

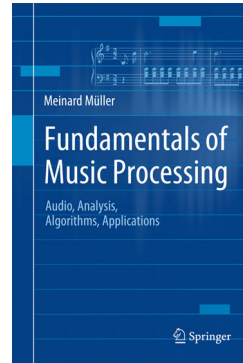
Workshop HfM Karlsruhe  
Music Information Retrieval

## Music Representations

Christof Weiß, Frank Zalkow, Meinard Müller  
International Audio Laboratories Erlangen

christof.weiss@audiolabs-erlangen.de  
frank.zalkow@audiolabs-erlangen.de  
meinard.mueller@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

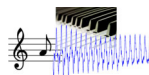
Chapter	Music Processing Scenario
1	Music Representations
2	Fourier Analysis of Signals
3	Music Synchronization
4	Music Structure Analysis
5	Chord Recognition
6	Tempo and Beat Tracking
7	Content-Based Audio Retrieval
8	Musically Informed Audio Decomposition

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)

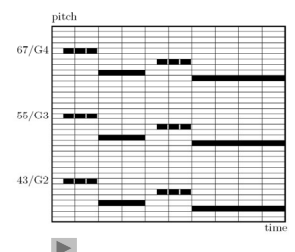
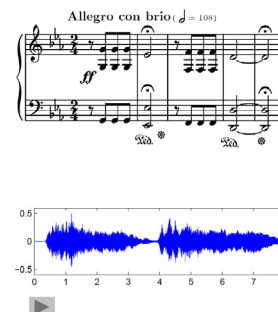
## Chapter 1: Music Representations

- 1.1 Sheet Music Representations
- 1.2 Symbolic Representations
- 1.3 Audio Representation
- 1.4 Further Notes



Musical information can be represented in many different ways. In Chapter 1, we consider three widely used music representations: sheet music, symbolic, and audio representations. This first chapter also introduces basic terminology that is used throughout the book. In particular, we discuss musical and acoustic properties of audio signals including aspects such as frequency, pitch, dynamics, and timbre.

## Music Representations



## Music Representations

- Sheet music representation
  - visual description of a musical score
  - image format (printed or scanned)
- Symbolic representations
  - description based on entities with explicit musical meaning
  - given in digital format that can be parsed by a computer
- Audio representation
  - physical description
  - encoding of sound wave

## Sheet Music Representation

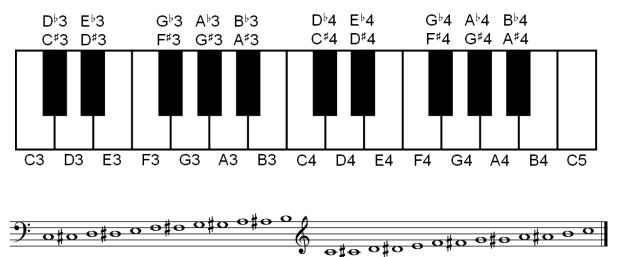
- Graphical-textual encoding of musical parameters
  - notes (onsets, pitches, durations)
  - tempo, measure, dynamics
  - instrumentation
  - ...
- Guide for performing music
- Leaves freedom for various interpretations

## Sheet Music Representation



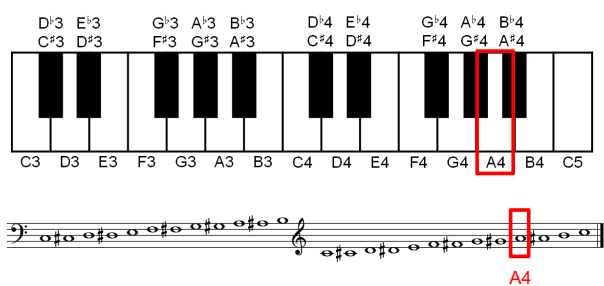
## Sheet Music Representation

### Piano keyboard and notes



## Sheet Music Representation

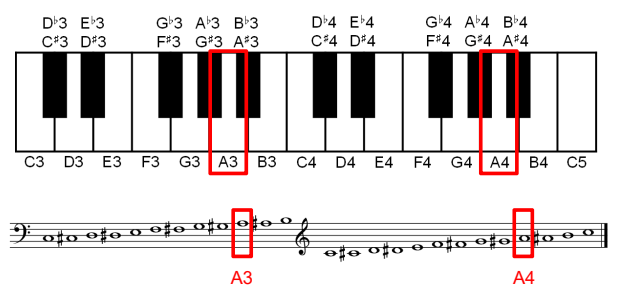
### Piano keyboard and notes



A ≙ pitch spelling attribute  
4 ≙ octave number

## Sheet Music Representation

### Piano keyboard and notes

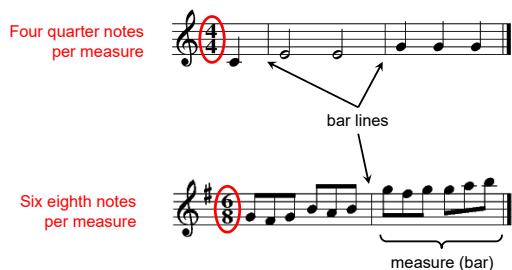


A ≙ pitch spelling attribute  
4 ≙ octave number



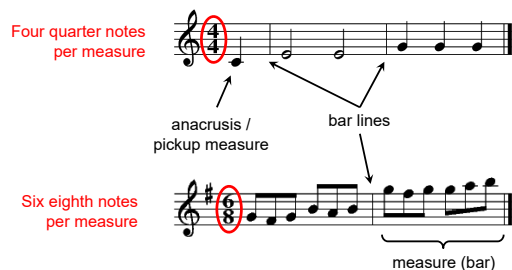
## Sheet Music Representation

### Time signature



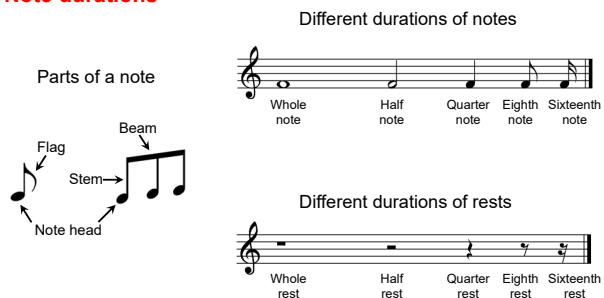
## Sheet Music Representation

### Time signature



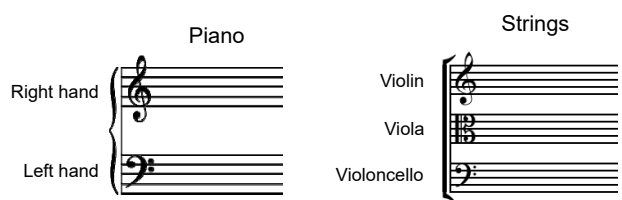
## Sheet Music Representation

### Note durations



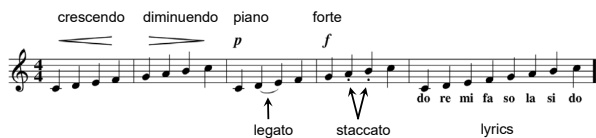
## Sheet Music Representation

### Staff systems



## Sheet Music Representation

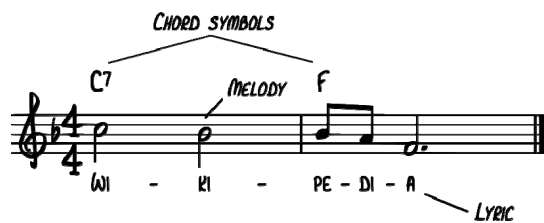
### Dynamics and articulation



## Sheet Music Representation



## Sheet Music Representation



## Sheet Music Representation



## Sheet Music Representation

### Types of score

- Full score: shows music for all instruments and voices; used by conductors
- Piano (reduction) score: transcription for piano  
Example: Liszt transcription of Beethoven symphonies
- Short score: reduction of a work for many instruments to just a few staves
- Lead sheet: specifies only melody, lyrics and harmonies (chord symbols); used for popular music to capture essential elements of a song

## Symbolic Representation

- Symbolic description of music
    - based on entities that have an explicit musical meaning
    - given in some digital format
    - can be parsed by a computer
  - Note:
    - Scanned sheet music based on pixels
    - Digital audio file based on samples
- are **not** regarded as being symbolic music formats

## Symbolic Representation

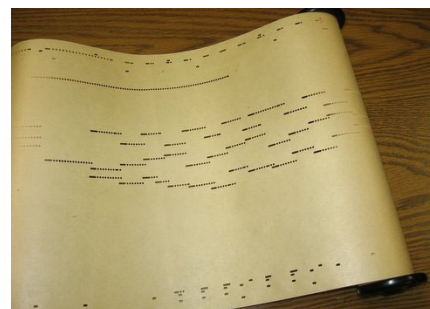
### MusicXML

```
<note>
  <pitch>
    <step>E</step>
    <alter>-1</alter>
    <octave>4</octave>
  </pitch>
  <duration>2</duration>
  <type>half</type>
</note>
```



## Symbolic Representation

### Piano roll representation



## Symbolic Representation

### Piano roll representation



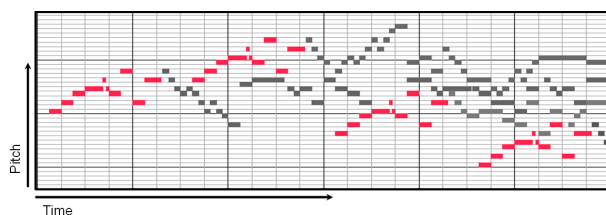
## Symbolic Representation

### Piano roll representation

- Piano roll: music storage medium used to operate a player piano
- Perforated paper rolls
- Holes in the paper encode the note parameters onset, duration, and pitch
- First pianola: 1895

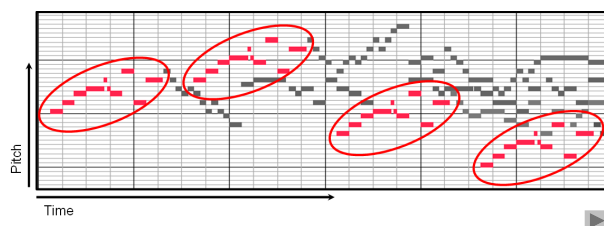
## Symbolic Representation

### Piano roll representation



## Symbolic Representation

### Piano roll representation



## Symbolic Representation

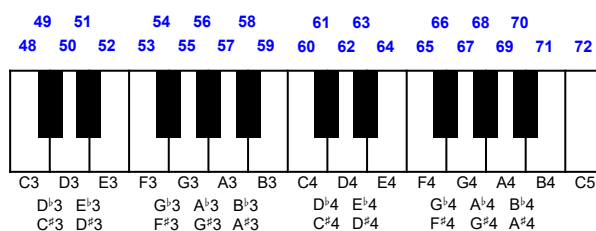
### MIDI representation

- Musical Instrument Digital Interface (MIDI)
- Standard protocol for controlling and synchronizing digital instruments
- Standard MIDI File (SMF) is used for collecting and storing MIDI messages
- SMF file is often called MIDI file

## Symbolic Representation

### MIDI representation

MIDI note numbers (MNN)  $\triangleq$  piano keys



## Symbolic Representation

### MIDI representation

- MIDI note number (pitch)
  - $p = 21, \dots, 108 \triangleq$  piano keys
  - $p = 69 \triangleq$  concert pitch A4
- Key velocity  $\triangleq$  intensity
- MIDI channel  $\triangleq$  instrument
- Note-on / note-off events  $\triangleq$  onset time & duration
- Tempo measured in clock pulses or ticks
  - (each MIDI event has a timestamp)
- Absolute tempo specified by
  - ticks per quarter note (musical time)
  - micro-seconds per tick (physical time)

## Symbolic Representation

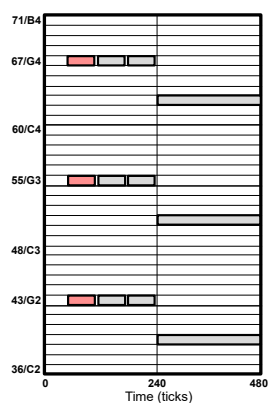
### MIDI representation



Time (Ticks)	Message	Channel	Note Number	Velocity
60	NOTE ON	1	67	100
0	NOTE ON	1	55	100
0	NOTE ON	2	43	100
55	NOTE OFF	1	67	0
0	NOTE OFF	1	55	0
0	NOTE OFF	2	43	0
5	NOTE ON	1	67	100
0	NOTE ON	1	55	100
0	NOTE ON	2	43	100
55	NOTE OFF	1	67	0
0	NOTE OFF	1	55	0
0	NOTE OFF	2	43	0
5	NOTE ON	1	67	100
0	NOTE ON	1	55	100
0	NOTE ON	2	43	100
55	NOTE OFF	1	67	0
0	NOTE OFF	1	55	0
0	NOTE OFF	2	43	0
5	NOTE ON	1	63	100
0	NOTE ON	2	51	100
0	NOTE ON	2	39	100
240	NOTE OFF	1	63	0
0	NOTE OFF	2	51	0
0	NOTE OFF	2	39	0

## Symbolic Representation

### MIDI representation



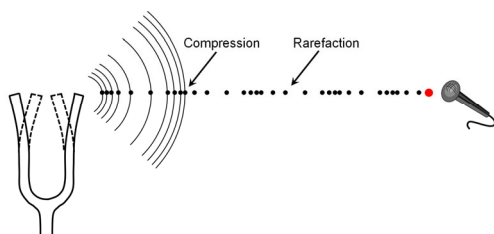
## Audio Representation

Various interpretations – Beethoven's Fifth

Bernstein	<a href="#">▶</a>
Karajan	<a href="#">▶</a>
Scherbakov (piano)	<a href="#">▶</a>
MIDI (piano)	<a href="#">▶</a>

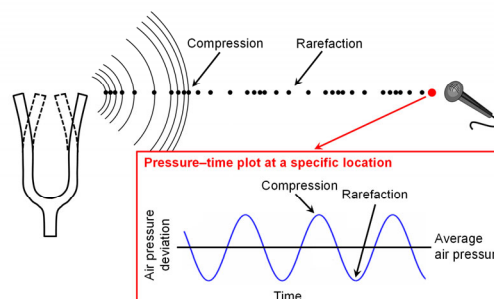
## Audio Representation

### Waveform



## Audio Representation

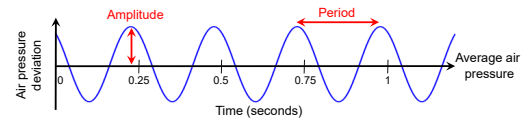
### Waveform



## Waveform

- Audio signal encodes change of air pressure at a certain location generated by a vibrating object (e.g. string, vocal cords, membrane)
- Waveform (pressure-time plot) is graphical representation of audio signal
- Parameters: amplitude, frequency / period

## Waveform



## Waveform

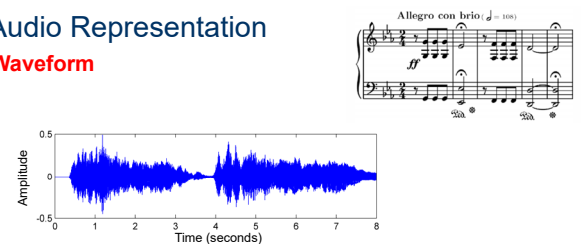
Pure tone (harmonic sound):

- Sinusoidal waveform
- Prototype of an acoustic realization of a musical note

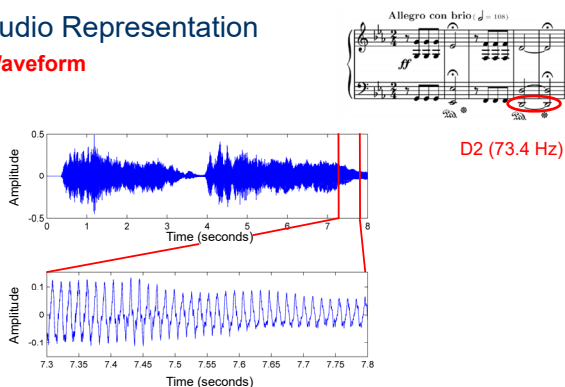
Parameters:

- Period  $p$  : time between to successive high pressure points
- Frequency  $f = \frac{1}{p}$  (measured in Hz)
- Amplitude  $a$  : air pressure at high pressure points

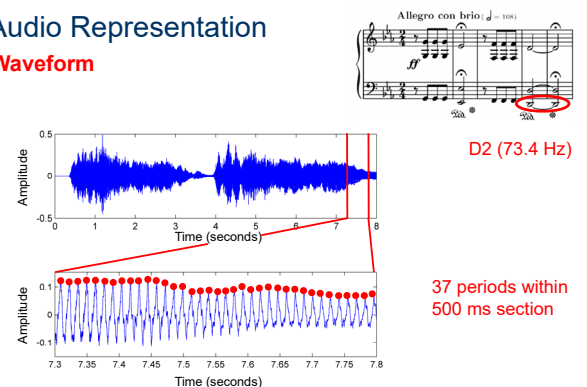
## Waveform



## Waveform



## Waveform





## Audio Representation

### Sound

- Sound: superposition of sinusoidals
- When realizing musical notes on an instrument one obtains a complex superposition of pure tones (and other noise-like components)
- Harmonics: integer multiples of fundamental frequency
  1. Harmonic  $\triangleq$  fundamental frequency (e.g. 440 Hz)
  2. Harmonic  $\triangleq$  first overtone (e.g. 880 Hz)
  3. Harmonic  $\triangleq$  second overtone (e.g. 1320 Hz)

## Audio Representation

### Pitch

- Property that correlates to the perceived frequency ( $\triangleq$  fundamental frequency)
- Example: A4 (also called concert pitch)  $\triangleq$  440 Hz
- Slight changes in frequency have no effect on perceived pitch (pitch  $\triangleq$  entire range of frequencies)
- Pitch perception: logarithmic in frequency  
Example: octave  $\triangleq$  doubling of frequency

## Audio Representation

### Pitch

**Equal-tempered scale:** A system of tuning in which every pair of adjacent notes has an identical frequency ratio

**Western music: 12-tone equal-tempered scale**

- Each octave is divided up into 12 logarithmically equal parts
- Notes correspond to piano keys:  $p = 21$  (A0) to  $p = 108$  (C8)
- Reference or standard pitch:  $p = 69$  (A4)  $\triangleq$  440 Hz
- Center frequency of a note with MIDI pitch  $p$

$$F_{\text{pitch}}(p) = 2^{(p-69)/12} \cdot 440 \quad (\text{Hz})$$

## Audio Representation

### Pitch

- **Semitone:** difference between two subsequent scale steps
- Ratio of frequencies one semitone apart is constant:

$$F_{\text{pitch}}(p+1)/F_{\text{pitch}}(p) = 2^{1/12} \approx 1.059463$$

- **Cent:** 1200 cents per octave (by definition)  
100 cents per semitone (equivalent definition)
- Ratio of frequencies one cent apart is constant:

$$2^{1/1200} \approx 1.0005777895$$

## Audio Representation

### Pitch

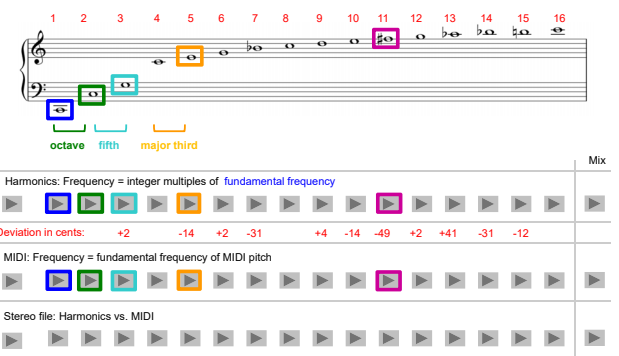
- Difference in cents between two frequencies  $\omega_1$  and  $\omega_2$ :

$$\log_2 \left( \frac{\omega_1}{\omega_2} \right) \cdot 1200$$

- Just noticeable difference = threshold of what is perceptible
  - varies from person to person
  - depends on other aspects such as the timbre
  - 25 cents recognizable by most people
  - 10 cents recognizable only by trained listeners

## Audio Representation

### Harmonics



Audio Representation

Dynamics

- Intensity of a sound
- Energy of the sound per time and area
- Loudness: subjective (psychoacoustic) perception of intensity (depends on frequency, timbre, duration)

Audio Representation

Dynamics

- $$\text{intensity} = \frac{\text{energy}}{\text{time} \cdot \text{area}} = \frac{\text{power}}{\text{area}} \quad \left( \frac{\text{W}}{\text{m}^2} \right)$$
- Decibel (dB): logarithmic unit to measure intensity relative to a reference level
- Reference level: threshold of hearing (THO)  $I_{\text{TOH}} := 10^{-12} \text{ W/m}^2$
- Intensity  $I$  measured in dB:  $\text{dB}(I) := 10 \cdot \log_{10} \left( \frac{I}{I_{\text{TOH}}} \right)$
- Examples:  
 $I = 10 \cdot I_{\text{TOH}} \rightarrow I$  has a sound level of 10 dB  
 $I = 100 \cdot I_{\text{TOH}} \rightarrow I$  has a sound level of 20 dB

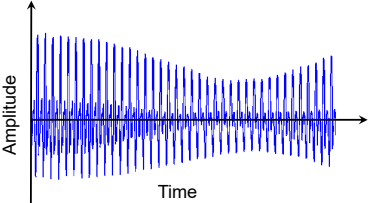
Audio Representation

Dynamics

Source	Intensity	Intensity level	× TOH
Threshold of hearing (TOH)	$10^{-12}$	0 dB	1
Whisper	$10^{-10}$	20 dB	$10^2$
Pianissimo	$10^{-8}$	40 dB	$10^4$
Normal conversation	$10^{-6}$	60 dB	$10^6$
Fortissimo	$10^{-2}$	100 dB	$10^{10}$
Threshold of pain	10	130 dB	$10^{13}$
Jet take-off	$10^2$	140 dB	$10^{14}$
Instant perforation of eardrum	$10^4$	160 dB	$10^{16}$

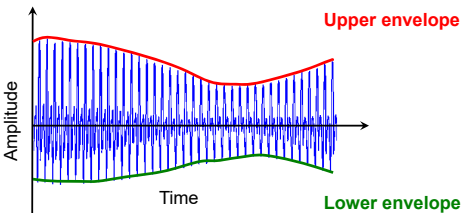
Audio Representation

Dynamics



Audio Representation

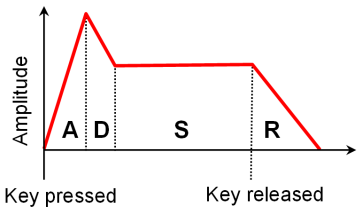
Dynamics



Audio Representation

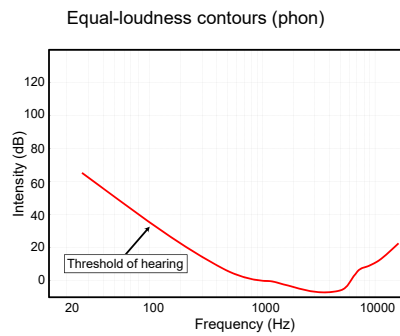
Dynamics

ADSR model: attack (A), decay (D), sustain (S), and release (R) phase



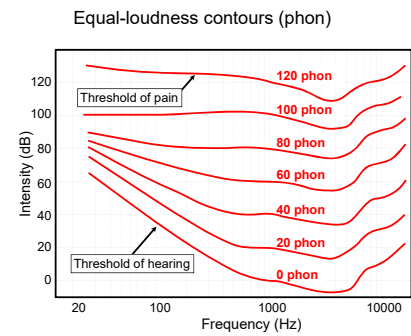
## Audio Representation

### Loudness



## Audio Representation

### Loudness



## Audio Representation

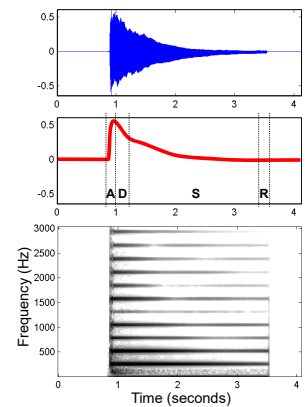
### Timbre

- Quality of musical sound that distinguishes different types of sound production such as voices or instruments
- Tone quality
- Tone color
- Depends on energy distribution in harmonics

## Audio Representation

### Timbre

Piano playing  
note C4 (261.6 Hz)



## Audio Representation

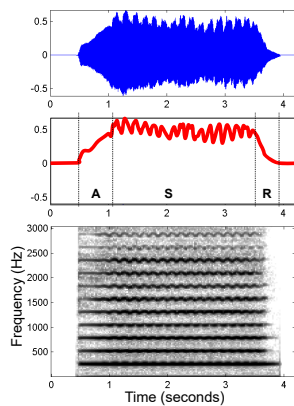
### Timbre

Violine playing  
note C4 (261.6 Hz)



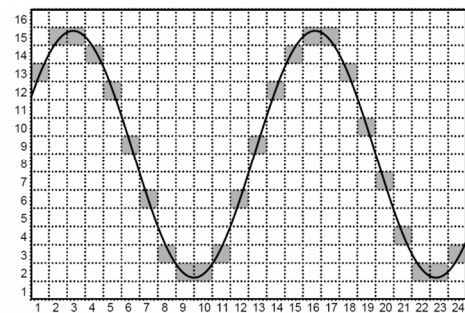
Vibrato:  
Frequency modulations

Tremolo:  
Amplitude modulations



## Audio Representation

### Digitization



## Audio Representation

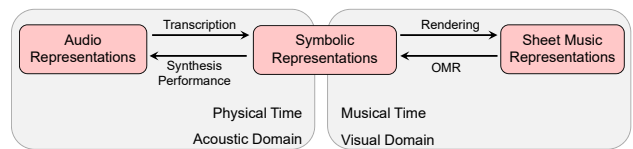
### Digitization

- Conversion of continuous-time (analog) signal into a discrete signal
- Sampling (discretization of time axis)
- Quantization (discretization of amplitudes)

Examples:

- Audio CD: 44100 Hz sampling rate  
16 bits (65536 values) used for quantization
- Telephone: 8000 Hz sampling rate  
8 bits (256 values) used for quantization

## Music Representations



OMR = optical music recognition

Process of transforming sheet music into a symbolic representation

## Music Representations

### OMR

Original score



OMR score



## Music Representations

### OMR

Original score



OMR score

