

Lecture

Music Processing

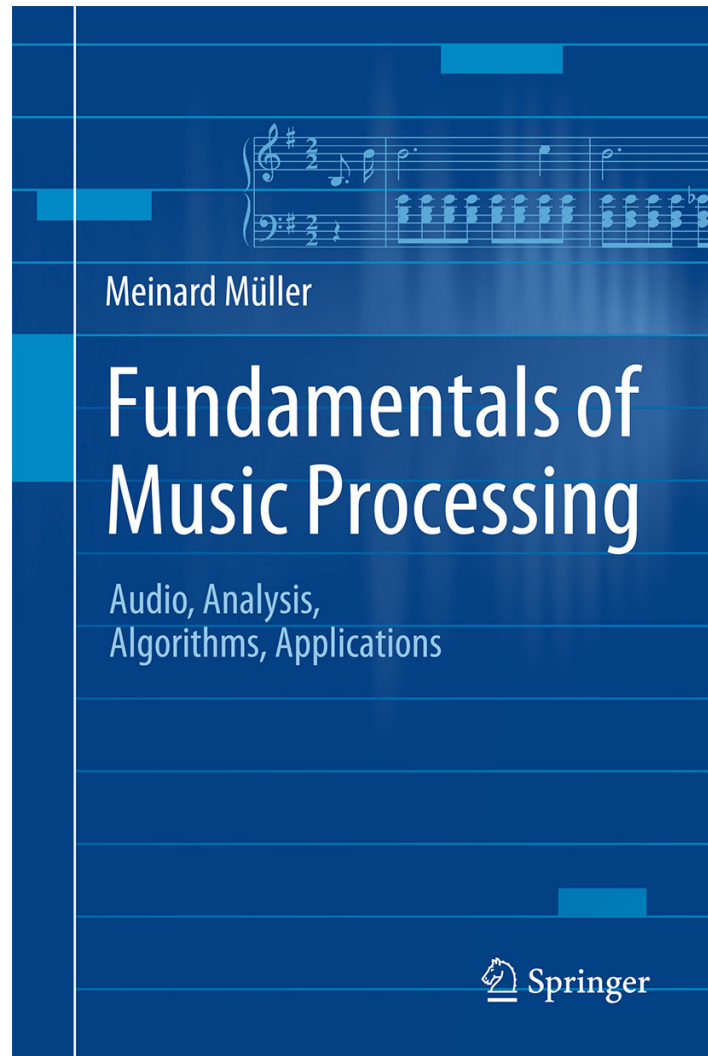
Music Representations

Meinard Müller

International Audio Laboratories Erlangen

meinard.mueller@audiolabs-erlangen.de

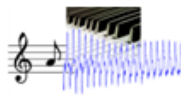

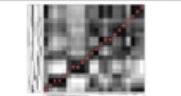
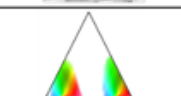

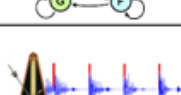


Book: Fundamentals of Music Processing



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

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

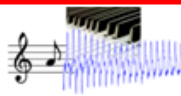

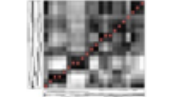


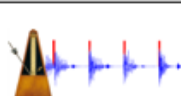
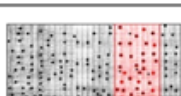
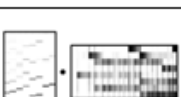
ISBN: 978-3-319-21944-8

Springer, 2015

Accompanying website:

www.music-processing.de

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

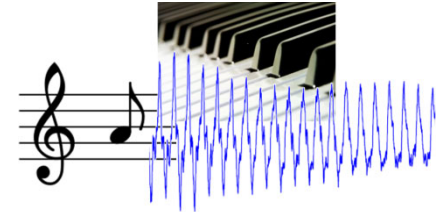
Springer, 2015

Accompanying website:

www.music-processing.de

Chapter 1: Music Representations

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



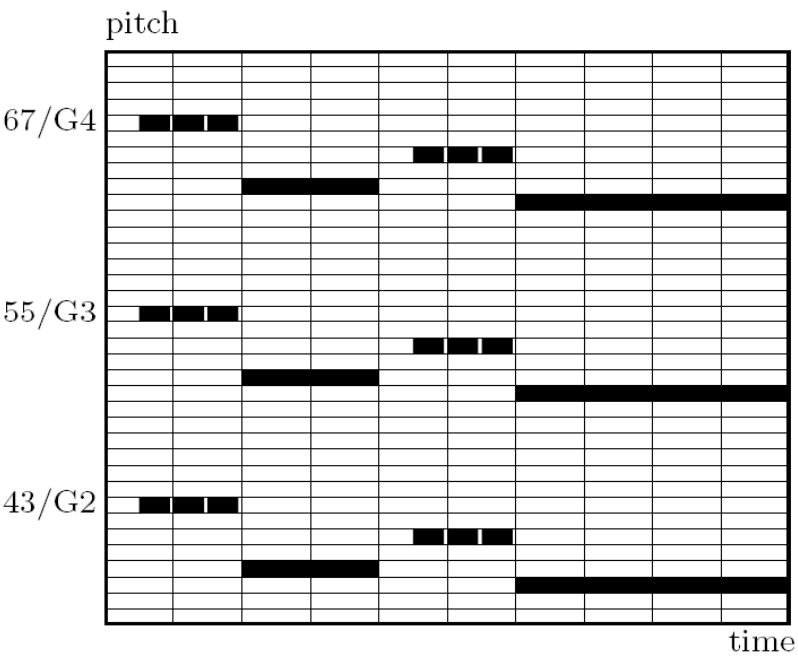
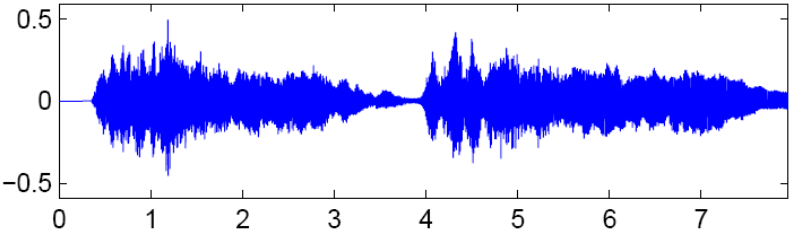
Musical information can be represented in many different ways. In Chapter 1, we consider three widely used music representations: sheet music, symbolic, and audio representations. This first chapter also introduces basic terminology that is used throughout the book. In particular, we discuss musical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

Music Representations

Allegro con brio ($\text{♩} = 108$)

ff

Ped. * Ped. *



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

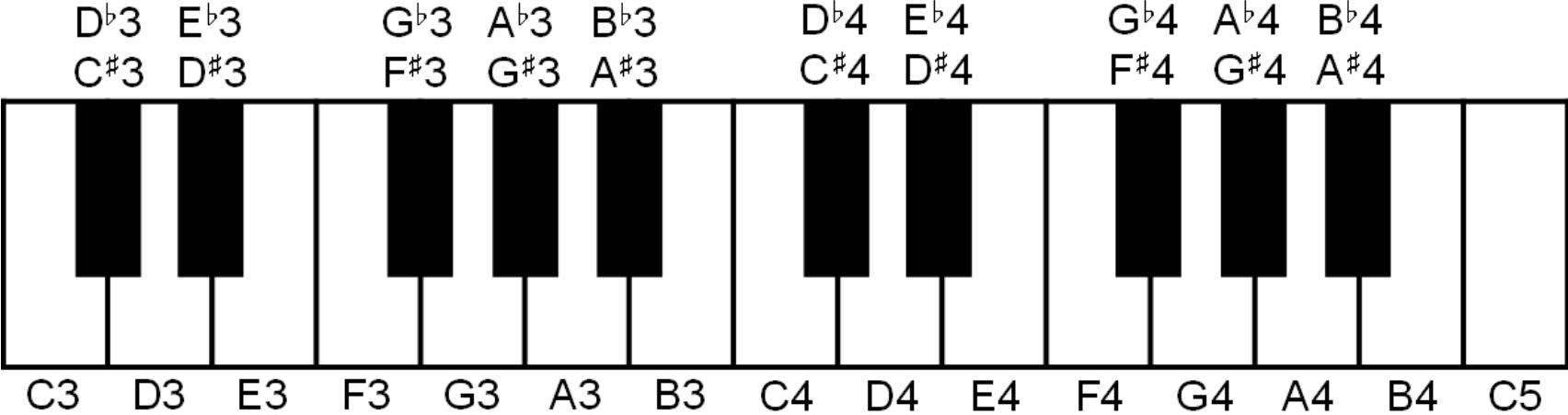
Sheet Music Representation

Allegro con brio (♩ = 108)

The image shows a musical score for piano in 2/4 time, marked *ff* (fortissimo). The tempo is *Allegro con brio* with a quarter note equal to 108 beats per minute. The score is written in a key signature of two flats (B-flat and E-flat). The music is divided into four measures. The first measure features a treble clef with a series of eighth notes and a bass clef with a series of eighth notes. The second measure has a treble clef with a half note and a bass clef with a half note. The third measure has a treble clef with a half note and a bass clef with a half note. The fourth measure has a treble clef with a half note and a bass clef with a half note. The *ff* marking is placed below the first measure. The *Ped.* marking with a flower-like symbol is placed below the second and fourth measures.

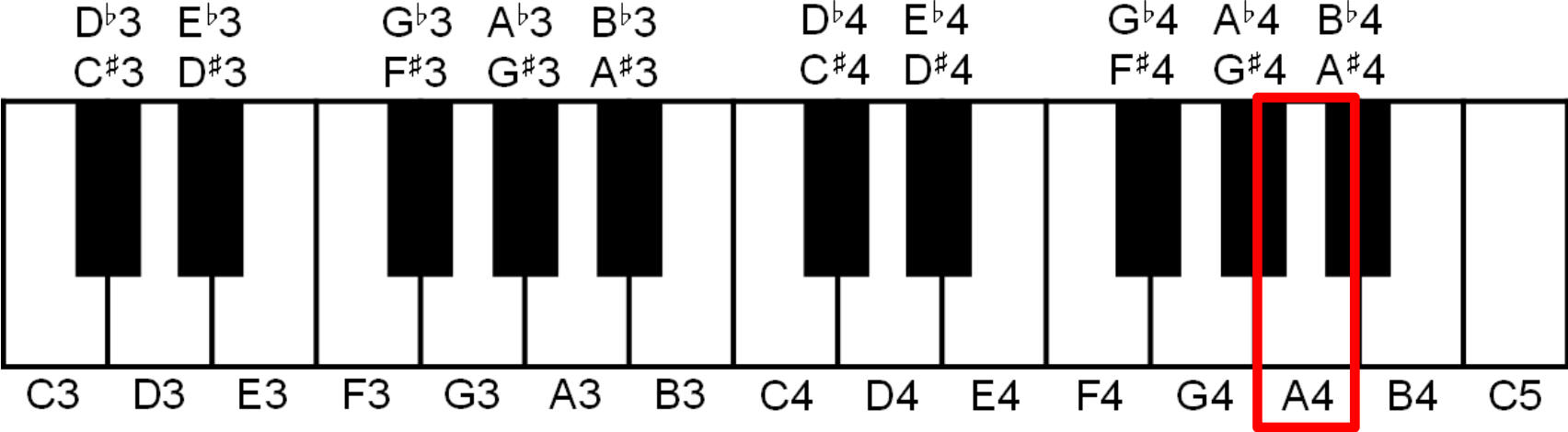
Sheet Music Representation

Piano keyboard and notes



Sheet Music Representation

Piano keyboard and notes

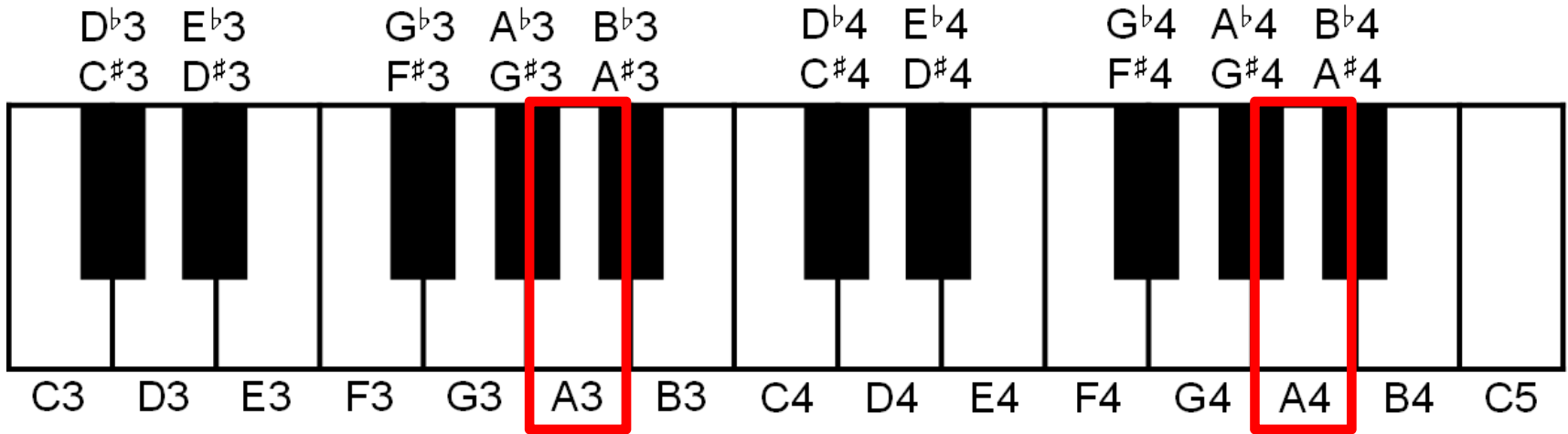


A4

A ≙ pitch spelling attribute
4 ≙ octave number

Sheet Music Representation

Piano keyboard and notes

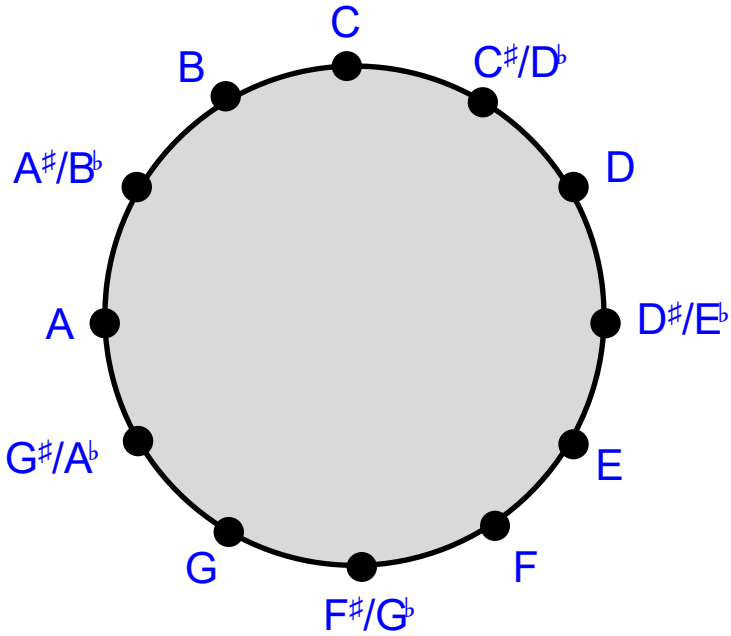


A \triangleq pitch spelling attribute
4 \triangleq octave number

Sheet Music Representation

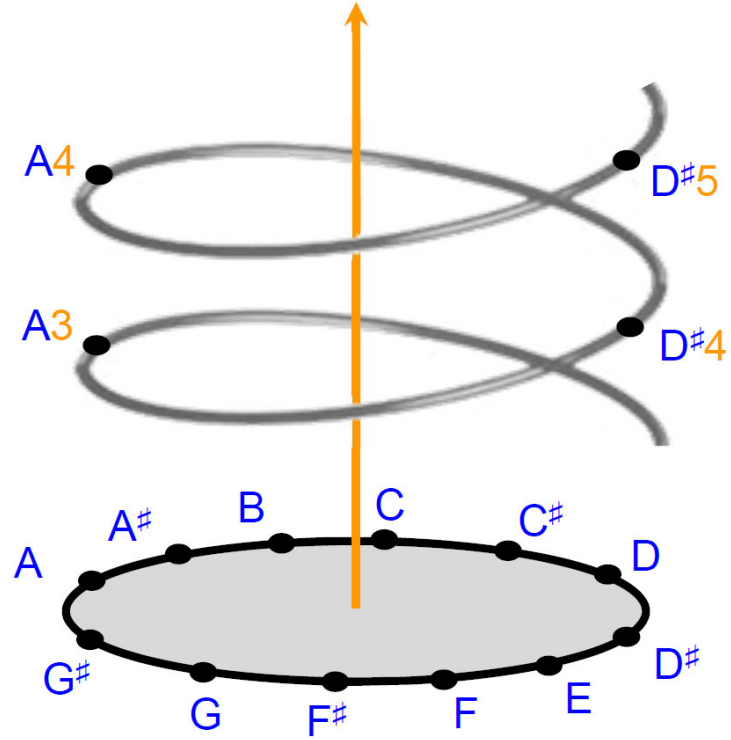
Chromatic circle

Chroma \triangleq pitch spelling attribute



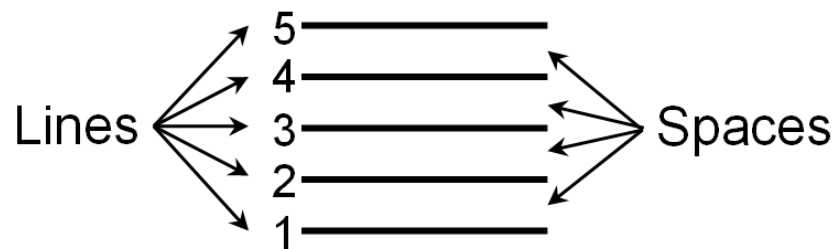
Shepard's helix of pitch

Tone height \triangleq octave number



Sheet Music Representation

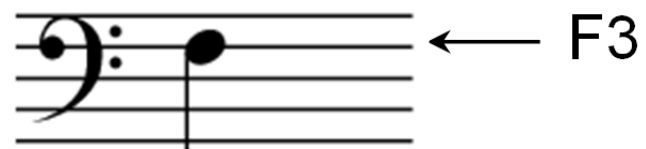
Staff



Staff with G-clef



Staff with F-clef



Sheet Music Representation

Musical score of a C-major scale



A musical score for a C-major scale, consisting of eight notes: C4, D4, E4, F4, G4, A4, B4, and C5. The notes are written on a single staff with a treble clef. The notes are represented by black dots with vertical stems, and the final note (C5) is marked with a double bar line. Below the staff, the notes are labeled with their corresponding scientific pitch notation: C4, D4, E4, F4, G4, A4, B4, and C5.

Sheet Music Representation

Musical score of a C-major scale



A musical staff in treble clef showing the C-major scale. The notes are C4, D4, E4, F4, G4, A4, B4, and C5. The notes are placed on the lines and spaces of the staff: C4 on the first space, D4 on the second line, E4 on the second space, F4 on the third line, G4 on the third space, A4 on the fourth line, B4 on the fourth space, and C5 on the fifth line. The notes are connected by a single horizontal line.

C4 D4 E4 F4 G4 A4 B4 C5

Musical score of a C-minor scale



A musical staff in treble clef showing the C-minor scale. The notes are C4, D4, E^b4, F4, G4, A^b4, B^b4, and C5. The notes are placed on the lines and spaces of the staff: C4 on the first space, D4 on the second line, E^b4 on the second space, F4 on the third line, G4 on the third space, A^b4 on the fourth line, B^b4 on the fourth space, and C5 on the fifth line. The notes are connected by a single horizontal line. The key signature consists of three flats: B^b4, E^b4, and A^b4.

C4 D4 E^b4 F4 G4 A^b4 B^b4 C5

Key signature consisting of three flats

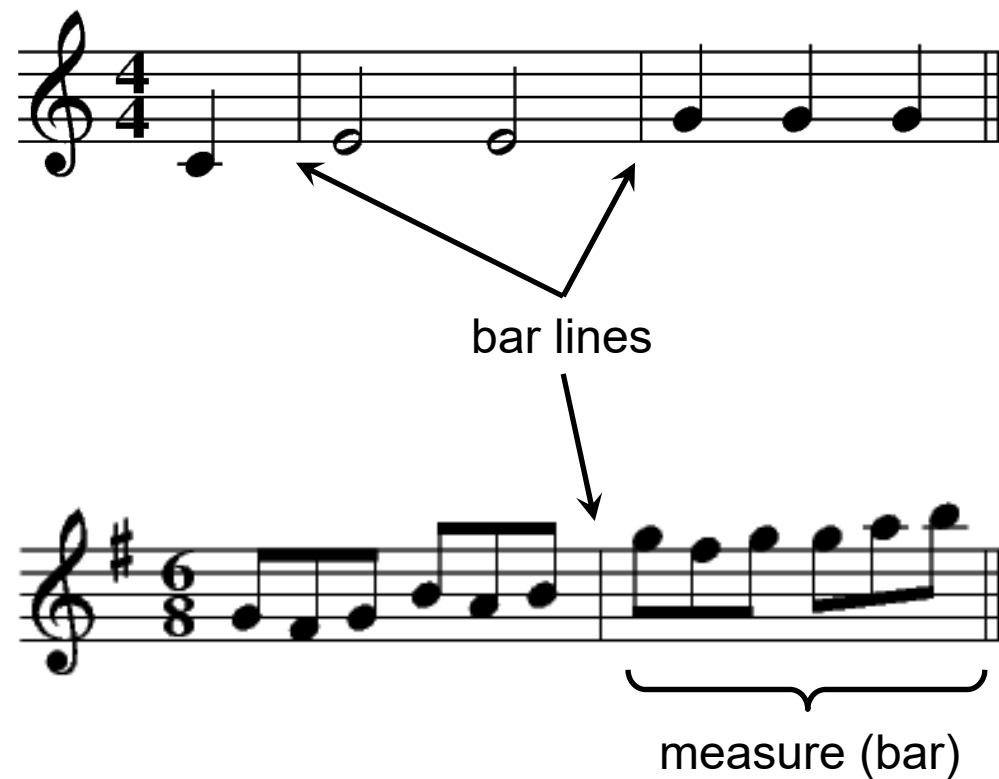
Sheet Music Representation

Time signature



Sheet Music Representation

Time signature



Sheet Music Representation

Time signature

Four quarter notes
per measure



bar lines

Six eighth notes
per measure

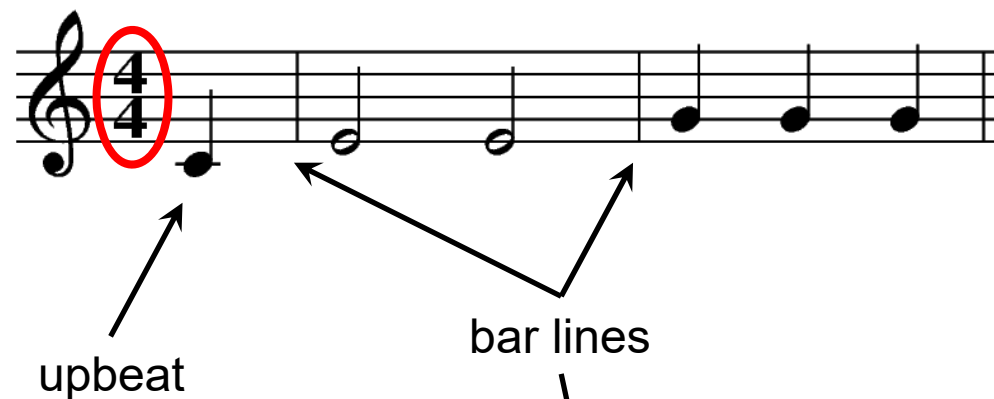


measure (bar)

Sheet Music Representation

Time signature

Four quarter notes
per measure



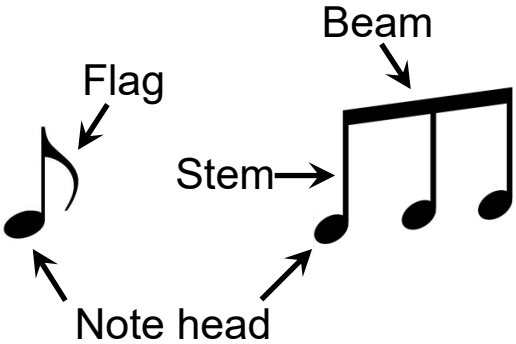
Six eighth notes
per measure



Sheet Music Representation

Note durations

Parts of a note



Different durations of notes

A musical staff in treble clef showing five notes of different durations: a whole note (filled oval), a half note (filled oval with a stem), a quarter note (filled oval with a stem and a flag), an eighth note (filled oval with a stem and a flag, beamed to the next note), and a sixteenth note (filled oval with a stem and a flag, beamed to the next note).

Whole note Half note Quarter note Eighth note Sixteenth note

Different durations of rests

A musical staff in treble clef showing five rests of different durations: a whole rest (horizontal bar), a half rest (horizontal bar), a quarter rest (vertical line with a flag), an eighth rest (vertical line with a flag, beamed to the next rest), and a sixteenth rest (vertical line with a flag, beamed to the next rest).

Whole rest Half rest Quarter rest Eighth rest Sixteenth rest

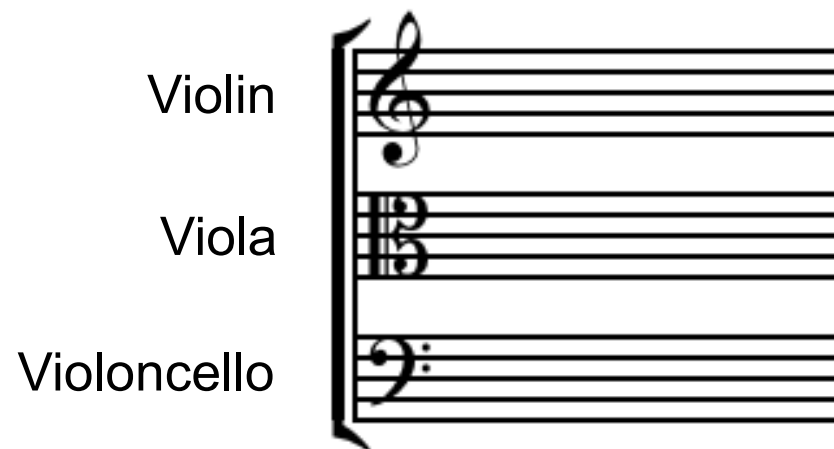
Sheet Music Representation

Staff systems

Piano



Strings



Sheet Music Representation

Dynamics and articulation

The image shows a single staff of music in 4/4 time, starting with a treble clef. The melody consists of quarter notes: C4, D4, E4, F4, G4, A4, B4, C5, B4, A4, G4, F4, E4, D4, C4. Above the staff, there are four dynamic markings: 'crescendo' with a wedge-shaped hairpin opening to the right, 'diminuendo' with a wedge-shaped hairpin closing to the right, 'piano' with the letter *p*, and 'forte' with the letter *f*. Below the staff, there are two articulation markings: 'legato' with an upward-pointing arrow and a slur under the notes G4 and A4, and 'staccato' with a downward-pointing arrow and a V-shaped mark under the notes B4 and C5. At the end of the staff, the lyrics 'do re mi fa so la si do' are written below the notes, with the word 'lyrics' centered under the final 'do'.

Sheet Music 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.

The image displays a page of sheet music for an orchestra. It features 13 staves, each labeled with an instrument. The tempo is marked as 'Allegro con brio' with a metronome marking of 108 quarter notes per minute. The time signature is 2/4, and the key signature has two flats (B-flat major). The music is written in a standard orchestral format. The first section of music (measures 1-10) is marked with a forte dynamic (ff). The second section (measures 11-20) is marked with a piano dynamic (p). The third section (measures 21-30) is also marked with a piano dynamic (p). The fourth section (measures 31-40) is marked with a forte dynamic (ff). The fifth section (measures 41-50) is marked with a piano dynamic (p). The sixth section (measures 51-60) is marked with a forte dynamic (ff). The seventh section (measures 61-70) is marked with a piano dynamic (p). The eighth section (measures 71-80) is marked with a forte dynamic (ff). The ninth section (measures 81-90) is marked with a piano dynamic (p). The tenth section (measures 91-100) is marked with a forte dynamic (ff). The eleventh section (measures 101-110) is marked with a piano dynamic (p). The twelfth section (measures 111-120) is marked with a forte dynamic (ff). The thirteenth section (measures 121-130) is marked with a piano dynamic (p).

Sheet Music Representation

A handwritten musical score on a single staff in 4/4 time, featuring a treble clef and a key signature of one flat (Bb). The lyrics are "WIKIPEDIA". The notation includes a whole note for "WI", a half note for "KI", and a quarter note followed by a dotted quarter note for "PE - DI - A". Handwritten annotations include "CHORD SYMBOLS" with lines pointing to "C7" above the first measure and "F" above the third measure. "MELODY" is written above the second measure with a line pointing to the note. "LYRIC" is written below the final note with a line pointing to it.

CHORD SYMBOLS

C7

MELODY

F

WI - KI - PE - DI - A

LYRIC

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

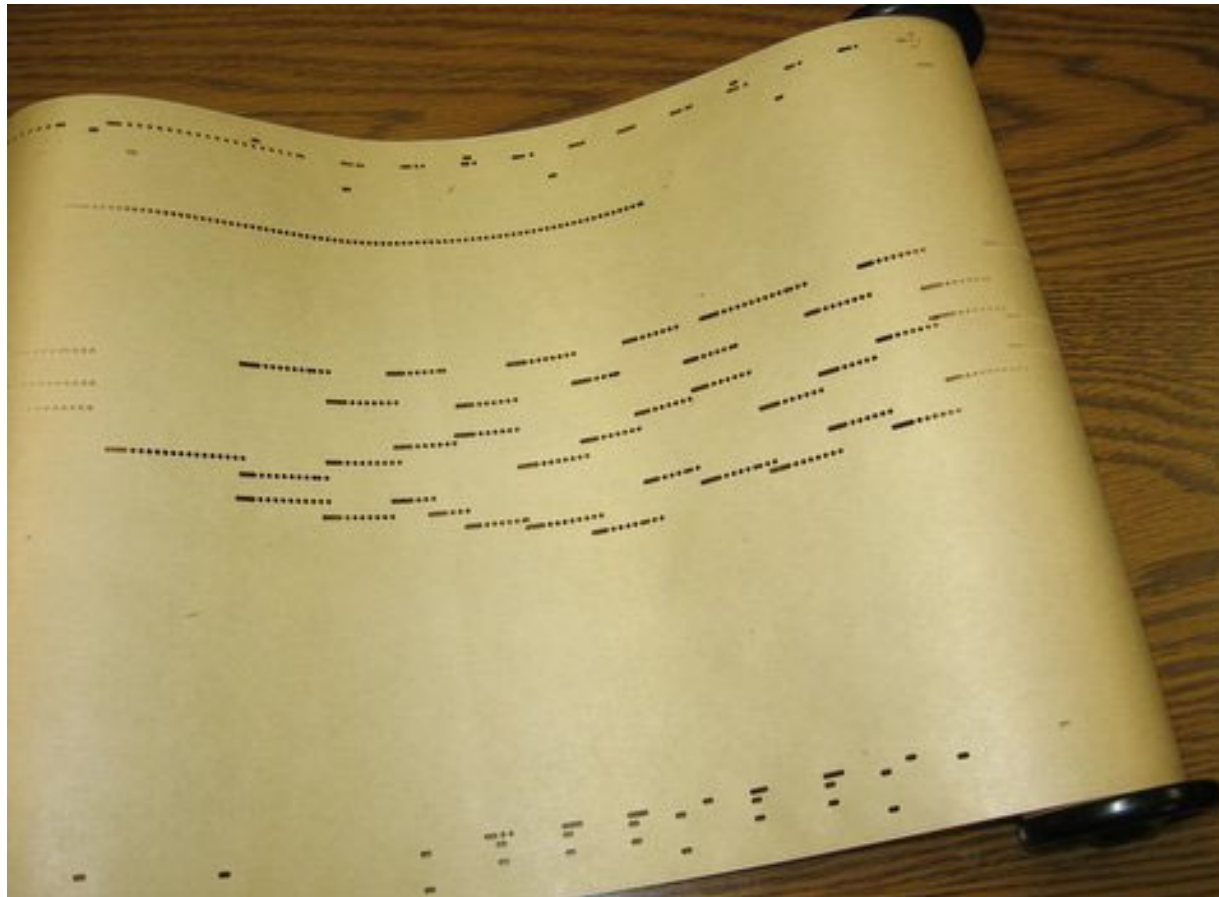
MusicXML

```
<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



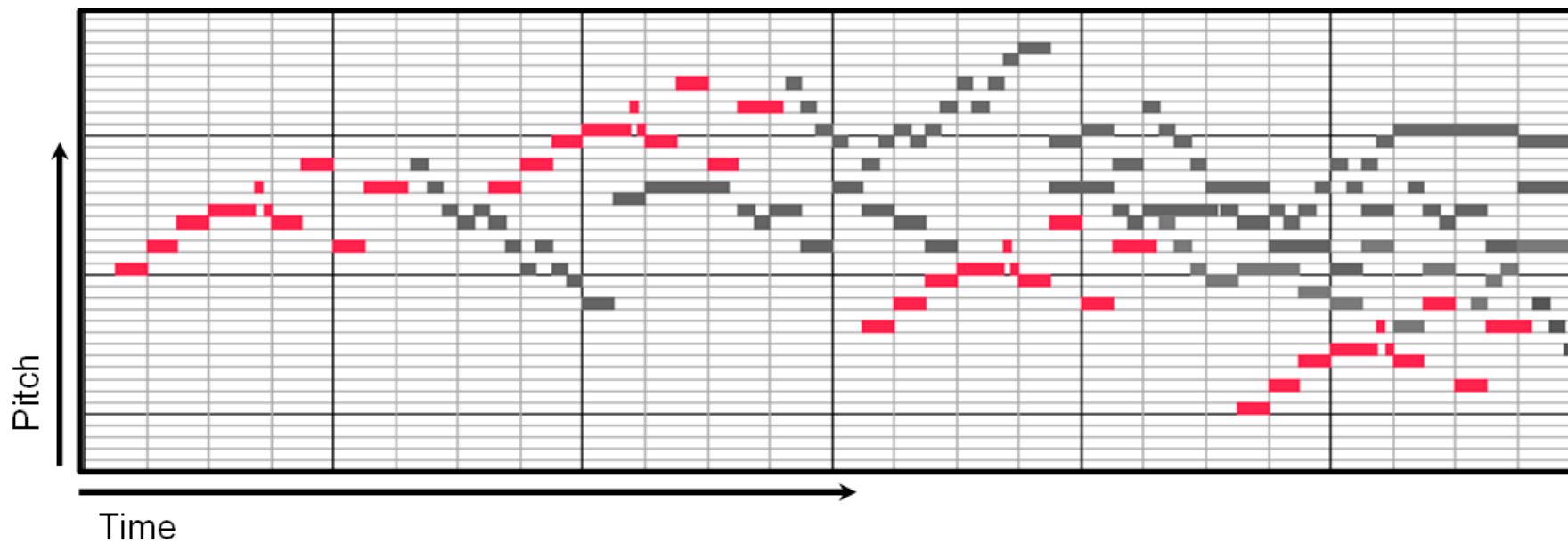
Symbolic Representation

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

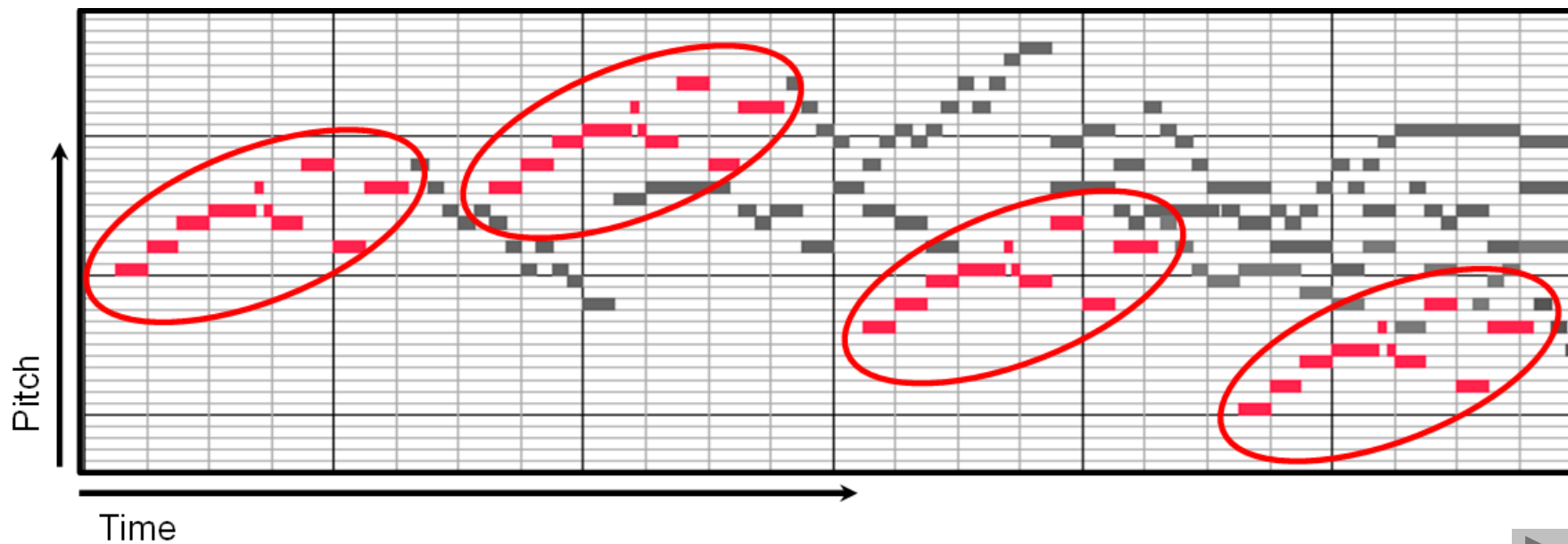
Symbolic Representation

Piano roll representation



Symbolic Representation

Piano roll representation



Symbolic Representation

MIDI representation

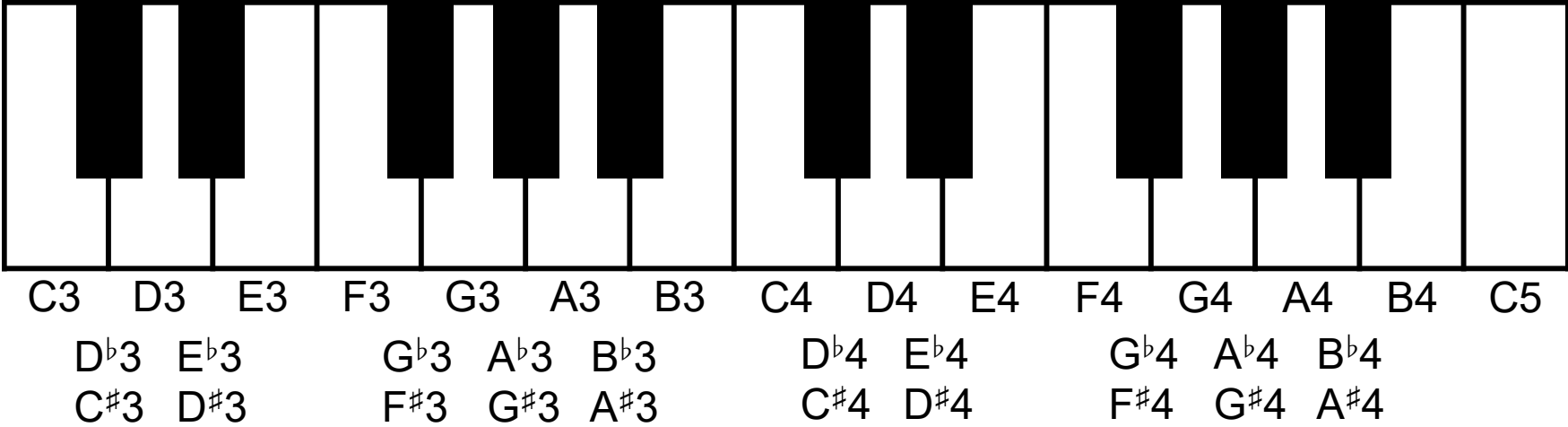
- Musical Instrument Digital Interface (MIDI)
- Standard protocol for controlling and synchronizing digital instruments
- Standard MIDI File (SMF) is used for collecting and storing MIDI messages
- SMF file is often called MIDI file

Symbolic Representation

MIDI representation

MIDI note numbers (MNN) \triangleq piano keys

49 51 54 56 58 61 63 66 68 70
48 50 52 53 55 57 59 60 62 64 65 67 69 71 72



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

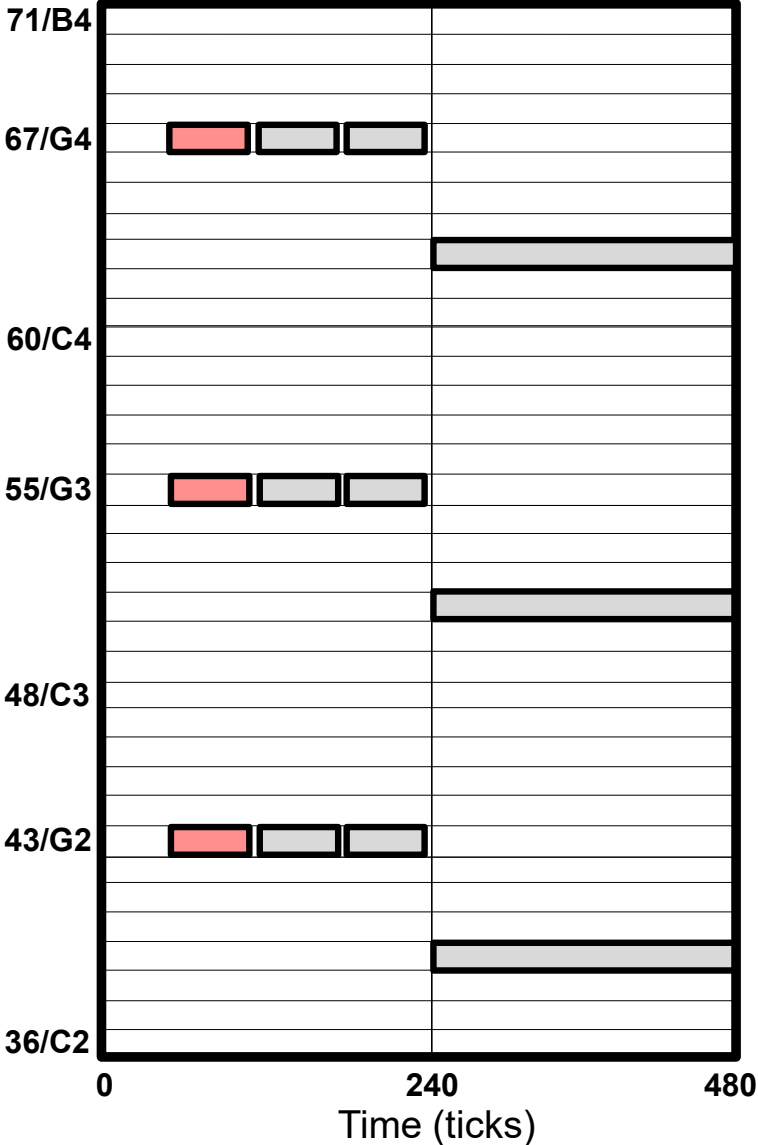
MIDI representation



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



Karajan



Scherbakov (piano)

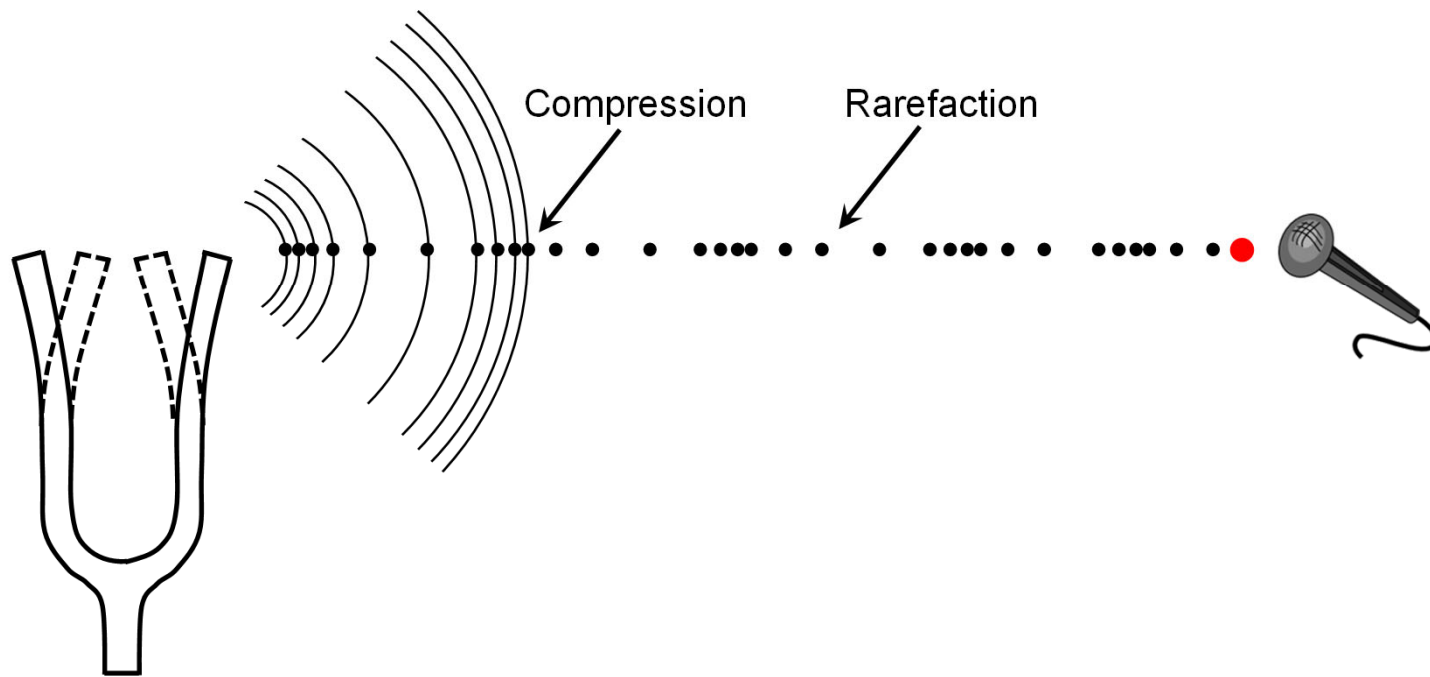


MIDI (piano)



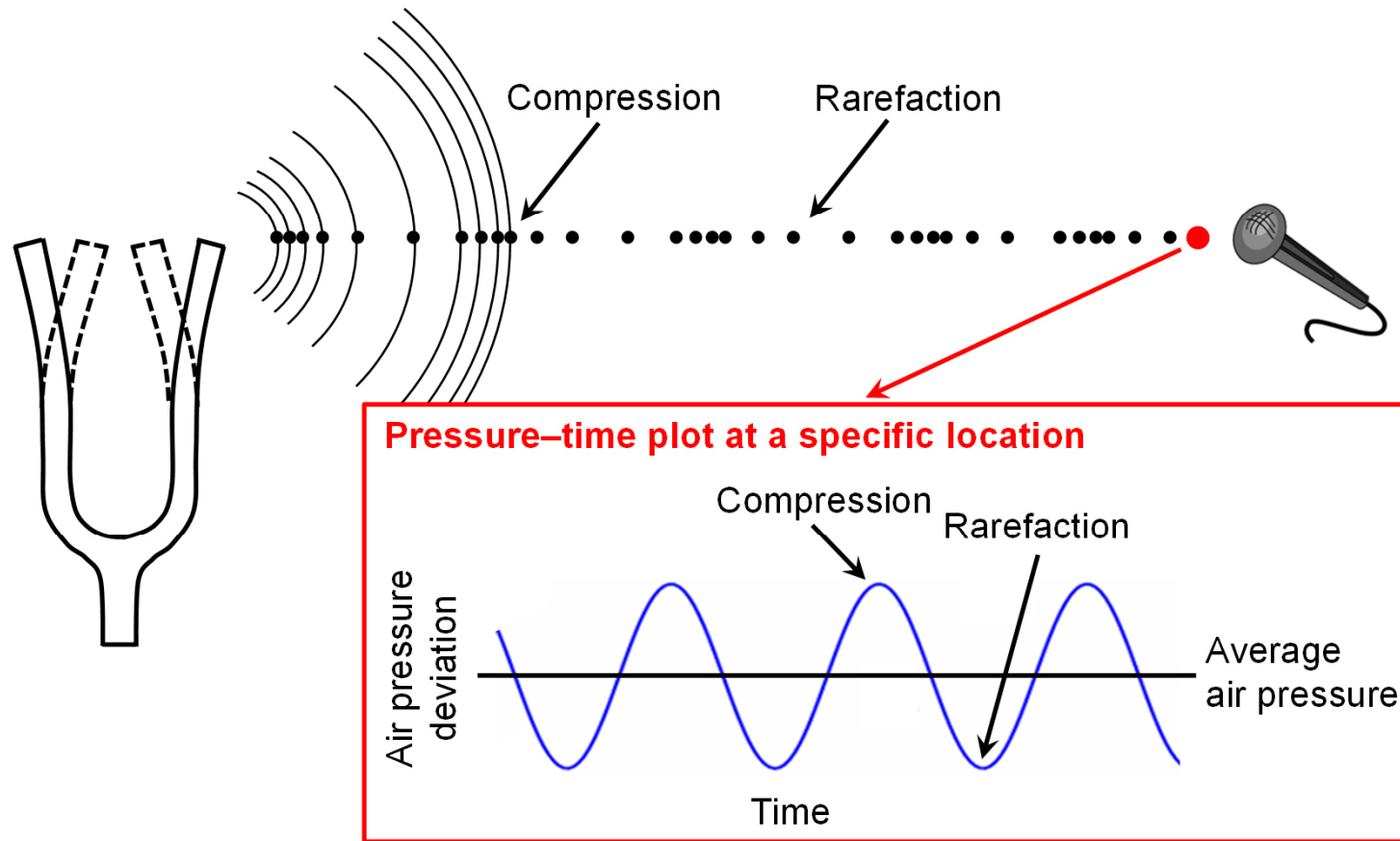
Audio Representation

Waveform



Audio Representation

Waveform



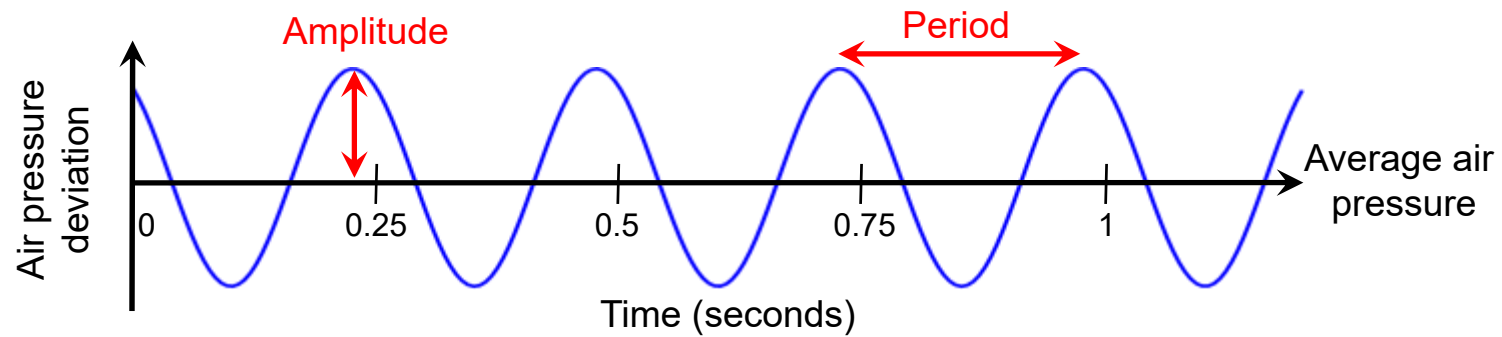
Audio Representation

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



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

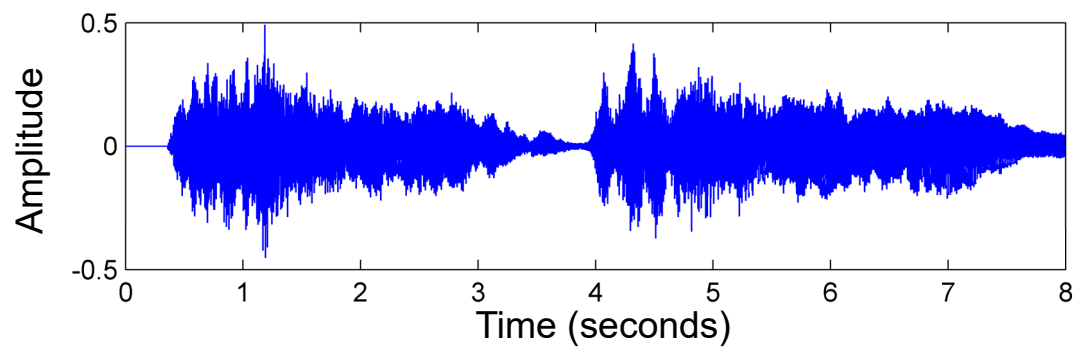
Audio Representation

Waveform

Allegro con brio ($\text{♩} = 108$)

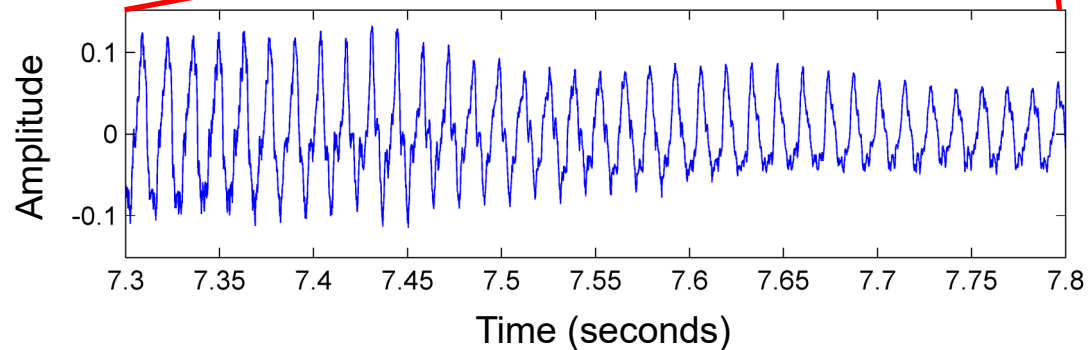
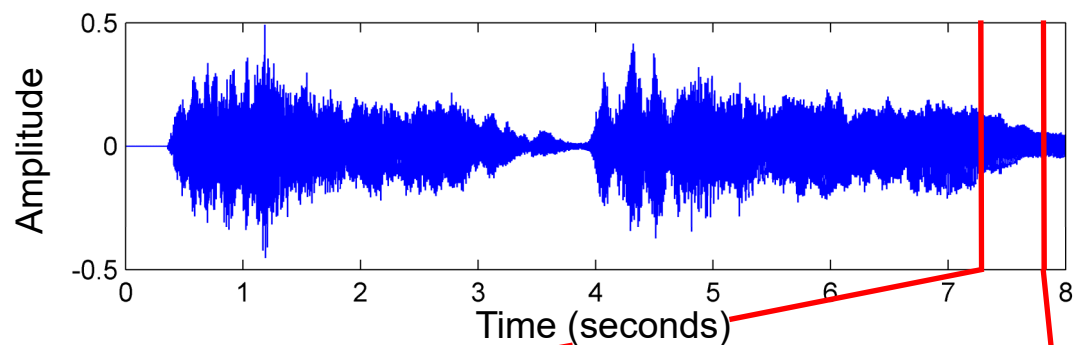


The musical score is for a piano piece in 2/4 time, marked *ff* (fortissimo) and *Allegro con brio* with a tempo of 108 beats per minute. The score consists of two staves: a treble clef staff and a bass clef staff. The treble staff begins with a series of eighth-note chords, followed by a half-note chord. The bass staff begins with a series of eighth notes, followed by a half-note chord. There are two *Rit.* markings with asterisks in the bass staff, one at approximately 4.5 seconds and another at approximately 7.5 seconds.



Audio Representation

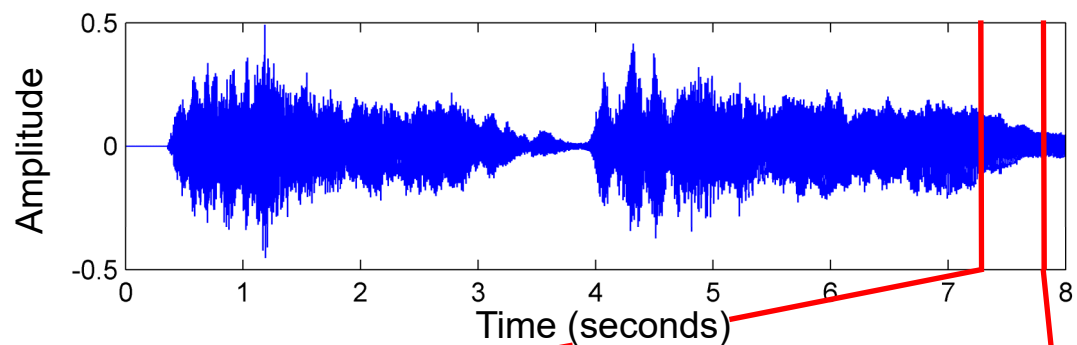
Waveform



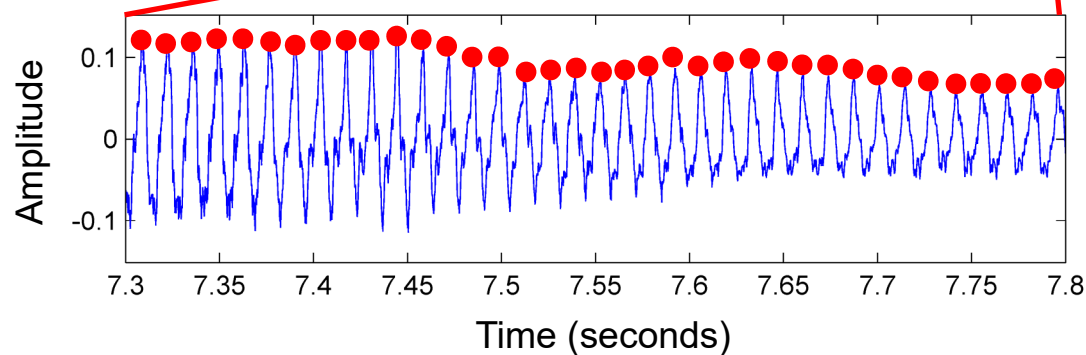
D2 (73.4 Hz)

Audio Representation

Waveform



D2 (73.4 Hz)



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)
- Referenz 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 ω_1 and ω_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



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



Mix

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 = $\frac{\text{energy}}{\text{time} \cdot \text{area}} = \frac{\text{power}}{\text{area}}$ $\left(\frac{\text{W}}{\text{m}^2} \right)$
- 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

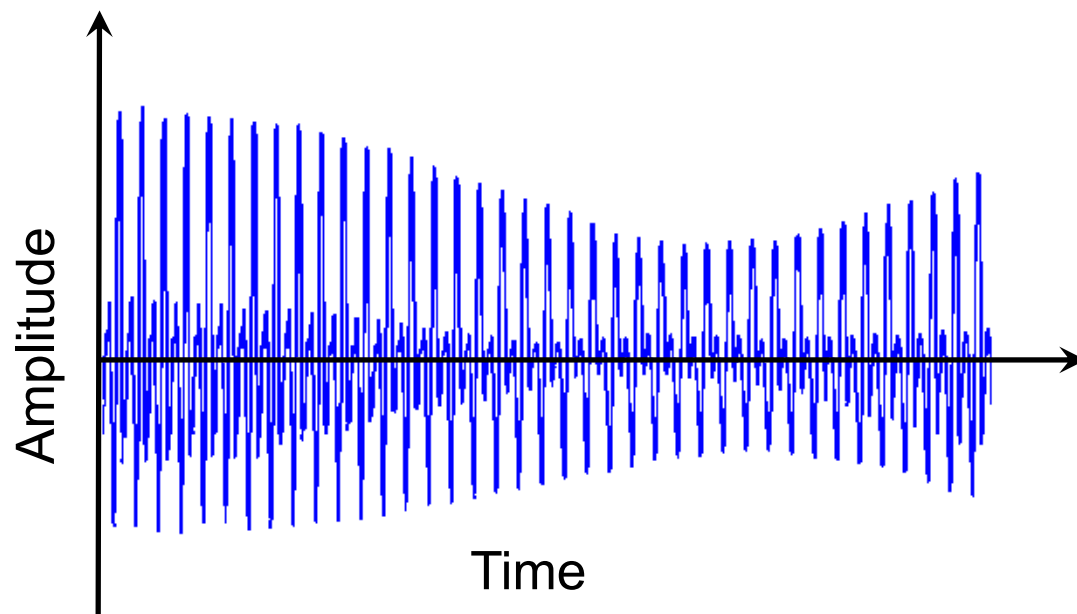
Audio Representation

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}

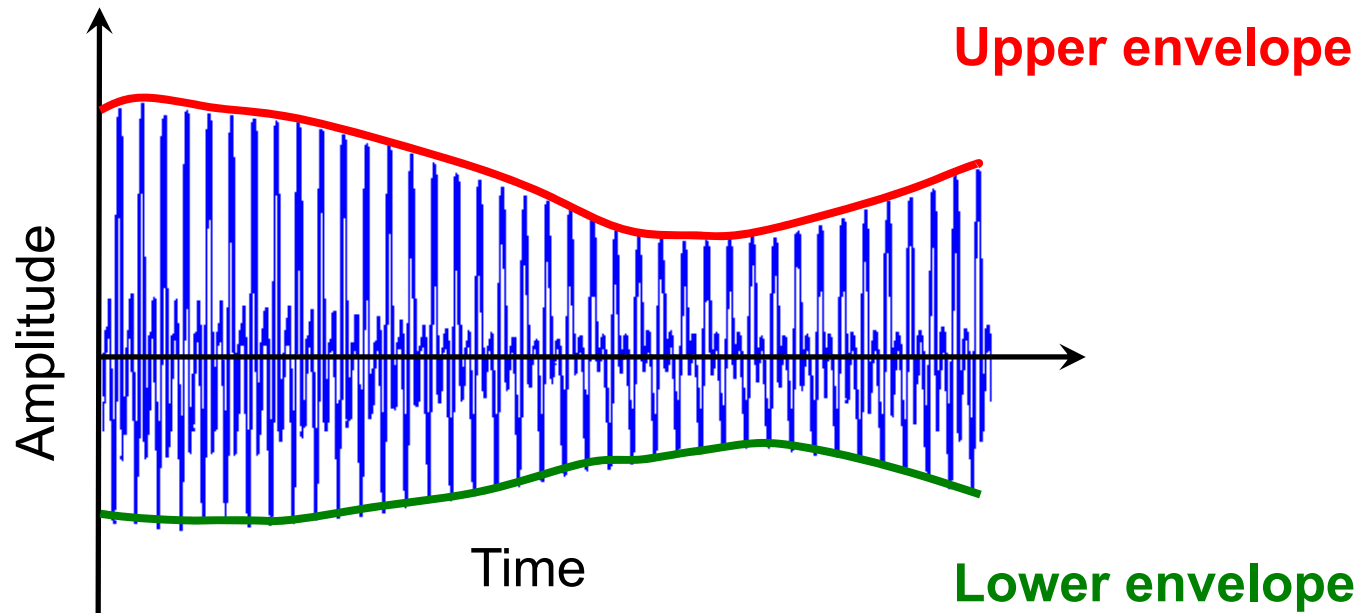
Audio Representation

Dynamics



Audio Representation

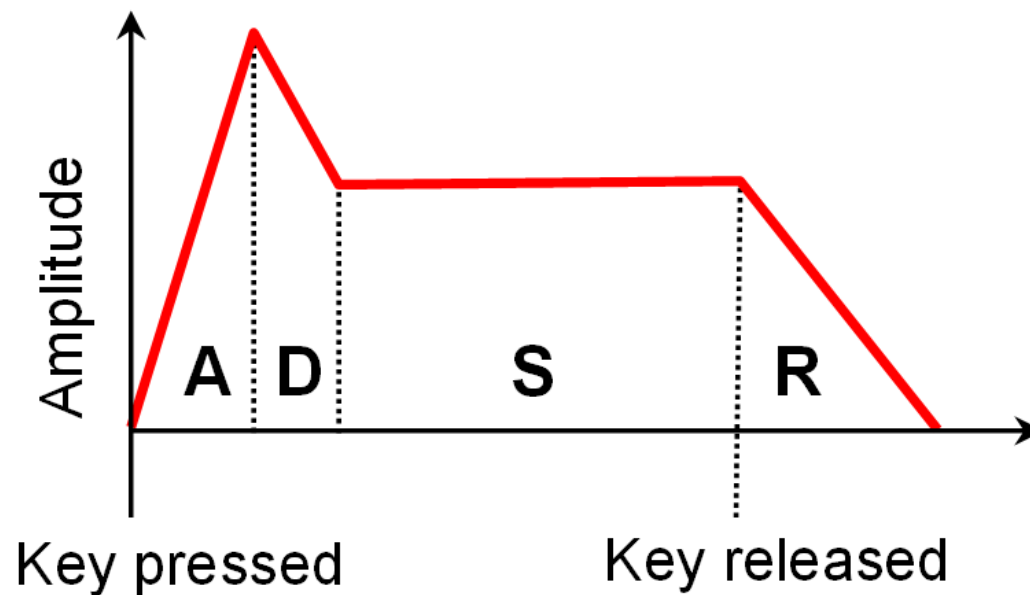
Dynamics



Audio Representation

Dynamics

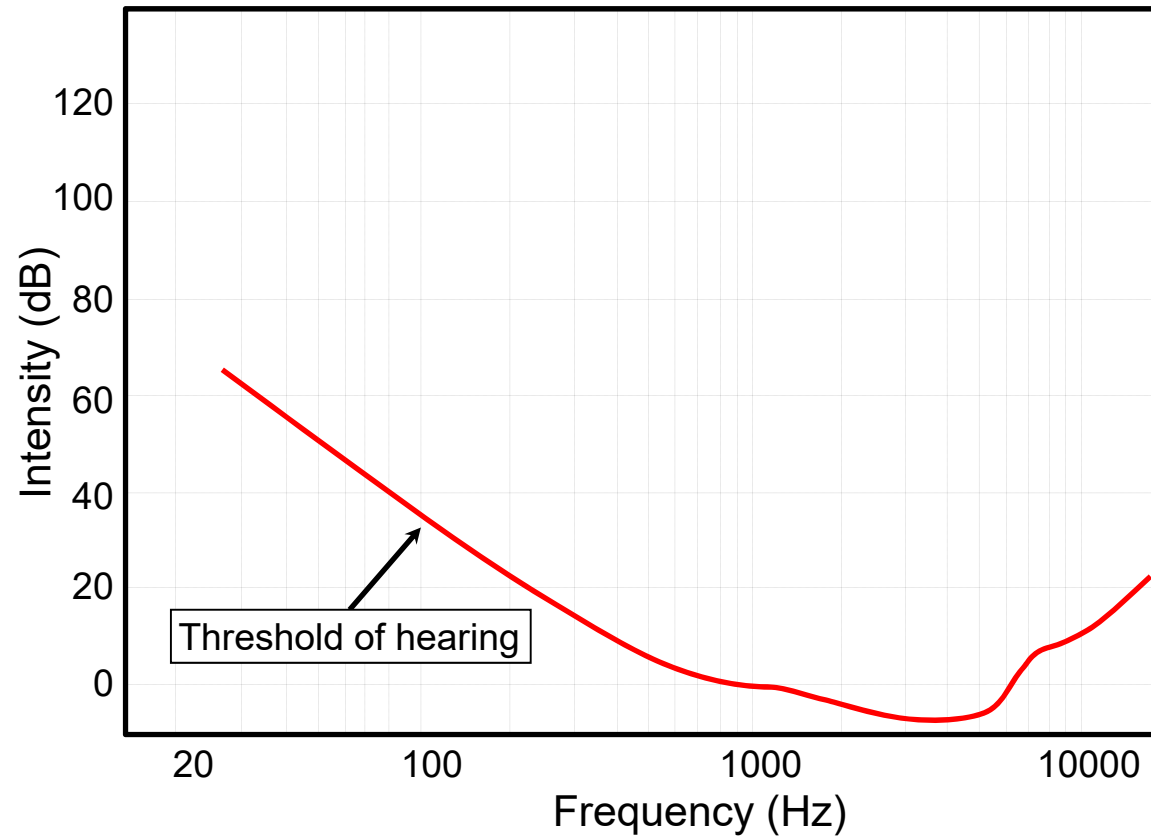
ADSR model: attack (A), decay (D), sustain (S), and release (R) phase



Audio Representation

Loudness

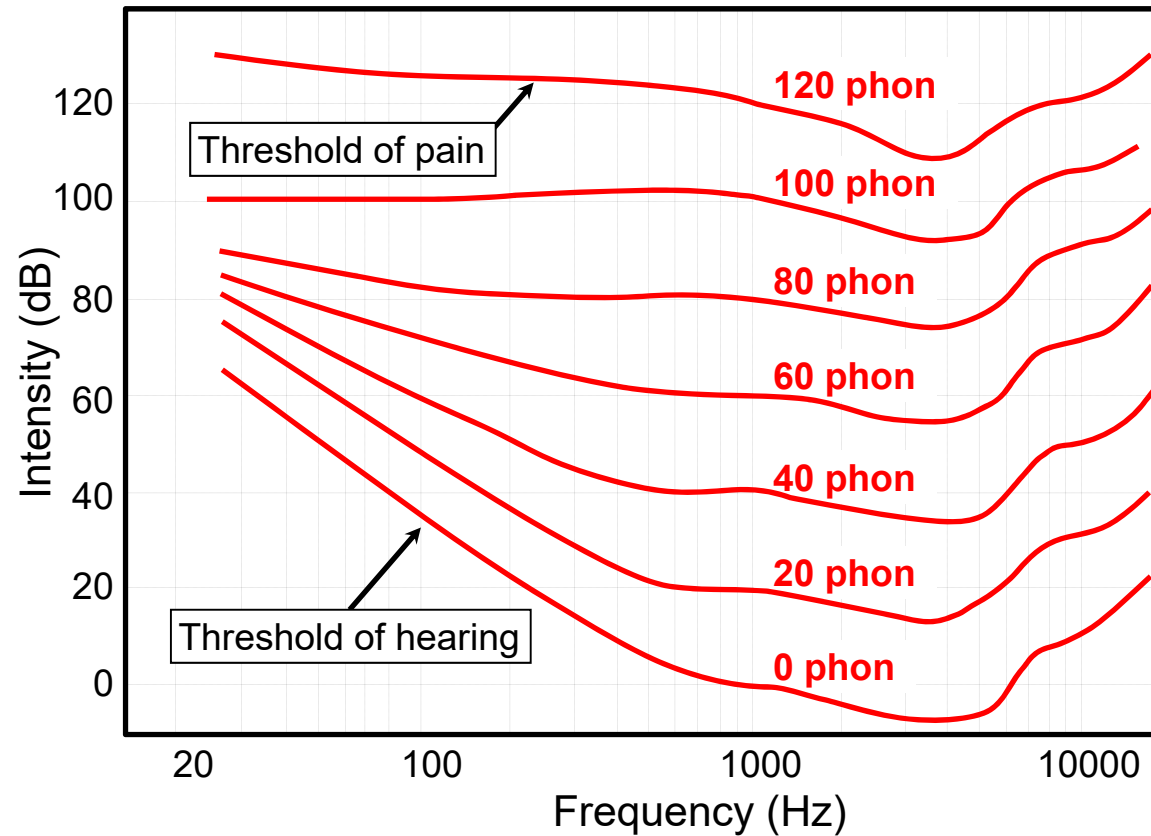
Equal-loudness contours (phon)



Audio Representation

Loudness

Equal-loudness contours (phon)



Audio Representation

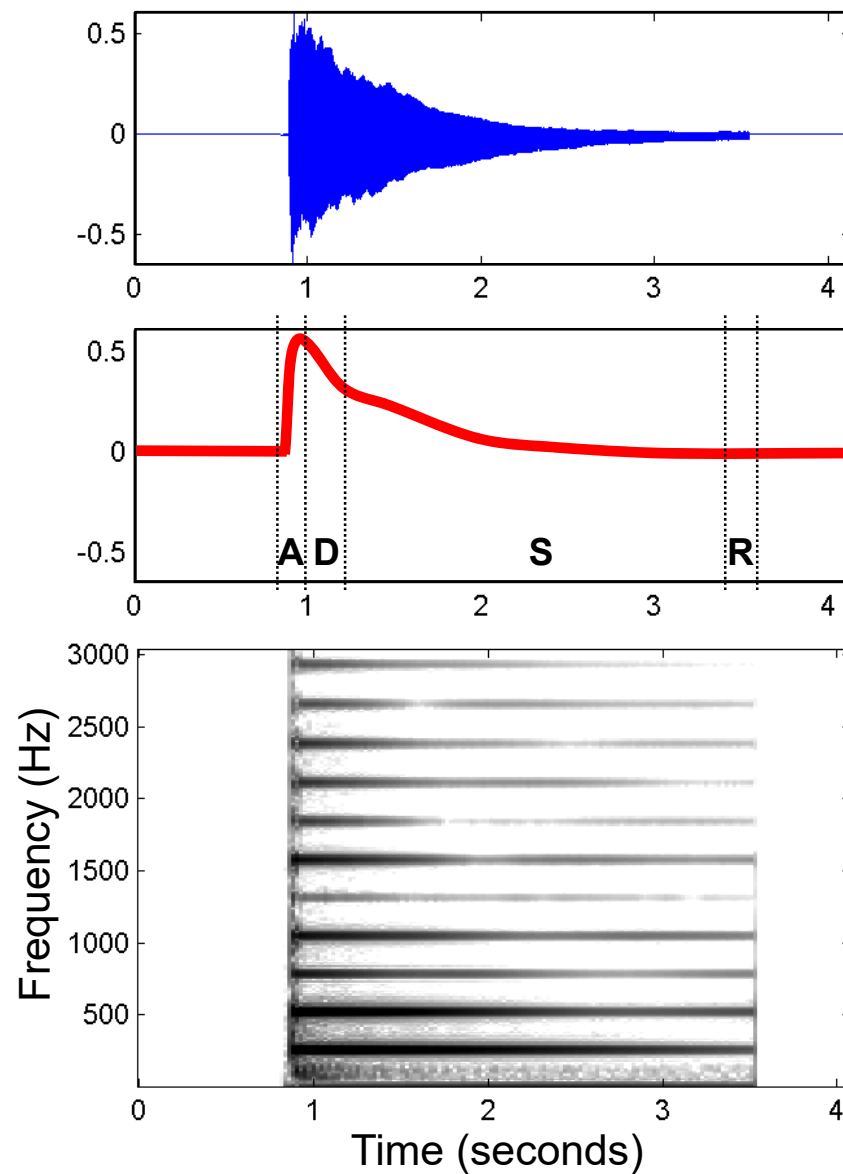
Timbre

- 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

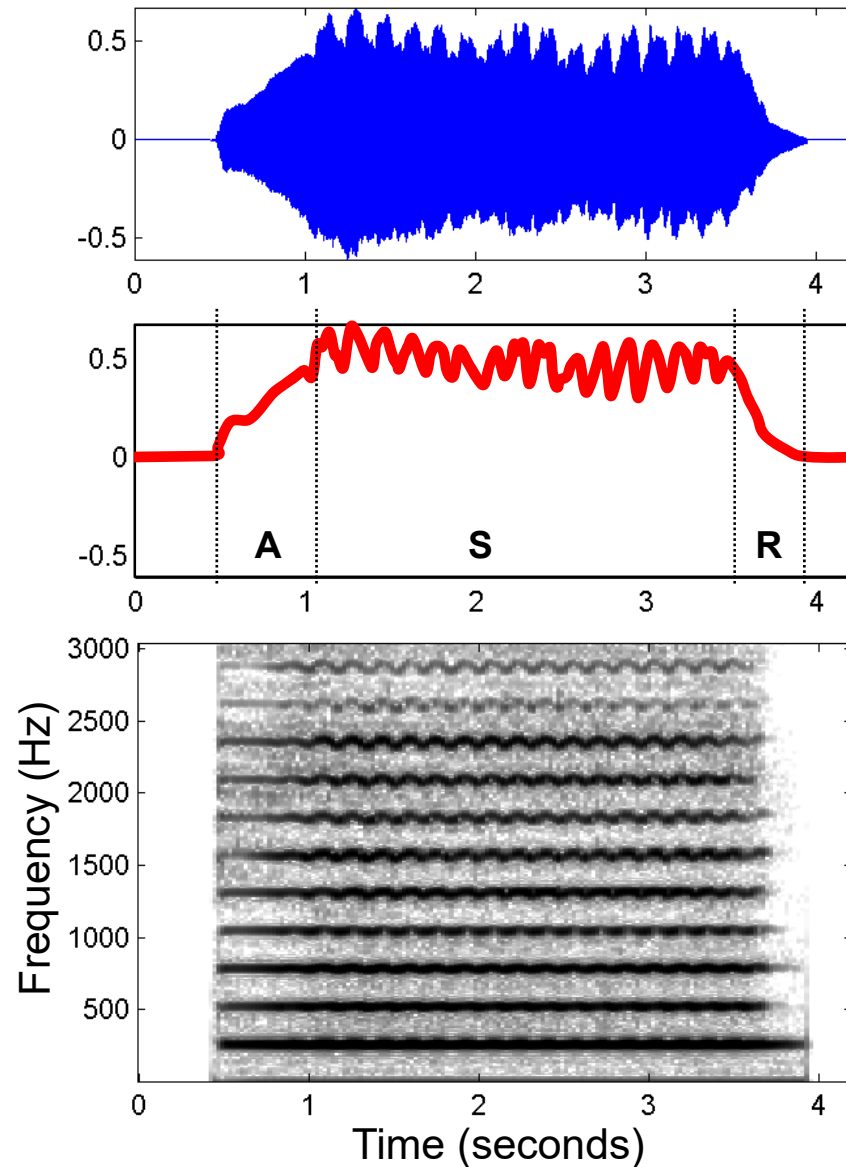
Timbre

Violine playing
note C4 (261.6 Hz)



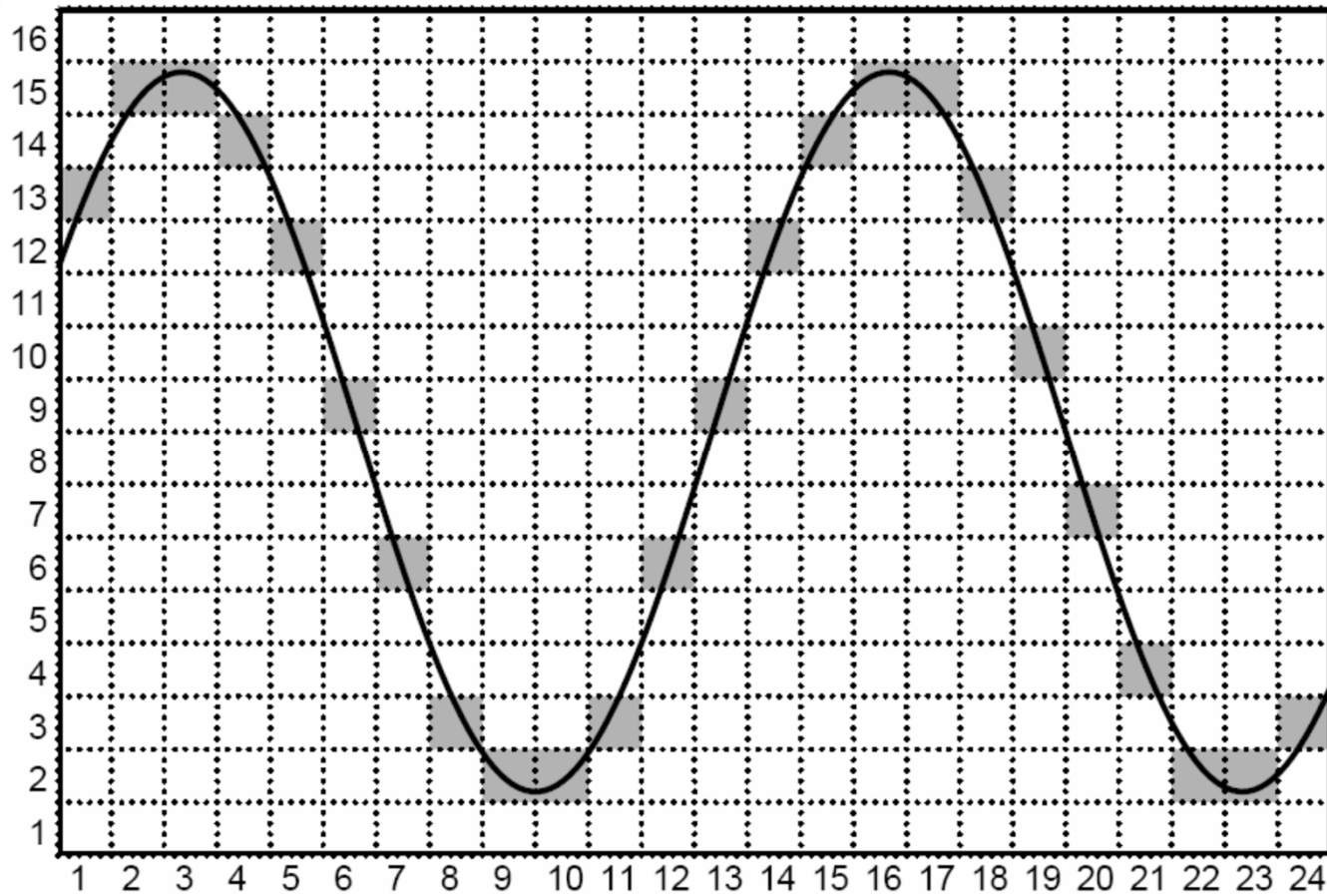
Vibrato:
Frequency modulations

Tremolo:
Amplitude modulations



Audio Representation

Digitization



Audio Representation

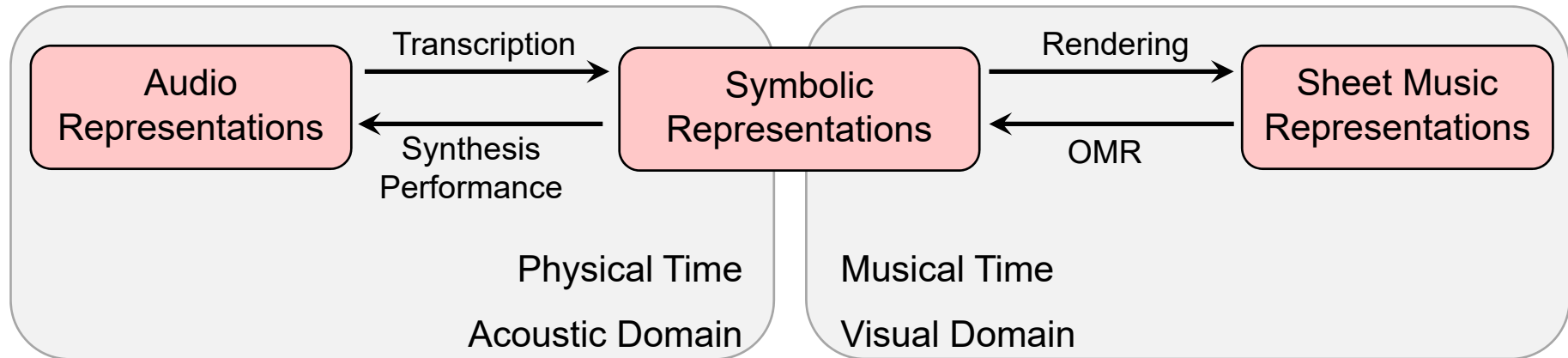
Digitization

- 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



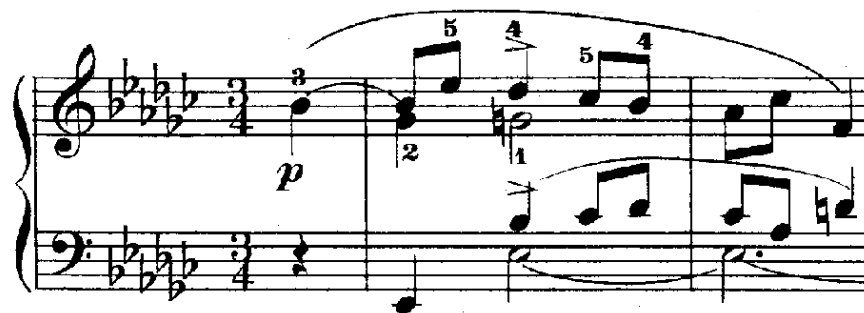
OMR = optical music recognition

Process of transforming sheet music into a symbolic representation

Music Representations

OMR

Original score



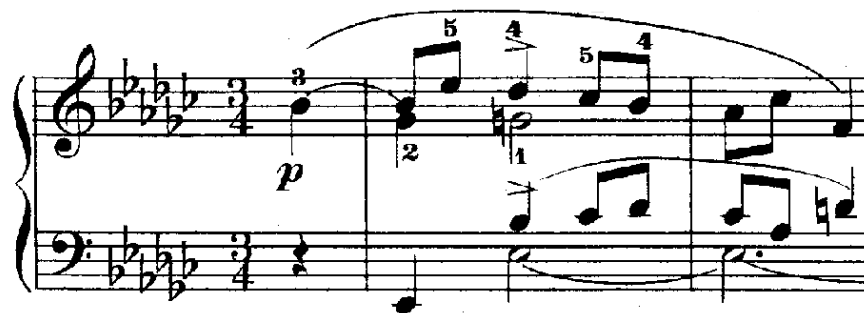
OMR score



Music Representations

OMR

Original score



OMR score



OMR errors