

Lecture

Music Processing

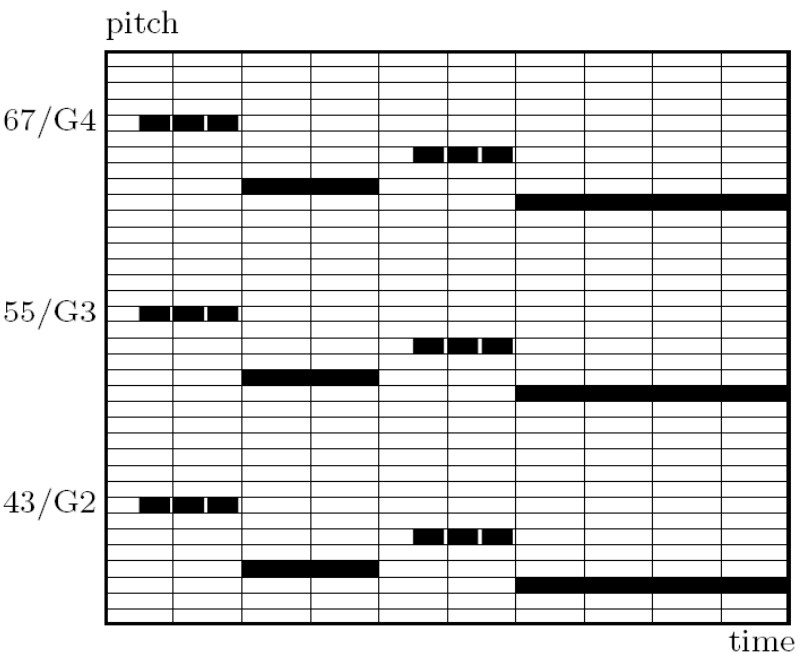
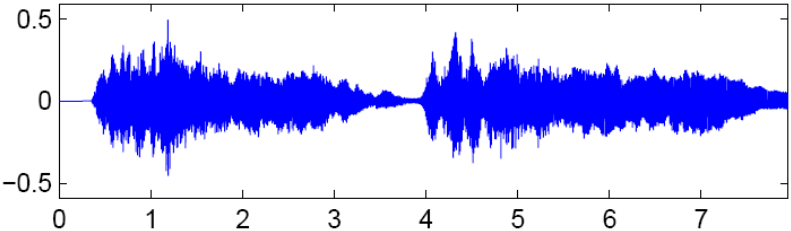
Music Representations

Meinard Müller

International Audio Laboratories Erlangen

meinard.mueller@audiolabs-erlangen.de

Music Representations



Music Representations

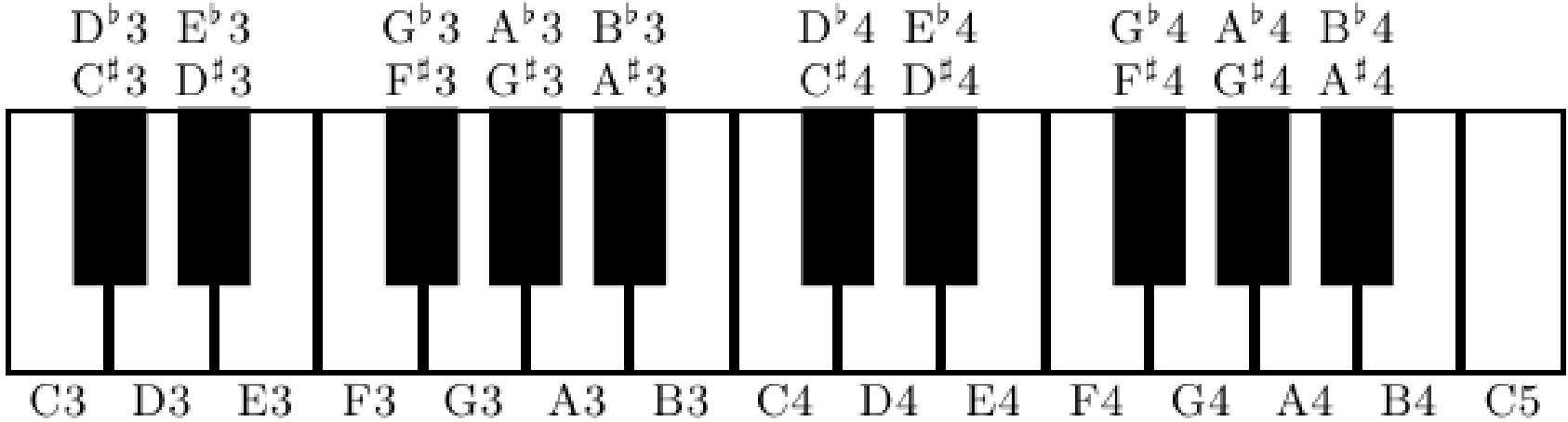
- Score representation: symbolic description
- MIDI representation: hybrid description
(models note events explicitly but may also encode performance subtleties)
- Audio representation: physical description
(encodes a sound wave)

Score Representation

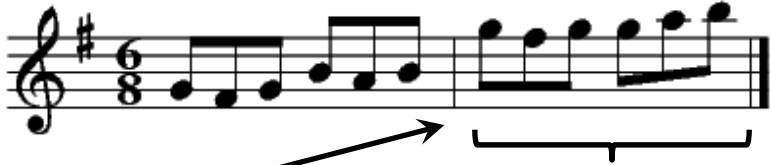
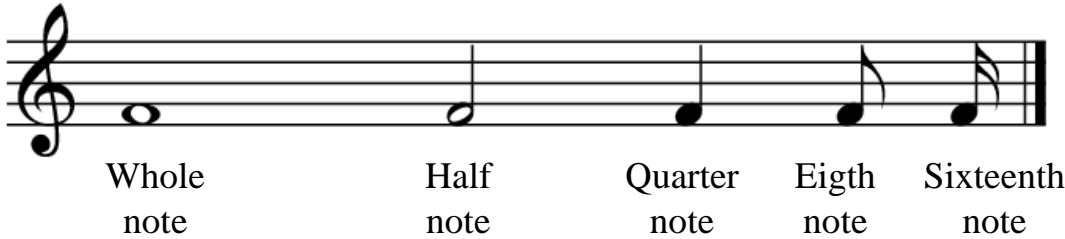
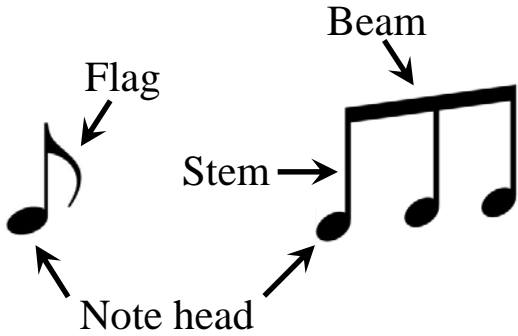
Allegro con brio (♩ = 108)

The image shows a musical score for piano in 2/4 time, featuring a forte (*ff*) dynamic and a 'Ped.' marking with a flower symbol. The score is written in a grand staff with a treble and bass clef. The key signature has two flats (B-flat and E-flat). The tempo is marked 'Allegro con brio' with a quarter note equal to 108 beats per minute. The music consists of two staves. The first staff (treble clef) starts with a forte (*ff*) dynamic and features a series of chords and a melodic line. The second staff (bass clef) features a melodic line with a 'Ped.' marking and a flower symbol. The score is divided into measures by vertical bar lines.

Score Representation



Score Representation



bar line

measure

Score Representation

Allegro con brio. $\text{♩} = 108.$

Flauti.

Oboi.

Clarineti in B.

Fagotti.

Corni in Es.

Trombe in C.

Timpani in C.G.

Allegro con brio. $\text{♩} = 108.$

Violino I.

Violino II.

Viola.

Violoncello.

Basso.

Score Representation

Symphony No. 5 C minor

LUDWIG VAN BEETHOVEN (1770-1827)
OP. 67 (1809)

Allegro con brio (♩ = 108) Piano Solo

The image displays a musical score for the Piano Solo of the first movement of Beethoven's Symphony No. 5 in C minor, Op. 67. The tempo is marked 'Allegro con brio' with a quarter note equal to 108 beats per minute. The score is written for piano and consists of two systems. The first system begins with a forte (ff) dynamic and transitions to piano (p). The second system starts at measure 13 and includes dynamics such as crescendo (cresc.), forte (f), fortissimo (ff), and piano (p). The score features complex rhythmic patterns and articulation marks like accents and slurs.

Score Representation

A hand-drawn musical score diagram illustrating the relationship between chord symbols, melody, and lyrics. The score is written on a single staff in 4/4 time, with a key signature of one flat (Bb). The lyrics are "WI - KI - PE - DI - A".

- CHORD SYMBOLS:** A bracket labeled "CHORD SYMBOLS" spans the first two measures. The first measure is labeled "C7" and the second measure is labeled "F".
- MELODY:** A line labeled "MELODY" points to the notes in the first two measures: a whole note G4 in the first measure and a whole note Bb4 in the second measure.
- LYRIC:** A line labeled "LYRIC" points to the text "WI - KI - PE - DI - A" written below the staff.

The musical notation shows a treble clef, a key signature of one flat, and a 4/4 time signature. The first measure contains a whole note G4, and the second measure contains a whole note Bb4. The third measure contains a quarter note D5, a quarter note E5, and a quarter note F5. The fourth measure contains a dotted half note G5. The lyrics "WI - KI - PE - DI - A" are aligned with the notes: "WI" under G4, "KI" under Bb4, "PE" under D5, "DI" under E5, and "A" under F5.

Score Representation



Score 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

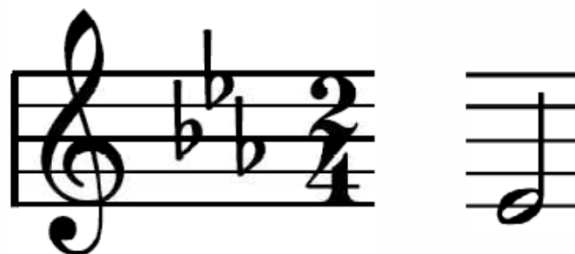
Score Representation

- Scanned image
- Various symbolic data formats
 - Lilypond
 - MusicXML
- Optical Music Recognition (OMR)
- Music notation software
 - Finale
 - Sibelius

Score Representation

MusicXML

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



Score Representation

Musical score / sheet music:

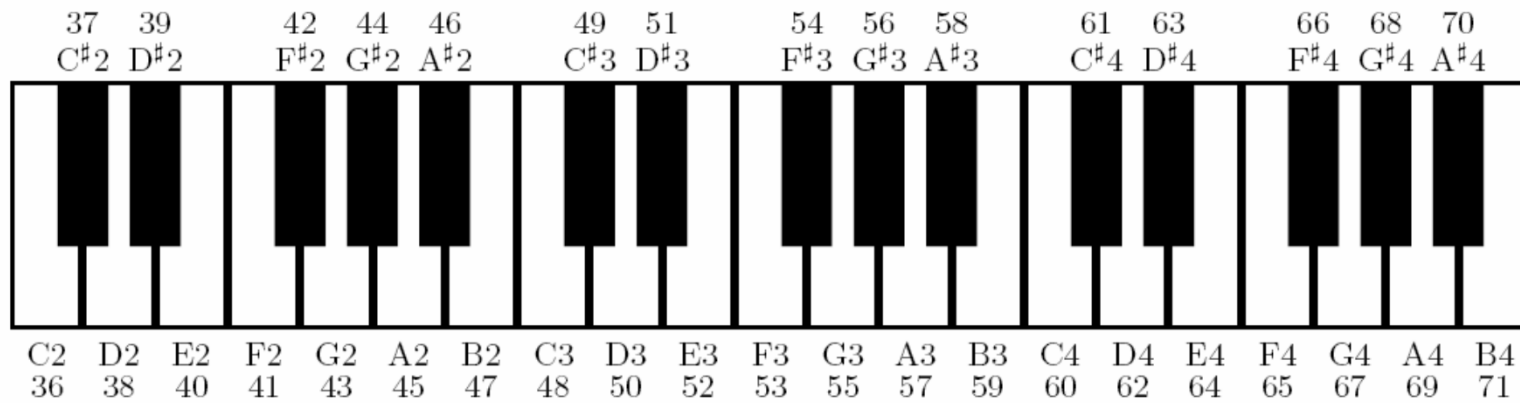
- Graphical / textual encoding of musical parameters (note onsets, pitches, durations, tempo, measure, dynamics, instrumentation)
- Guide for performing music
- Leaves freedom for various interpretations

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

MIDI Representation

MIDI note numbers (MNN) \triangleq piano keys

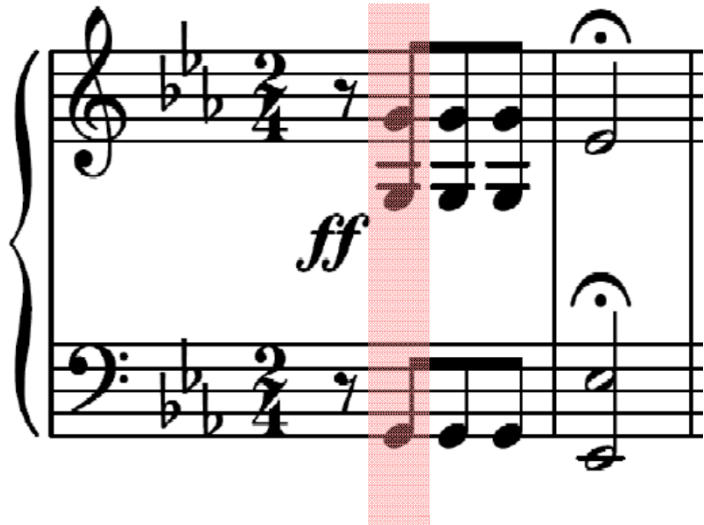


MIDI Representation

MIDI parameters:

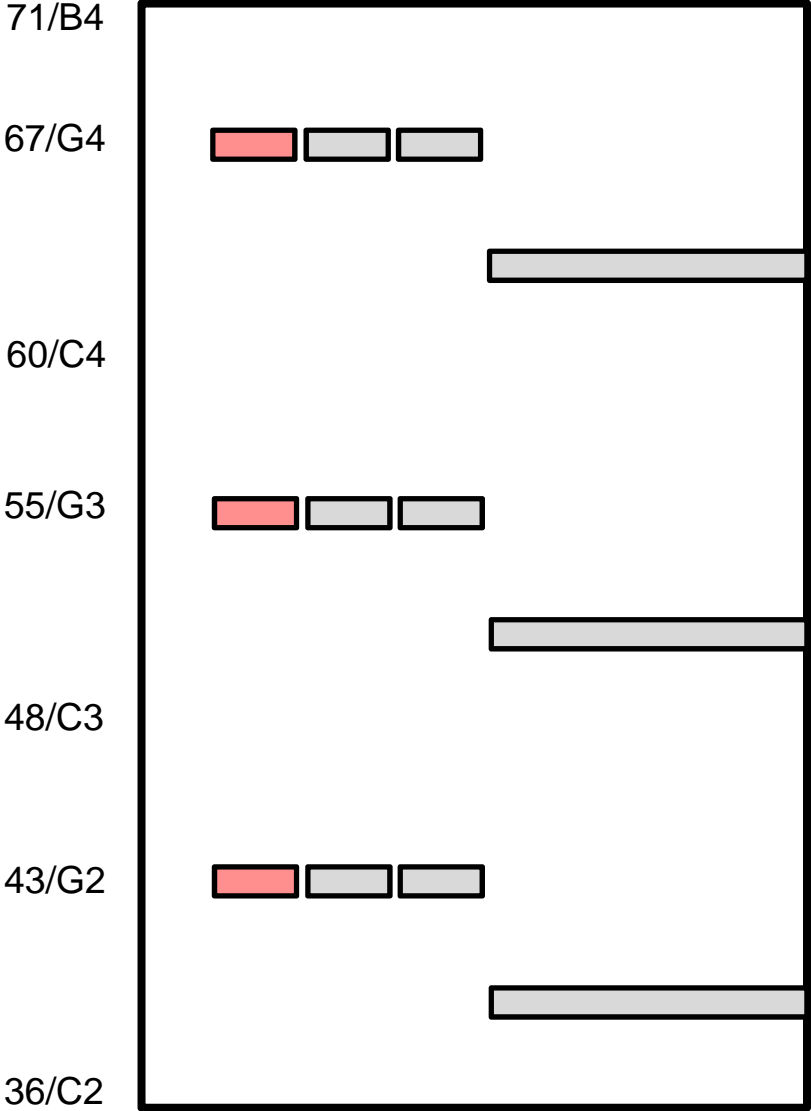
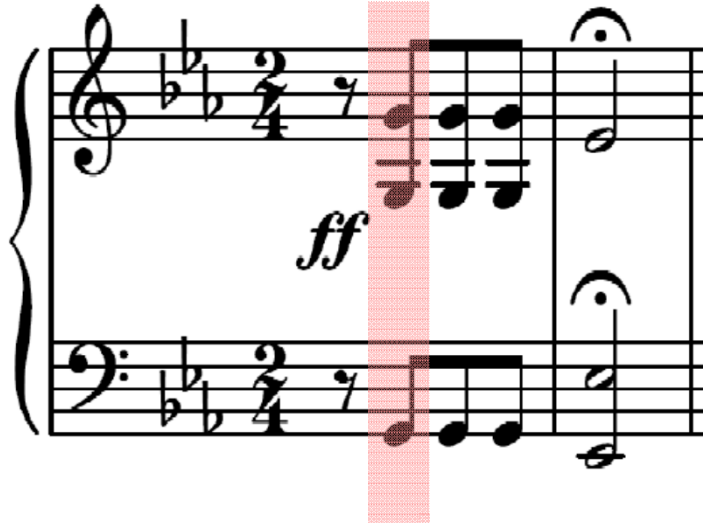
- MIDI note number (pitch) [0:127]
 $p = 21, \dots, 108 \triangleq$ „piano keys“
 $p = 69 \triangleq$ concert pitch A (440Hz)
- Key velocity [0:127] \triangleq intensity
- MIDI channel [0:15] \triangleq instrument
- Note-on / note-off events \triangleq onset time & duration
- Tempo measured in clock 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)

MIDI Representation

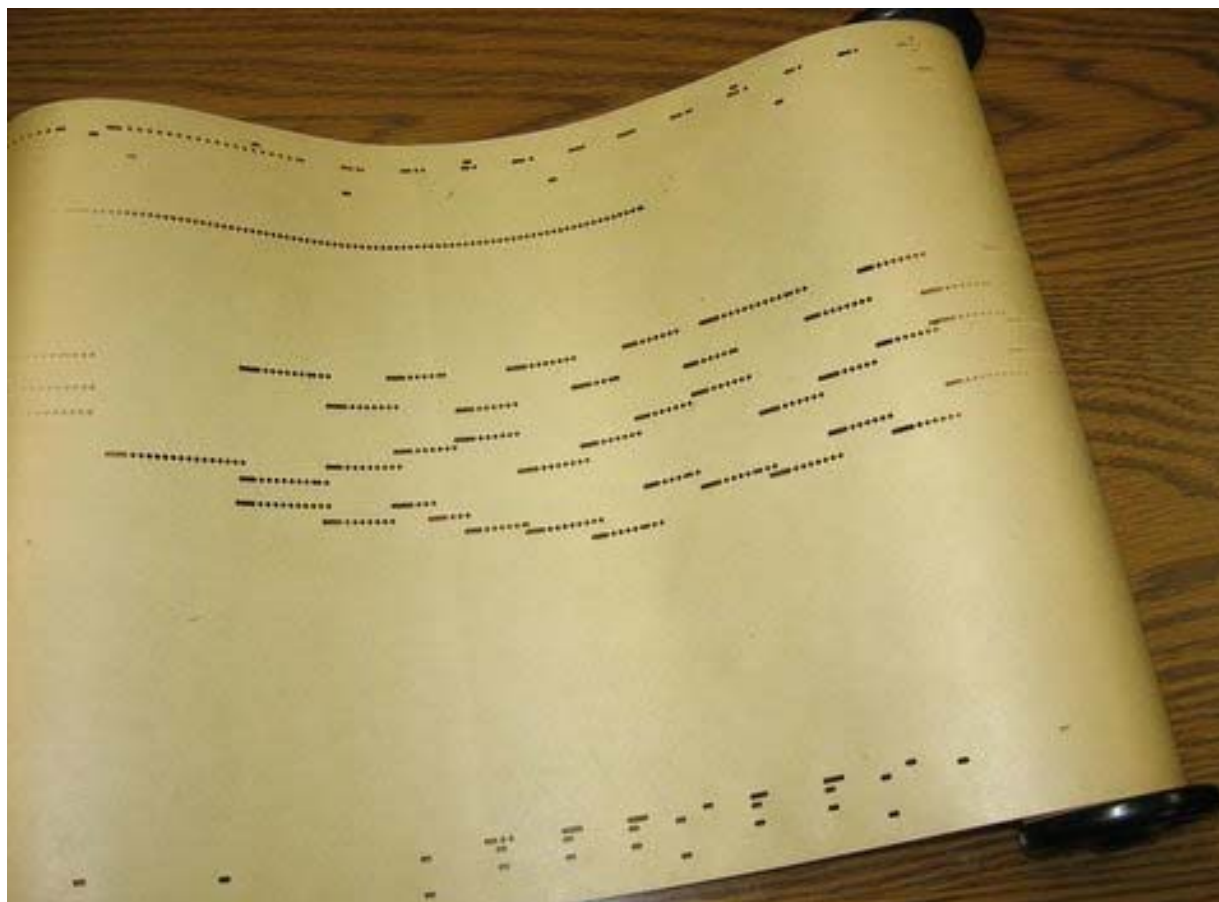


Time (Ticks)	Message	Channel	Note Number	Velocity
60	NOTE ON	1	67	100
0	NOTE ON	1	55	100
0	NOTE ON	2	43	100
55	NOTE OFF	1	67	0
0	NOTE OFF	1	55	0
0	NOTE OFF	2	43	0
5	NOTE ON	1	67	100
0	NOTE ON	1	55	100
0	NOTE ON	2	43	100
55	NOTE OFF	1	67	0
0	NOTE OFF	1	55	0
0	NOTE OFF	2	43	0
5	NOTE ON	1	67	100
0	NOTE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

MIDI Representation



MIDI Representation



MIDI Representation



MIDI 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

Audio Representation

Various interpretations – Beethoven's Fifth

Bernstein



Karajan



Scherbakov (piano)

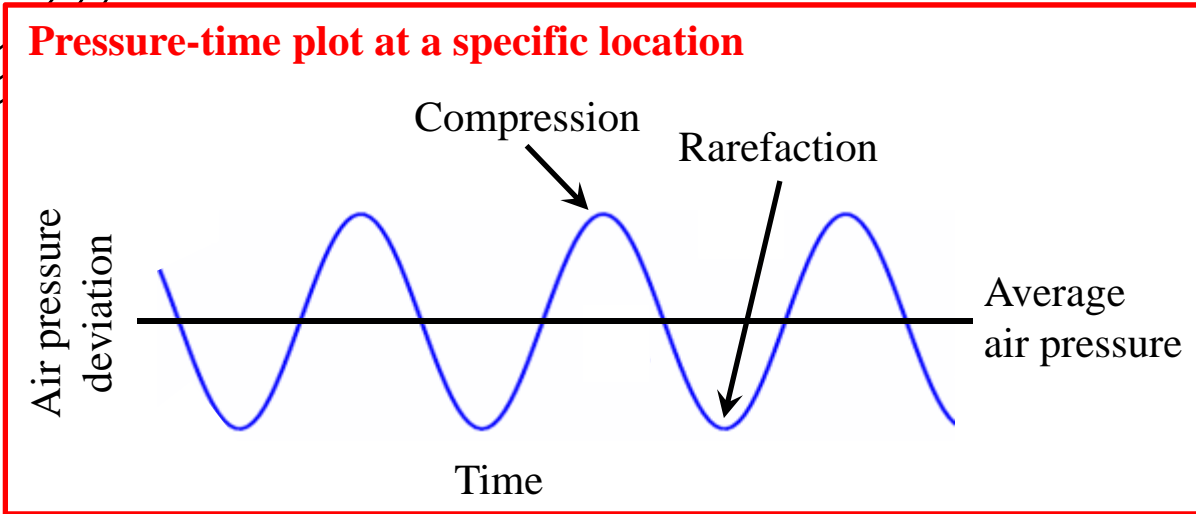
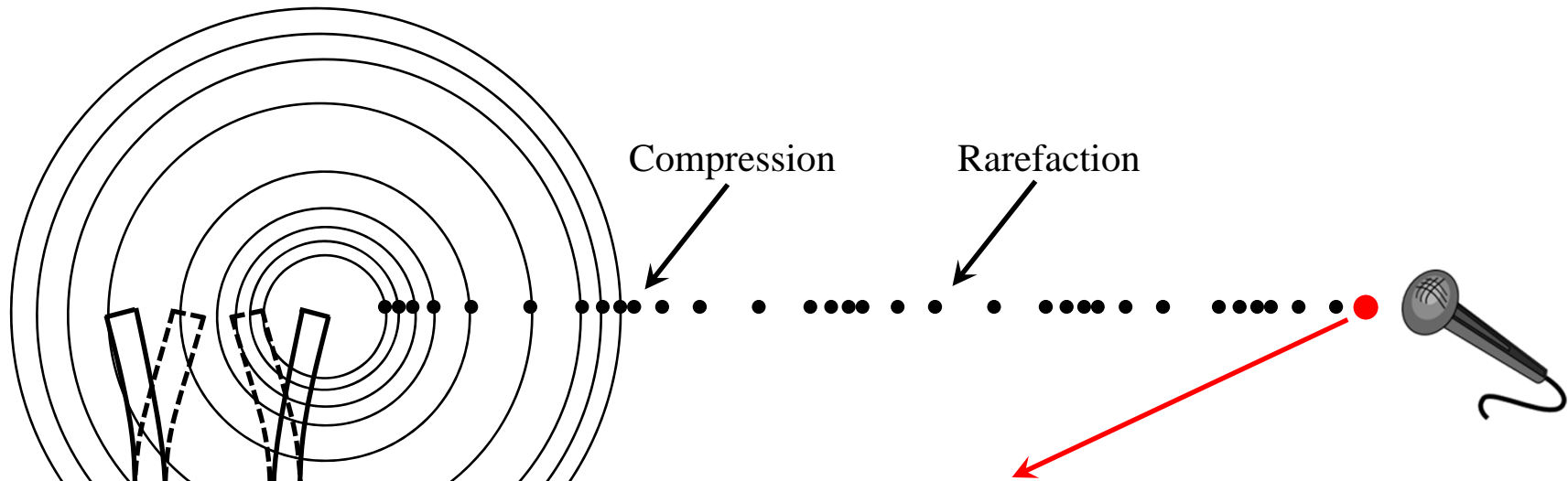


MIDI (piano)



Audio Representation

Waveform



Audio Representation

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

Audio Representation

Waveform

Pure tone (harmonic sound):

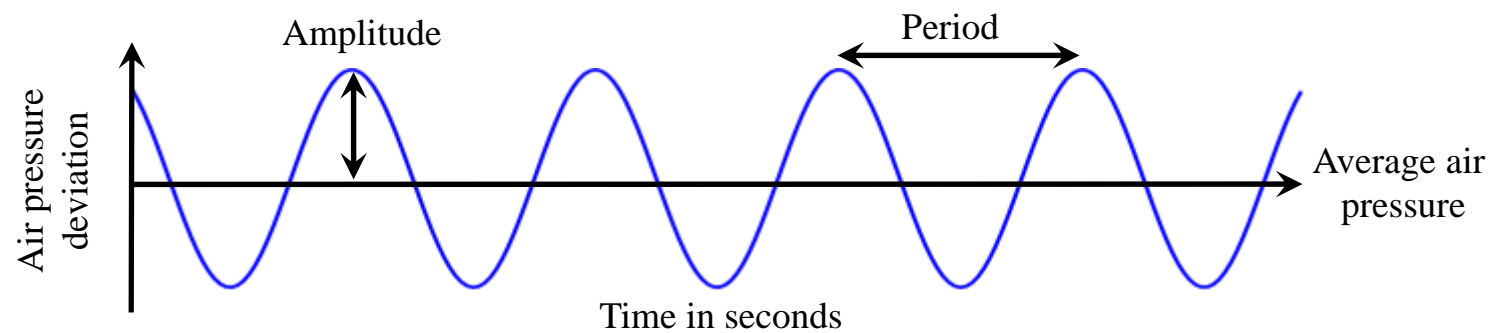
- Sinusoidal waveform
- Prototype of an acoustic realization of a musical note

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

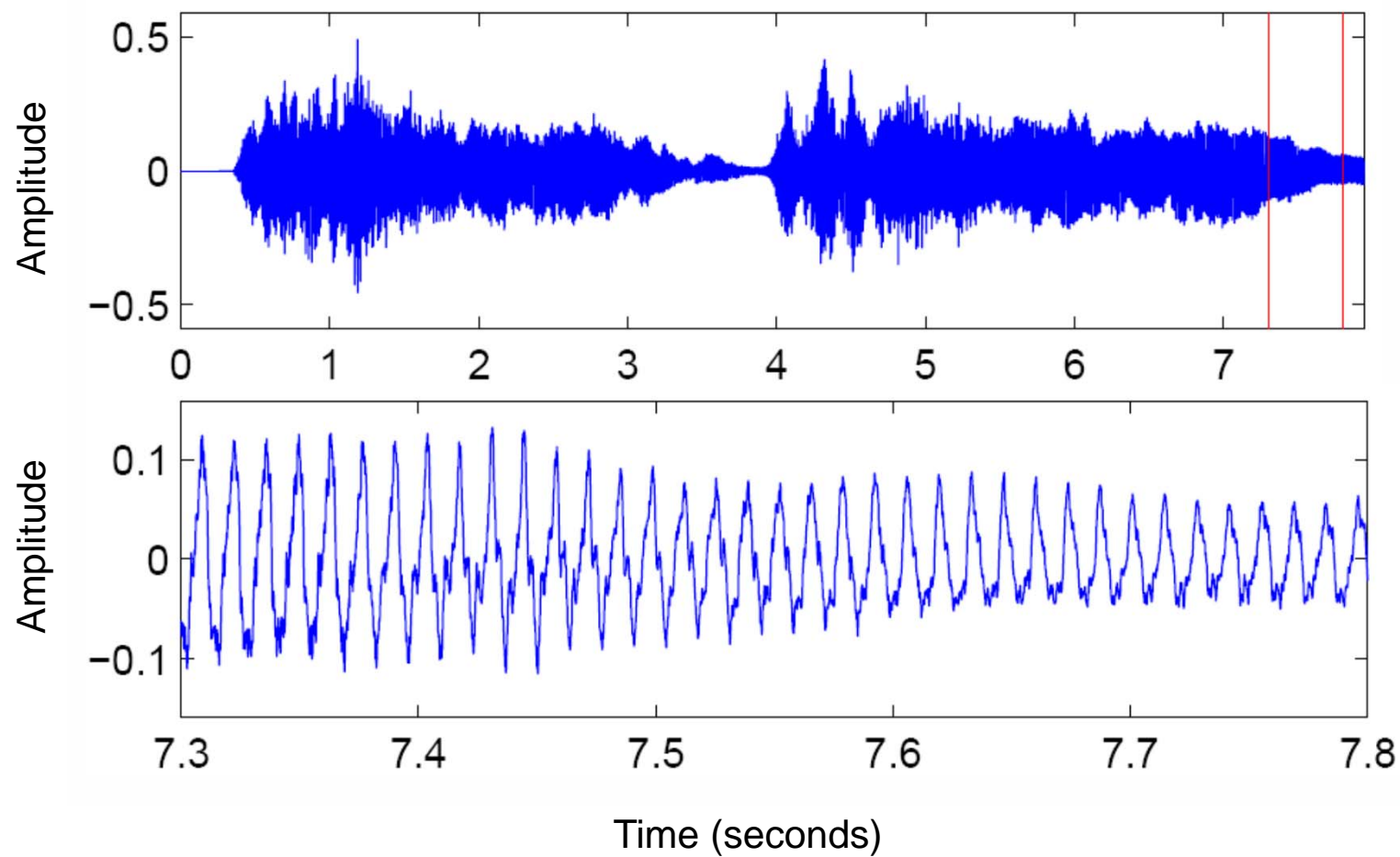
Audio Representation

Waveform



Audio Representation

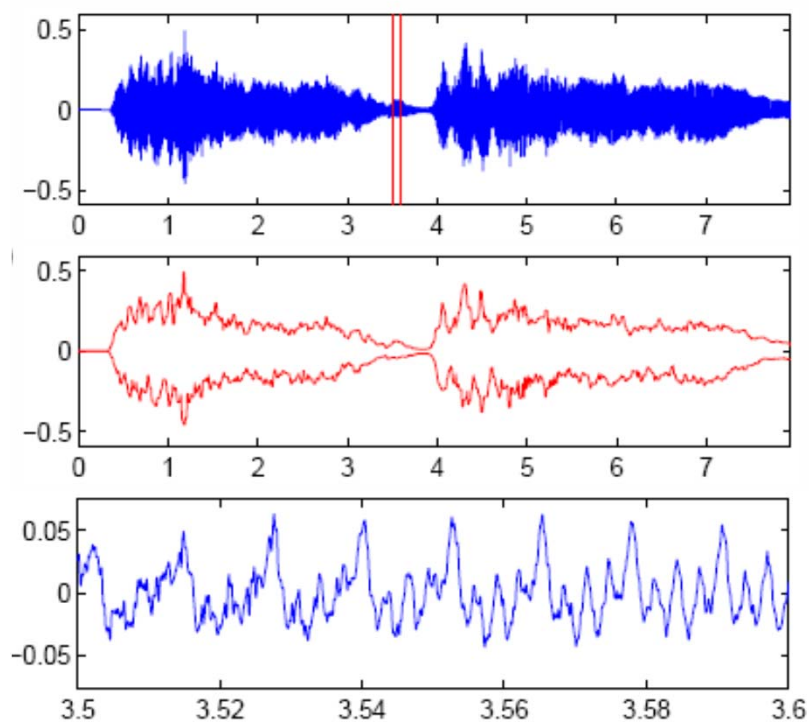
Waveform



Audio Representation

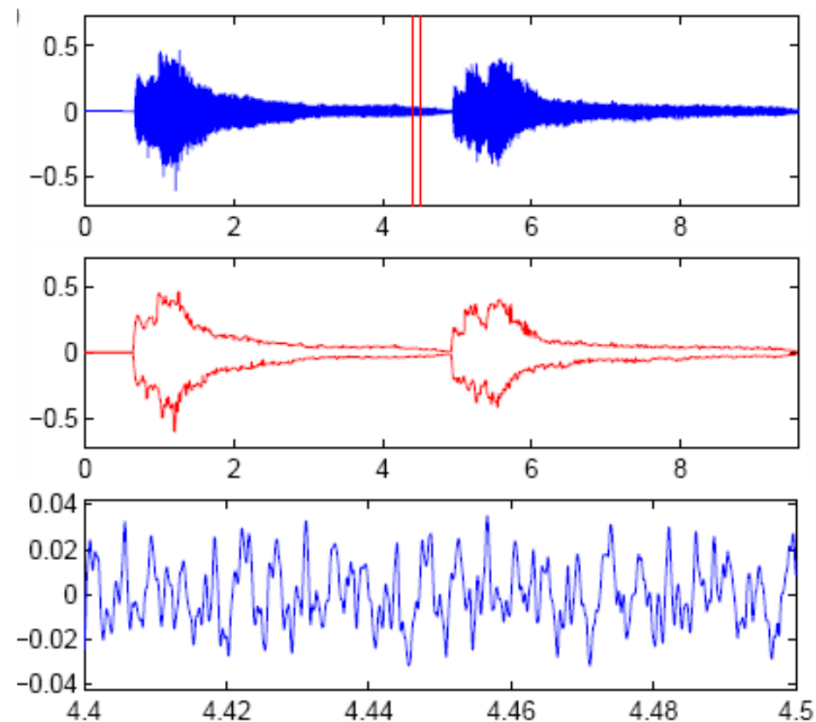
Waveform

Bernstein (orchestra)



Time (seconds)

Glen Gould (piano)



Time (seconds)

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: middle A or 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)
- Referenz: standard pitch $p = 69$ (A4) \triangleq 440 Hz
- Frequency of a note with MIDI pitch p

$$f_{\text{MIDI}}(p) = 2^{\frac{p-69}{12}} \cdot 440$$

Audio Representation

Harmonics



Mix

Harmonics: Frequency = integer multiples of fundamental frequency



Deviation in cents: +2 -14 +2 -31 +4 -14 -49 +2 +41 -31 -12

MIDI: Frequency = fundamental frequency of MIDI pitch



Stereo file: Harmonics vs. MIDI



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

- $intensity = \frac{energy}{time \cdot area} = \frac{power}{area} \quad \left(\frac{W}{m^2} \right)$
- Decibel (dB): logarithmic unit to measure intensity relative to a reference level
- Reference level: threshold of hearing (THO) $P_0 = 1 \cdot 10^{-12} \frac{W}{m^2}$
- Intensity P_1 measured in dB: $dB(P_1) = 10 \cdot \log_{10} \left(\frac{P_1}{P_0} \right)$
- Examples:
 - $P_1 = 10 \cdot P_0 \rightarrow P_1$ has a sound level of 10 dB
 - $P_2 = 100 \cdot P_0 \rightarrow P_2$ has a sound level of 20 dB

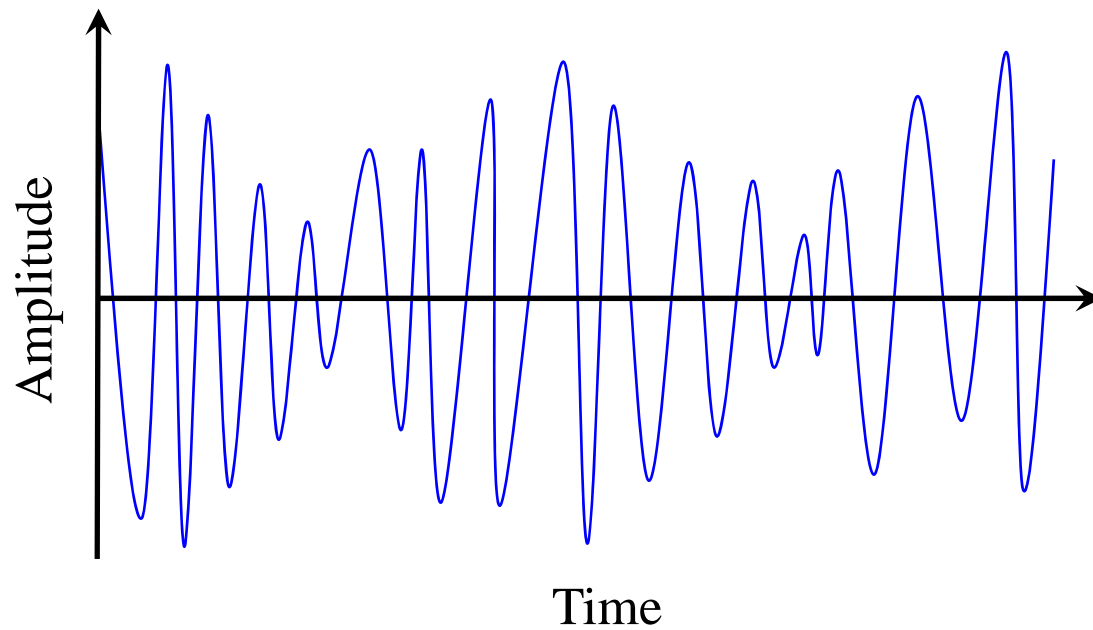
Audio Representation

Dynamics

Source	Intensity	Intensity level	× TOH
Threshold of hearing (TOH)	10^{-12}	0 dB	0
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}

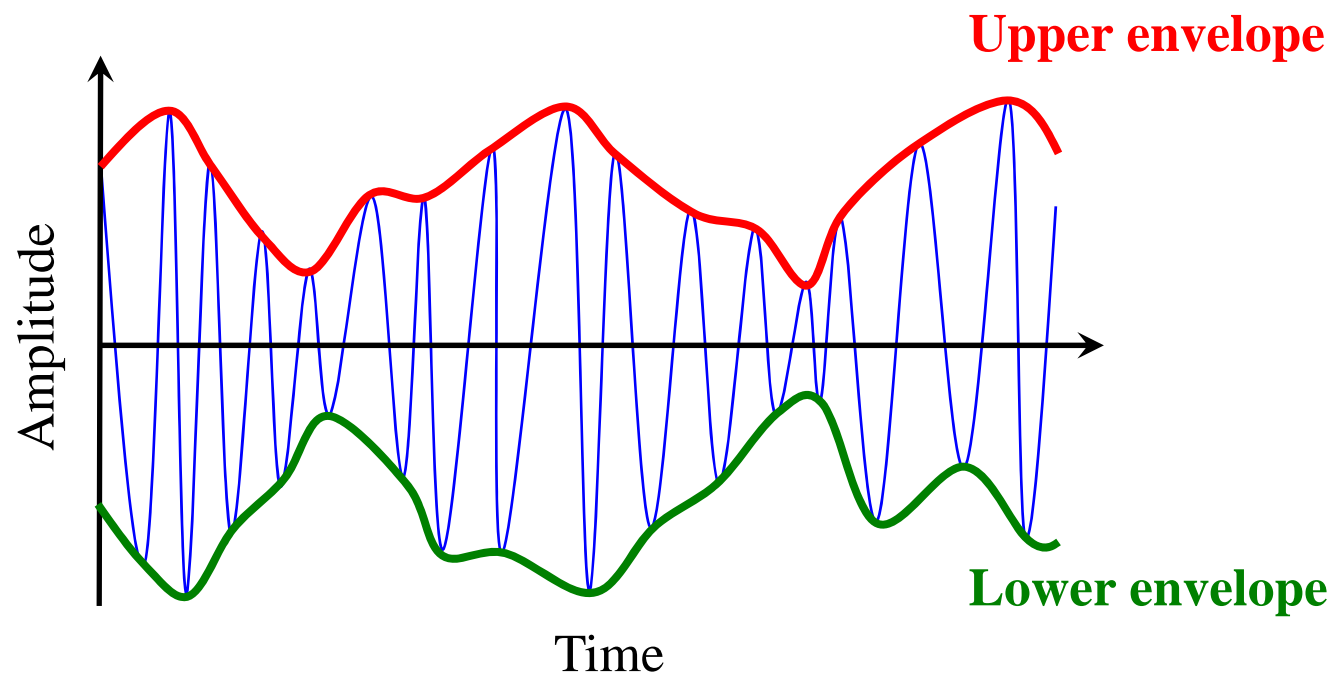
Audio Representation

Dynamics



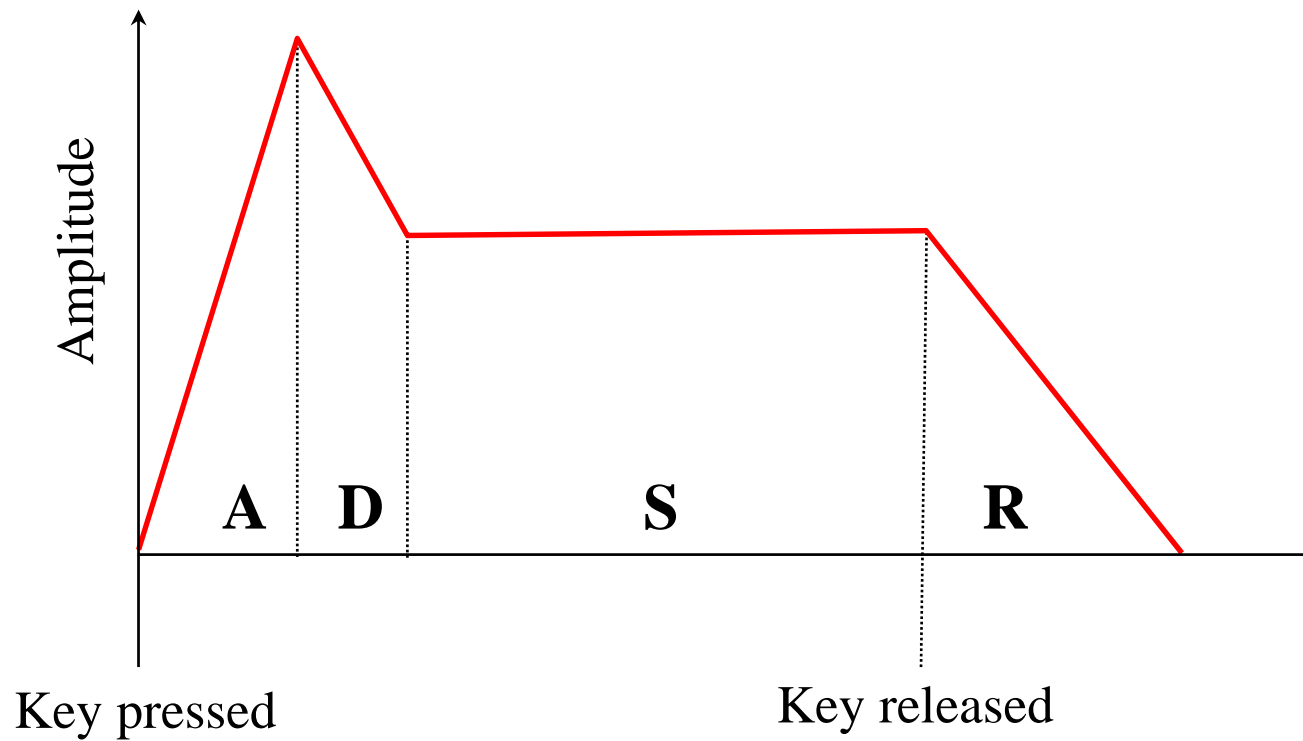
Audio Representation

Dynamics



Audio Representation

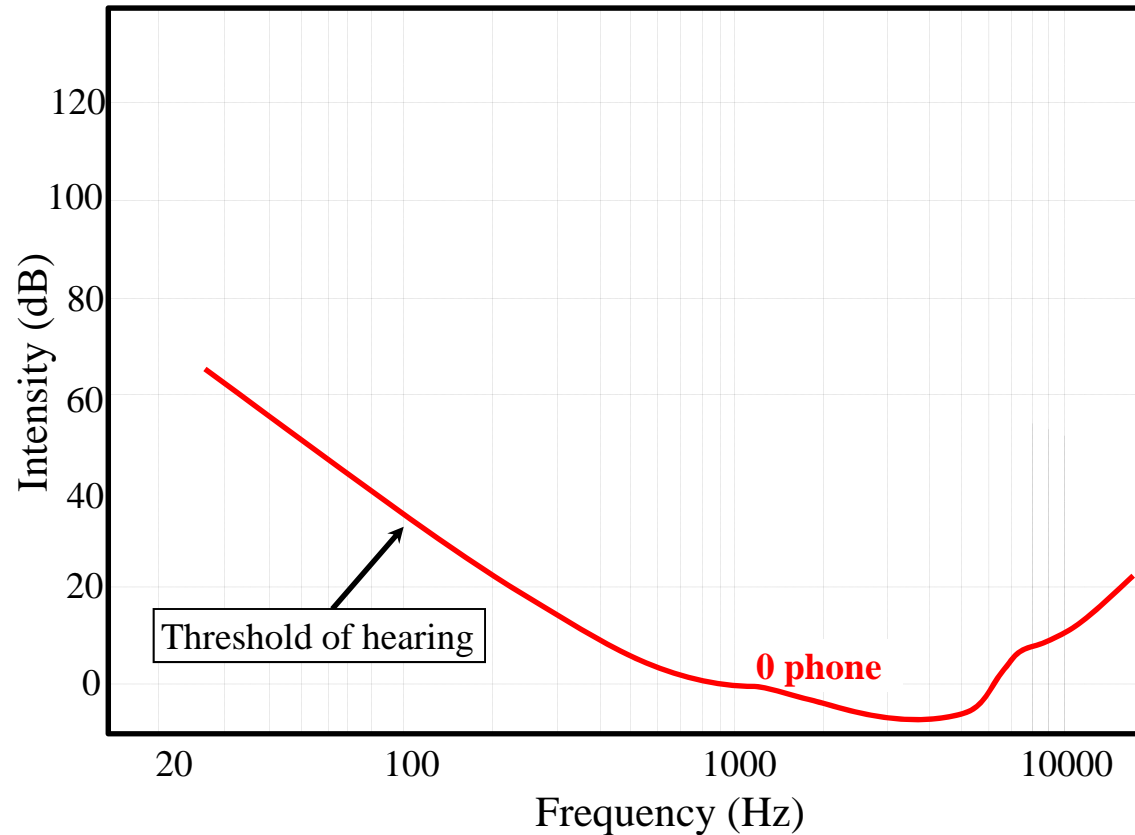
Dynamics



Audio Representation

Loudness

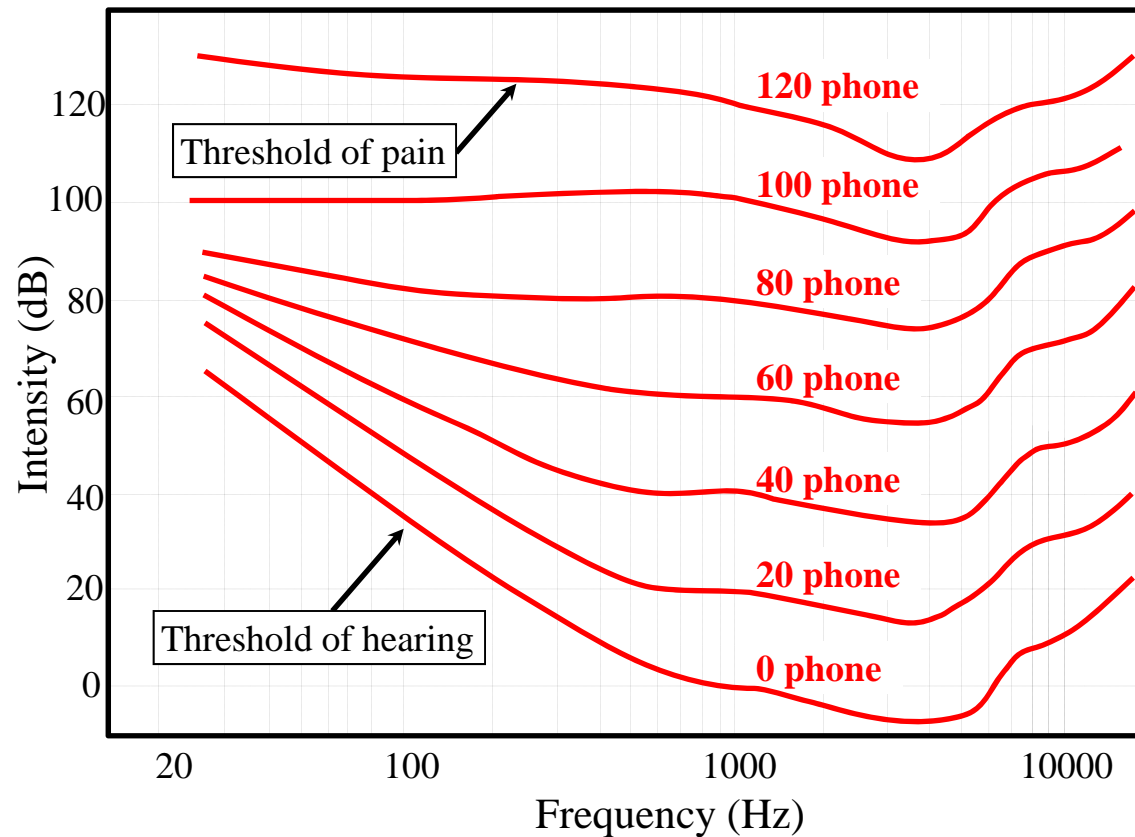
Equal-loudness contours (phone)



Audio Representation

Loudness

Equal-loudness contours (phone)



Audio Representation

Timbre

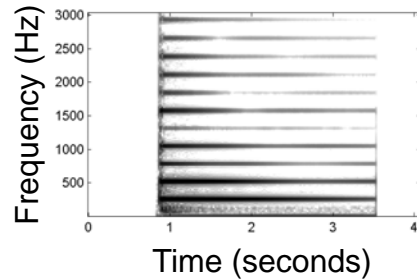
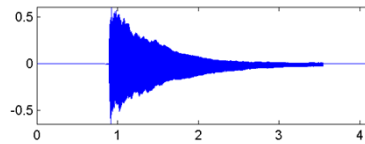
- Quality of musical sound that distinguishes different types of sound production such as voices or instruments
- Tone quality
- Tone color
- Depends on energy distribution in harmonics

Audio Representation

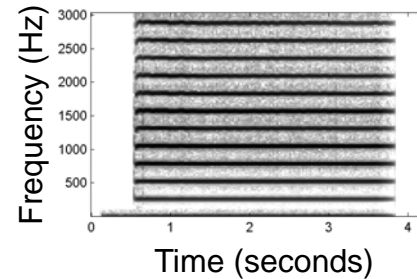
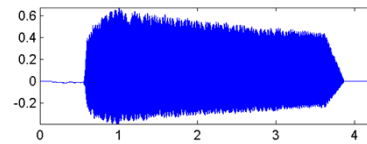
Timbre

All instruments play the same note C4 (261.6 Hz)

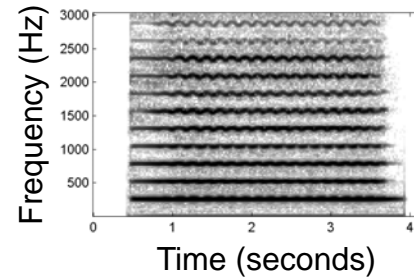
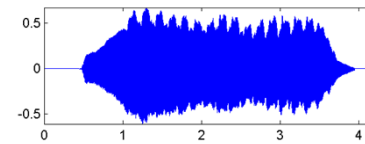
Piano ▶



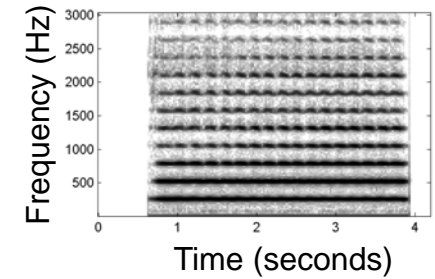
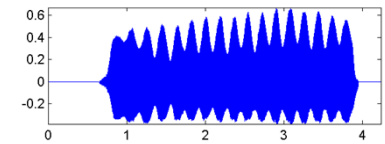
Trumpet ▶



Violine ▶



Flute ▶

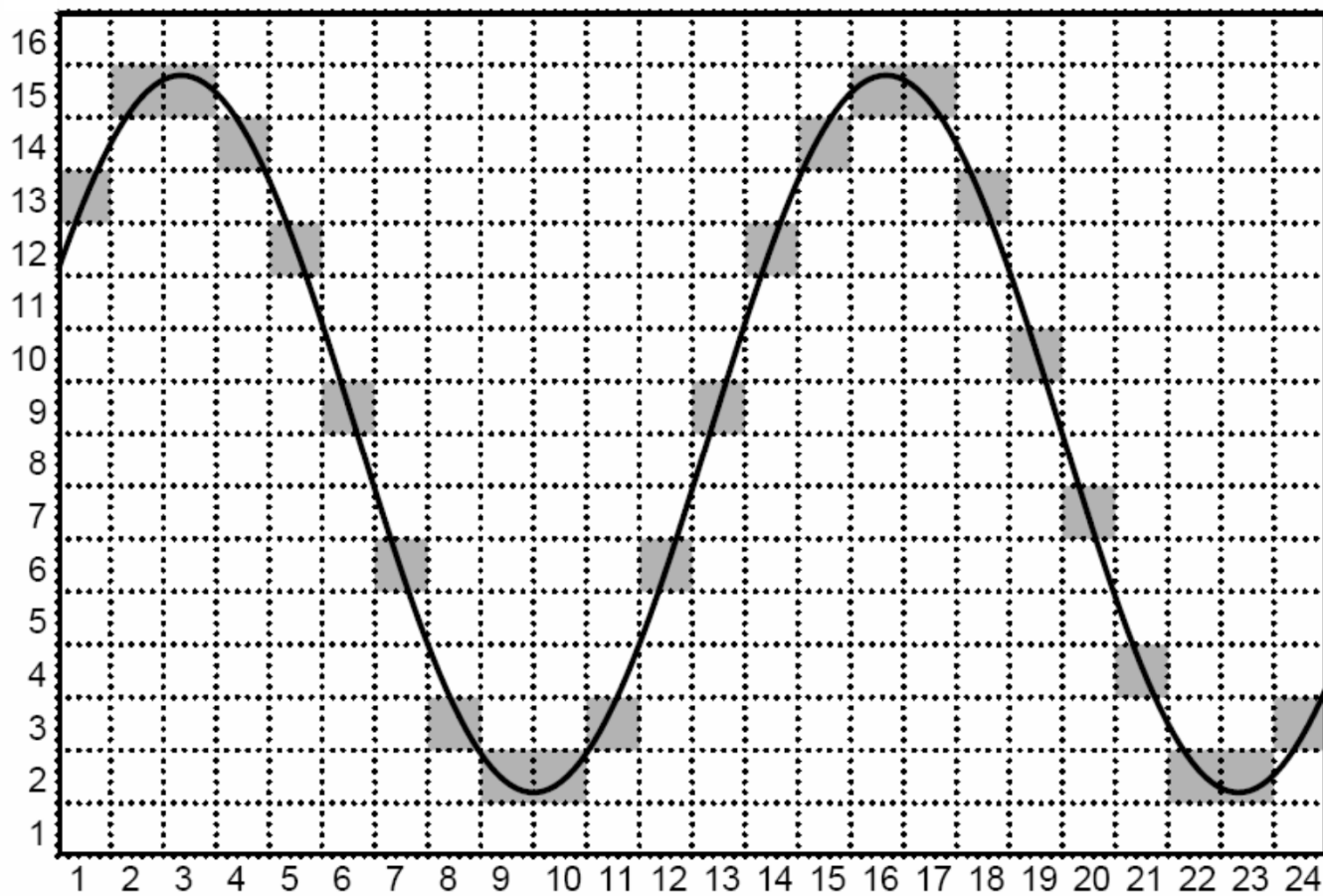


Vibrato:
Frequency
modulation

Tremolo:
Amplitude
modulation

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

