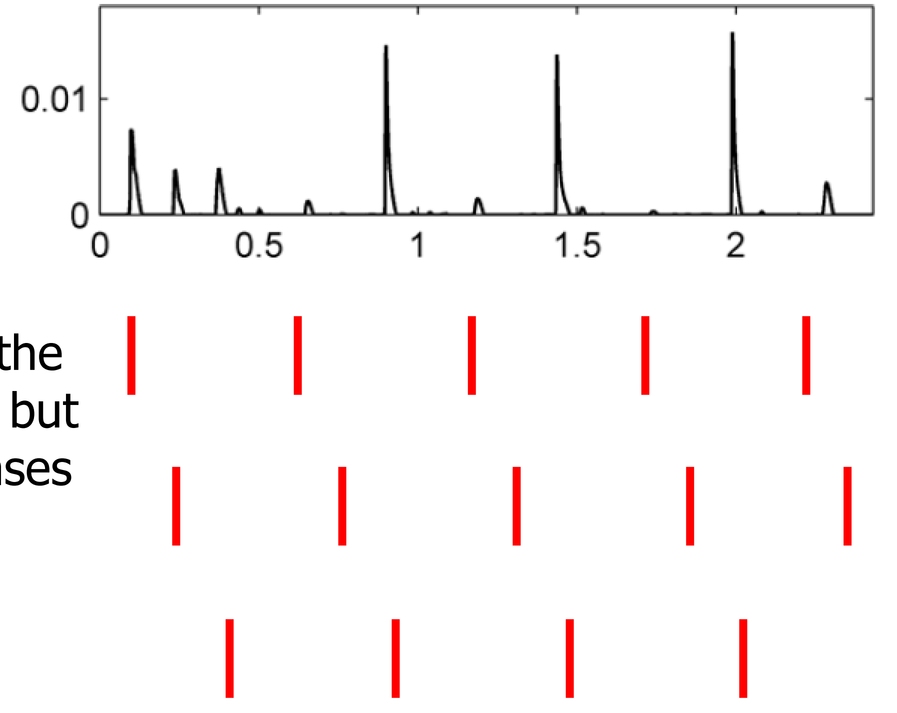
- A 2-step approach
  - Step 1: Tempo estimation
  - Step 2: Identify beats from onsets using the tempo
    - Create an impulse train (i.e., “comb”) with the tempo
    - Cross-correlate the “comb” with the onset strength curve.
    - The lag that gives us the highest cross-correlation value tells us the beat phase.

Beat Tracking

- A 2-step approach, illustration

Onset strength curve

Combs with the same tempo but different phases



- Problem: too rigid about beat spacing

Beat Tracking by Dynamic Programming

- Beat tracking: finding a sequence of beat locations such that they
  - 1) are well aligned with strong onsets
  - 2) mostly regularly spaced

[Ellis, 2007]

Score function

$$\mathbf{S}(B) := \sum_{\ell=1}^{L} \Delta(b_\ell) + \lambda \sum_{\ell=2}^{L} P_{\hat{\delta}}(b_\ell - b_{\ell-1})$$

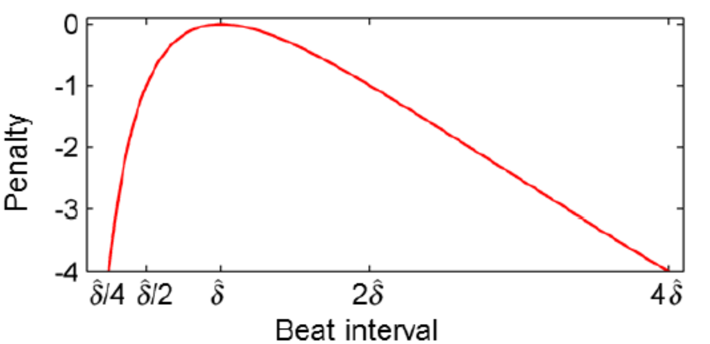
Beat sequence

Onset strength

Regularity penalty function

Rough estimate of beat spacing

$$P_{\hat{\delta}}(\delta) := -\left(\log_2(\delta/\hat{\delta})\right)^2$$



- Find $B = (b_1, b_2, \cdots, b_L)$ that maximizes $\boldsymbol{S}(B)$

Beat Tracking by Dynamic Programming

- Suppose beat locations are precise to audio frames, and suppose there are N frames, then how many possible sequences?
  - 2^N (although many are bad ones!)
  - Can't enumerate all!

- Key idea: reuse calculations by recursion!

Beat Tracking by Dynamic Programming

- Consider a beat sequence $B_n = (b_1, b_2, \cdots, b_L)$ where $b_L = n$.
- Let $D(n)$ be the maximal score over all such sequences ending at n, with various lengths.
- Assume B_n is the optimal sequence
- Then

if $L = 1$

$$\mathbf{D}(n) = \Delta(n)$$

recursion

if $L > 1$

$$\mathbf{D}(n) = \Delta(n) + \lambda P_{\hat{\delta}}(n - b_{L-1}) + \boxed{\mathbf{D}(b_{L-1})}$$

Beat Tracking by Dynamic Programming

- Considering the two cases, we have

$$\mathbf{D}(n) = \Delta(n) + \max \begin{cases} 0, \\ \max_{m \in [1:n-1]} \left\{ \mathbf{D}(m) + \lambda P_{\hat{\delta}}(n-m) \right\} \end{cases}$$

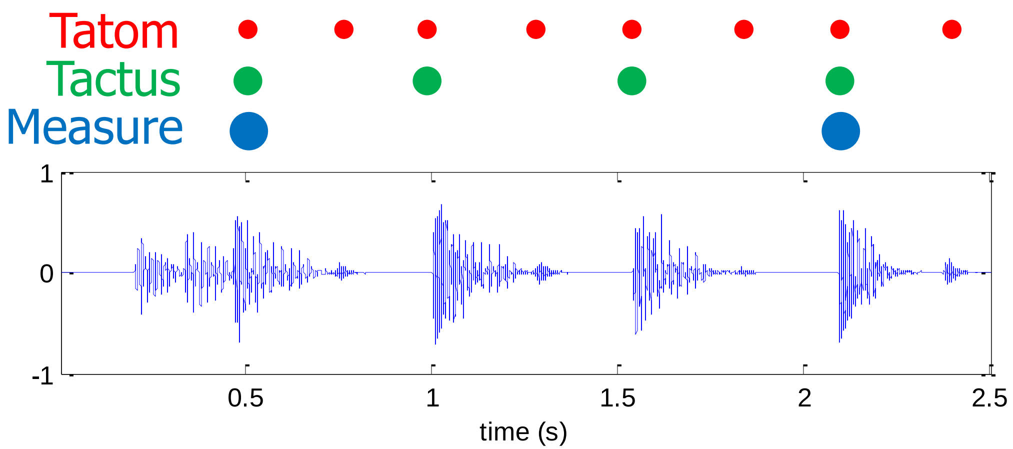
  - We can calculate $D(n)$ from $D(1) = \Delta(1)$.
- Record the preceding beat

$$\mathbf{P}(n) := \underset{m \in [1:n-1]}{\operatorname{argmax}} \left\{ \mathbf{D}(m) + \lambda P_{\hat{\delta}}(n-m) \right\}$$

- Best score $\mathbf{S}(B^*) = \max_{n \in [0:N]} \mathbf{D}(n)$

- Trace back from $b_L = n^*$ to get the best sequence

Rhythmic Structure



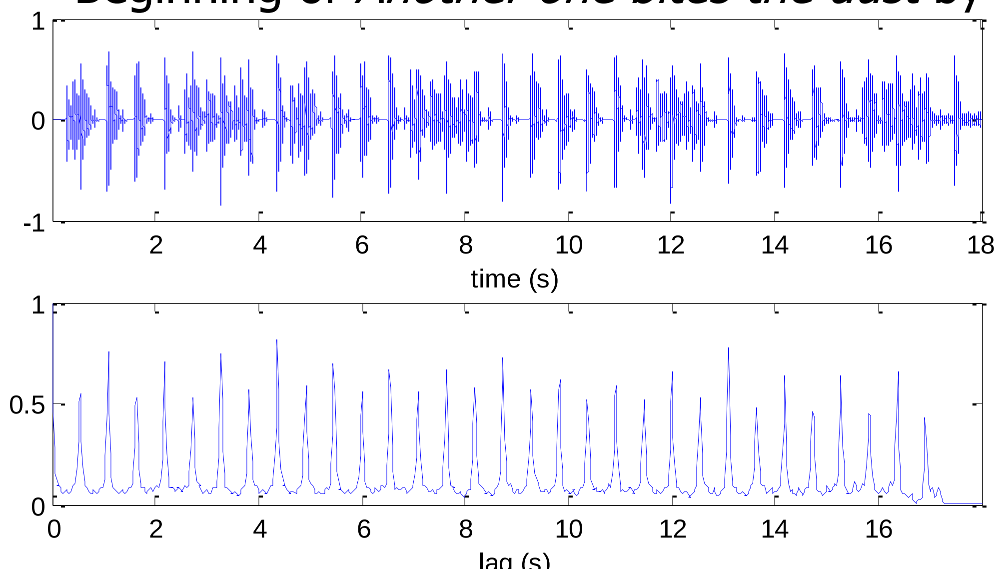
Beginning of *Another one bites the dust* by Queen.

- One approach: detect onsets; analyze tempo and beats at different levels.
- Another approach: analyze repetition of spectral content
  - Beat spectrum

Beat Spectrum

- Definition
  - Using the autocorrelation function, we can derive the beat spectrum [Foote et al., 2001]

Beginning of *Another one bites the dust* by

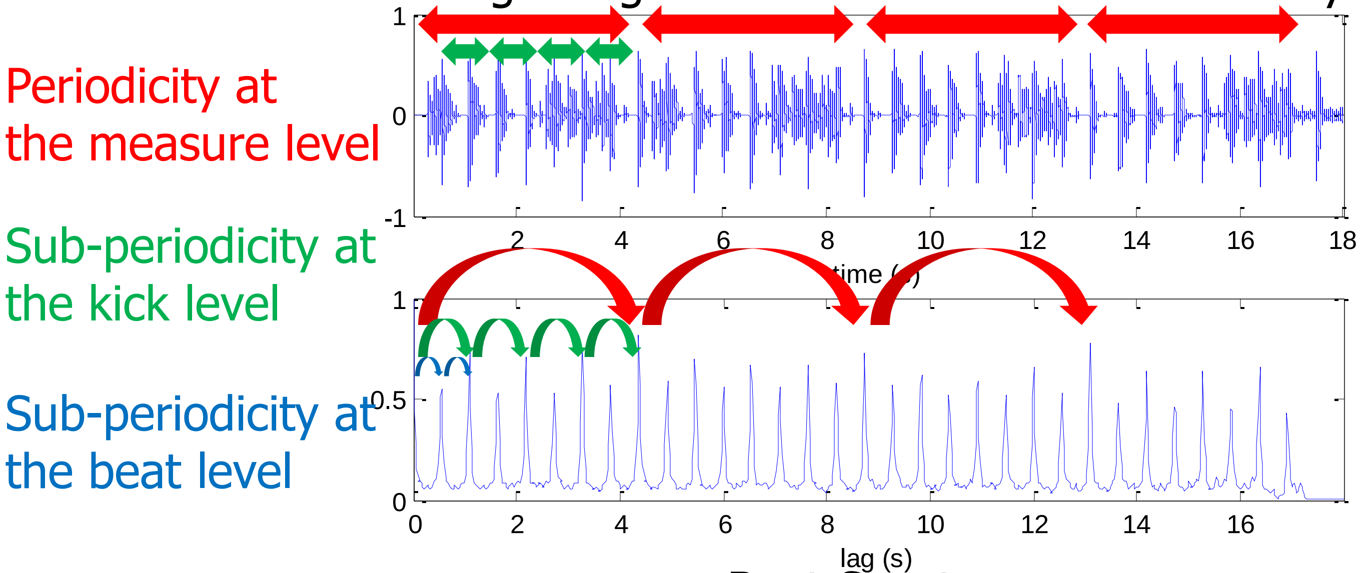


Beat Spectrum.

Beat Spectrum

- Use
  - The beat spectrum reveals the hierarchically periodically repeating structure of the audio

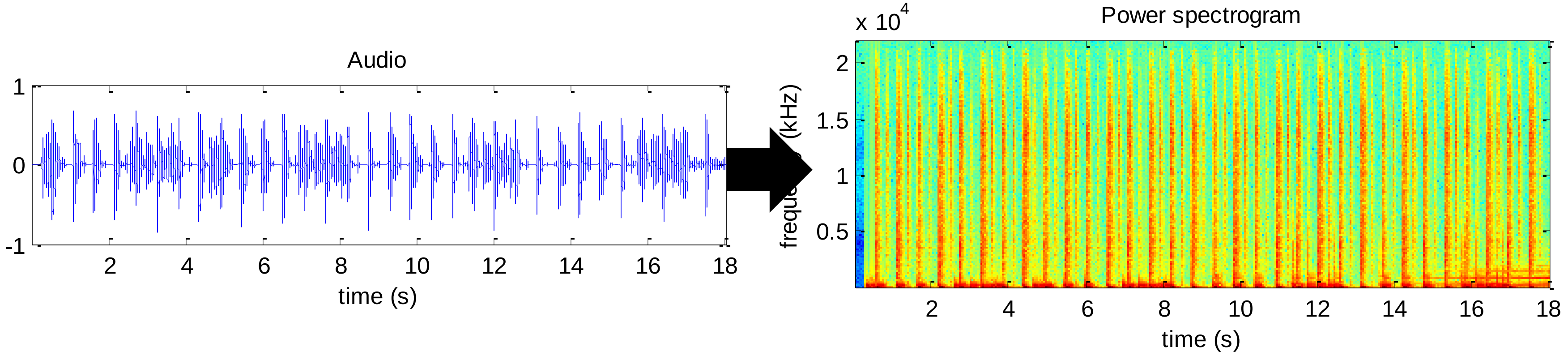
Beginning of *Another one bites the dust* by



Beat Spectrum.

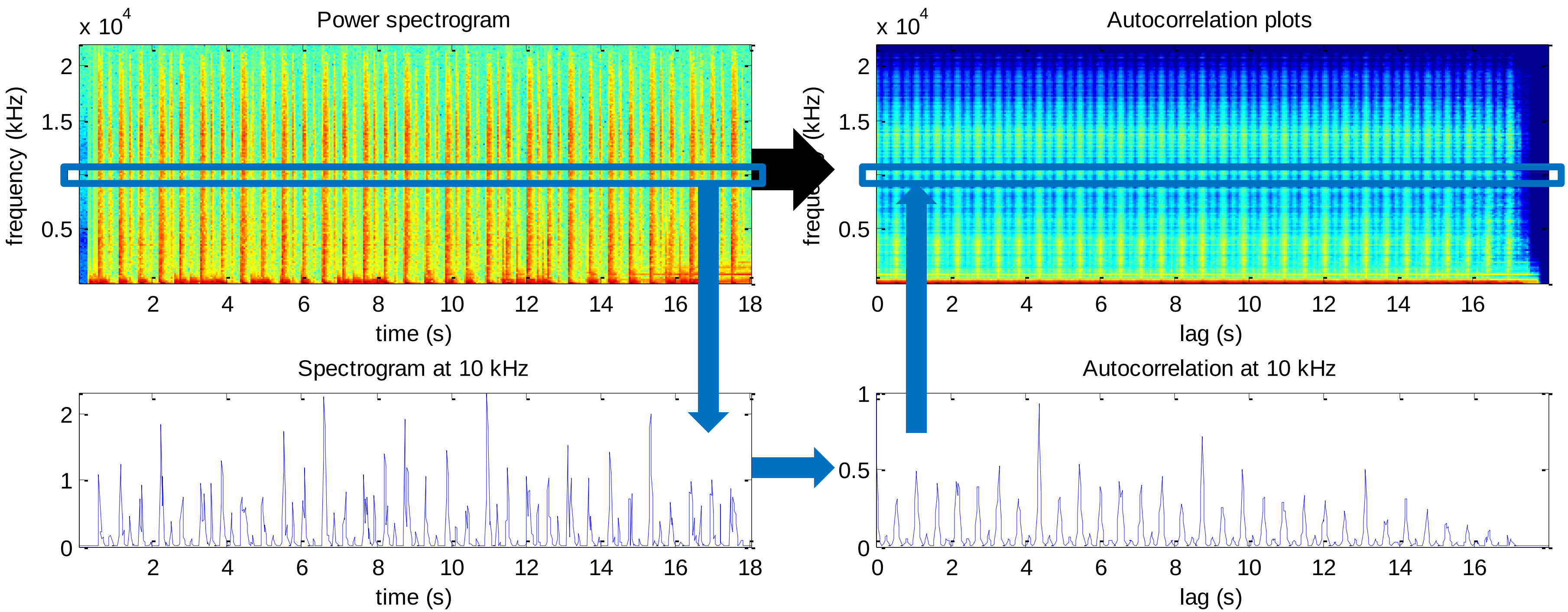
Beat Spectrum

- Calculation
  - Compute the power spectrogram from the audio using the STFT (square of magnitude spectrogram)



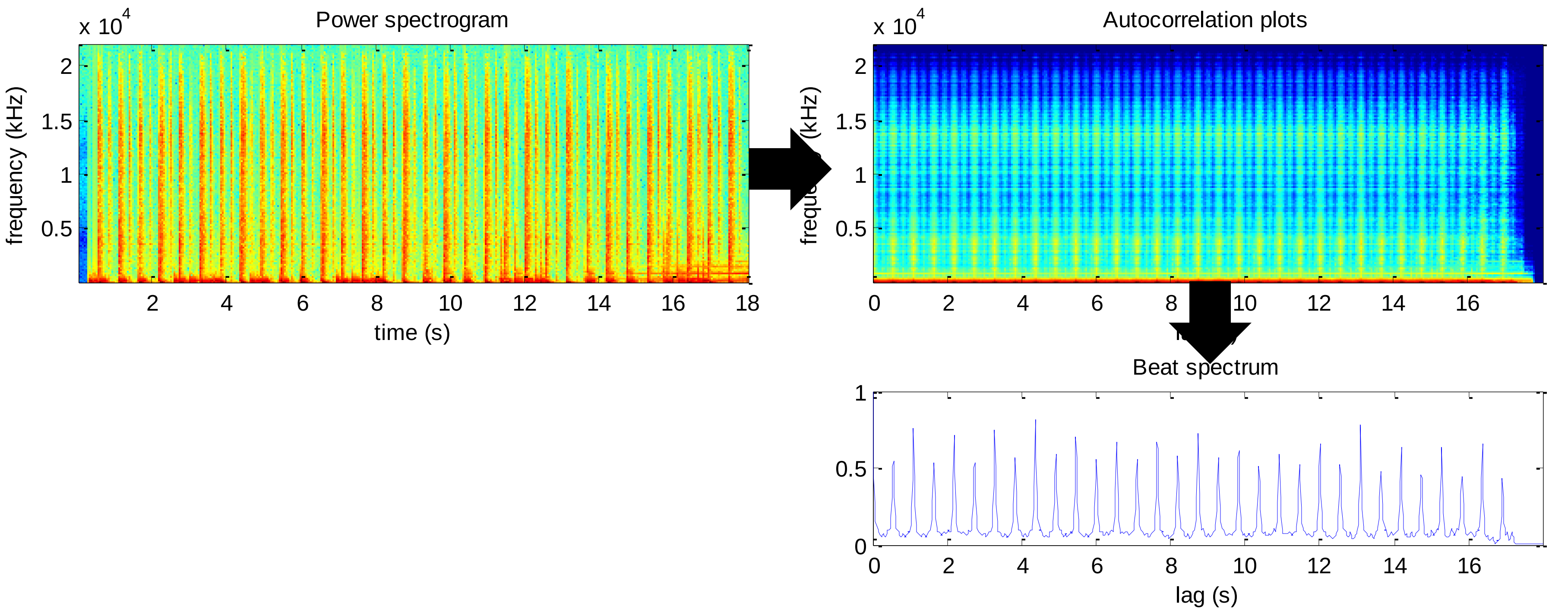
Beat Spectrum

- Calculation
  - Compute the autocorrelation of the rows of the spectrogram



Beat Spectrum

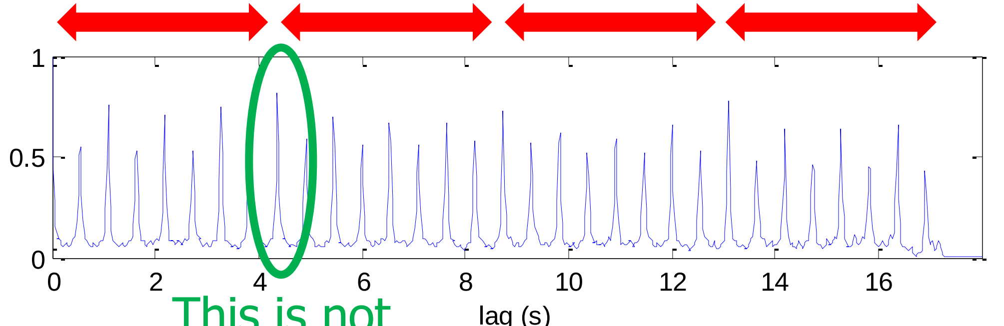
- Calculation
  - Compute the mean of the autocorrelations (of the rows)



Beat Spectrum

- Notes
  - The first highest peak in the beat spectrum does not always correspond to the repeating period!
  - The beat spectrum does not indicate where the beats are or when a measure starts!

This is how you find the period…



This is not the period

Beat Spectrum.

State-of-the-Art

- Some interesting links
  - Dannenberg's articles on beat tracking: http://www.cs.cmu.edu/~rbd/bib-beattrack.html
  - Goto's work on beat tracking: http://staff.aist.go.jp/m.goto/PROJ/bts.html
  - Ellis' Matlab codes for tempo estimation and beat tracking: http://labrosa.ee.columbia.edu/projects/beattrack/
  - MIREX's annual evaluation campaign for Music Information Retrieval (MIR) algorithms, including tasks such as onset detection, tempo extraction, and beat tracking: http://www.music-ir.org/mirex/wiki/MIREX_HOME

References

- J. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M. B. Sandler, "A tutorial on onset detection in music signals," *IEEE Trans. on Speech and Audio Processing*, vol. 13, no. 5, 2005.
- R. B. Dannenberg, "Music Understanding by Computer," *1987/1988 Computer Science Research Review*, Carnegie Mellon School of Computer Science, pp. 19-28, 1987.
- J. Foote, "Visualizing Music and Audio using Self-Similarity," in *7th ACM International Conference on Multimedia (Part 1)*, Orlando, FL, USA, pp. 77-80, October 30-November 05, 1999.
- J. Foote, "Automatic Audio Segmentation using a Measure of Audio Novelty," in *IEEE International Conference on Multimedia and Expo*, New York, NY, USA, vol.1, pp. 452-455, July 30-August 02, 2000.
- J. Foote and S. Uchihashi, "The Beat Spectrum: A New Approach to Rhythm Analysis," in *IEEE International Conference on Multimedia and Expo*, Tokyo, Japan, pp. 881-884, August 22-25, 2001.
- M. Goto, "An Audio-based Real-time Beat Tracking System for Music With or Without Drum-sounds," *Journal of New Music Research*, vol. 30, no. 2, pp. 159-171, 2001.
- D. P. W. Ellis, "Beat Tracking by Dynamic Programming," *Journal of New Music Research*, vol. 36, no. 1, pp. 51-60, 2007.
- M. Müller, D. P. W. Ellis, A. Klapuri, and G. Richard, "Signal Processing for Music Analysis," *IEEE Journal of Selected Topics in Signal Processing*, vol. 5, no. 6, pp. 1088-1110, October 2011.
- Wikipedia, "Rhythm," http://en.wikipedia.org/wiki/Rhythm, 2012.
- Wikipedia, "Meter," http://en.wikipedia.org/wiki/Metre_(music), 2012.