

Multimedia Systems

1. Introduction to SMC

António Sá Pinto

FEUP

Agenda

1. Introduction to SMC

1.0. What is SMC?

1.1. Audio Signal Fundamentals

1.2. Music Representations

1.3. Sound and Music Descriptors

— break (20 mins) —

Practical

Software Tools

Sonic Visualiser, Audacity and FMP Notebooks

What is Sound and Music Computing (SMC)?

Sound and Music Computing (SMC) research approaches the whole sound and music communication chain from a multidisciplinary point of view. By combining scientific, technological and artistic methodologies it aims at understanding, modelling and generating sound and music through computational approaches.

This definition is generally considered to include all types of sounds and human communication processes except speech. Speech research has its own aims and methodologies and is outside the SMC field.

from:[Sound and Music Computing Network \(2007\)](#)

What is Sound and Music Computing (SMC)?

Disciplines

- Music Composition
- Musicology
- Music Performance
- Physics/Acoustics
- Mathematics
- Psychology
- **Engineering**

from:[Sound and Music Computing Network \(2007\)](#)

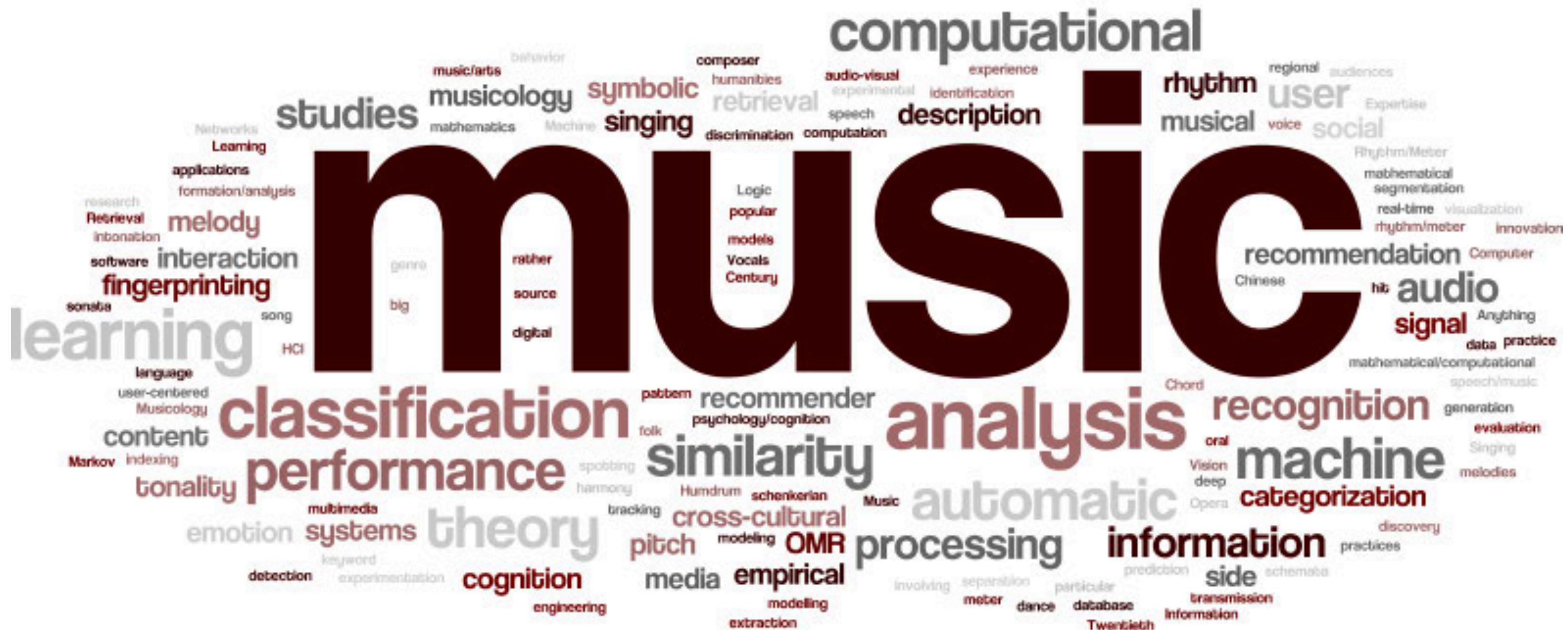
What is Sound and Music Computing (SMC)?

Areas of Application

- Digital Music Instruments
- Music Production
- **Music Information Retrieval**
- Digital Music Libraries
- Interactive Multimedia Systems
- Auditory Interfaces
- Augmented Action and Perception

from:[Sound and Music Computing Network \(2007\)](#)

What is Music Information Retrieval (MIR)?



from:Women in Music Information Retrieval

What is MIR?

Music Information Retrieval aims at extending the understanding and usefulness of music data, through the research, development and application of computational approaches and tools.

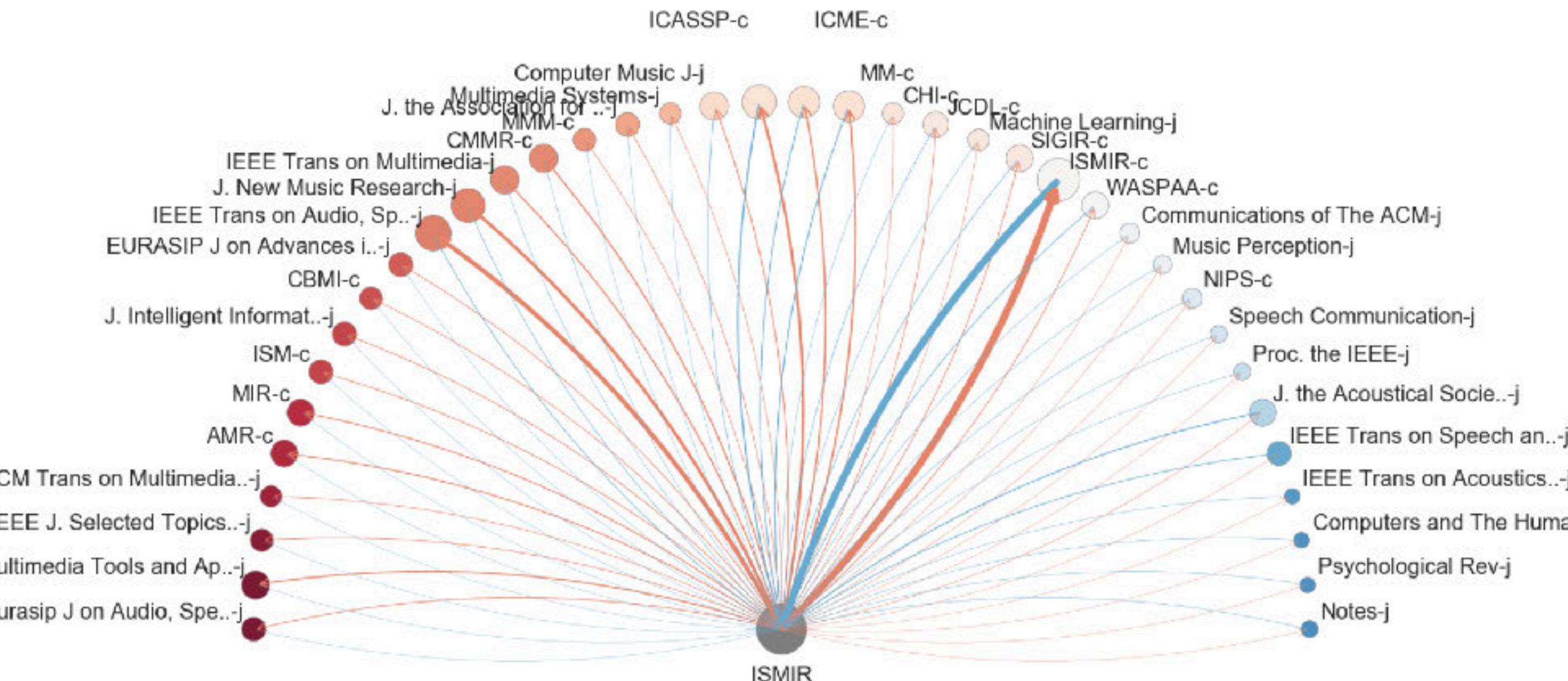
from: (Bello) [MIR Lecture Notes](#) - New York University

*MIR is the **interdisciplinary** science of retrieving information from music. Those involved in MIR may have a background in musicology, psychoacoustics, psychology, academic music study, signal processing, informatics, machine learning, optical music recognition, computational intelligence or some combination of these. MIR is a small but growing field of research with many real-world **applications**.*

from: Wikipedia

What is MIR?

- Main event: International Society for Music Information Retrieval conference (ISMIR)
- ISMIR has been running since 1999
- **Highly** multidisciplinary: Electronic Engineering, Library and Information Science, Computer Science, Music/Musicology, Psychology, Law
- Papers and sessions at other conferences: ICASSP, ICMC, SMC, DAFx, NIME, WASPAA, CMMR, AES, etc.



What is MIR?

Ever-changing: *Music Informatics Research*?

Feature Engineering —> Machine Learning

System-Centred —> User-Centred

Evaluation

What is MIR?

Music Information Retrieval Evaluation eXchange (MIREX) since 2008

- [2018:Audio Classification \(Train/Test\) Tasks](#)
 - Audio US Pop Genre Classification
 - Audio Latin Genre Classification
 - Audio Music Mood Classification
 - Audio Classical Composer Identification
- [2018:Audio K-POP Mood Classification](#)
- [2018:Audio K-POP Genre Classification](#)
- [2018:Audio Fingerprinting](#)
- [2018:Multiple Fundamental Frequency Estimation & Tracking](#)
- [2018:Set List Identification](#)
- [2018:Audio Melody Extraction](#)
- [2018:Audio Onset Detection](#)
- [2018:Audio Beat Tracking](#)
- [2018:Audio Key Detection](#)
- [2018:Audio Downbeat Estimation](#)
- [2018:Real-time Audio to Score Alignment \(a.k.a Score Following\)](#)
- [2018:Audio Cover Song Identification](#)
- [2018:Audio Chord Estimation](#)
- [2018:Automatic Lyrics-to-Audio Alignment](#)
- [2018:Drum Transcription](#)
- [2018:Patterns for Prediction](#)
- [2018:Audio Tempo Estimation](#)

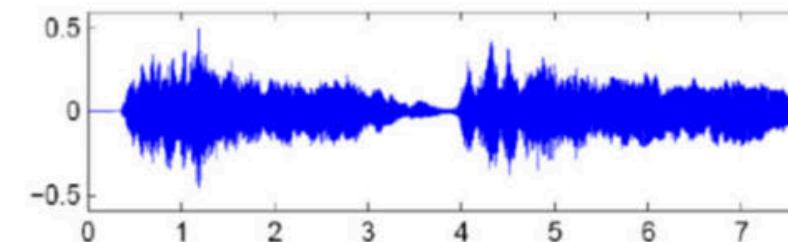
What is MIR?

Music Information

Sheet Music (Image)



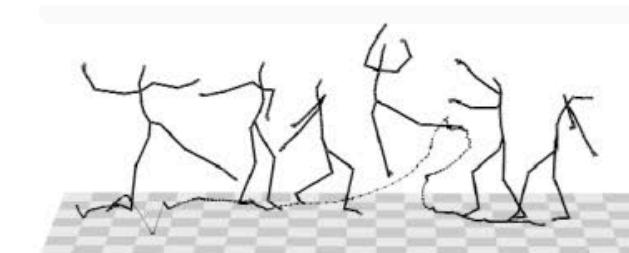
CD / MP3 (Audio)



MusicXML (Text)

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

Dance / Motion (Mocap)



Music

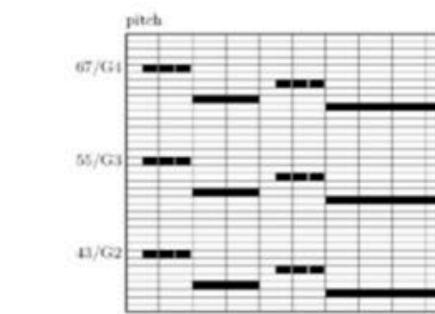
Singing / Voice (Audio)



Music Film (Video)



MIDI

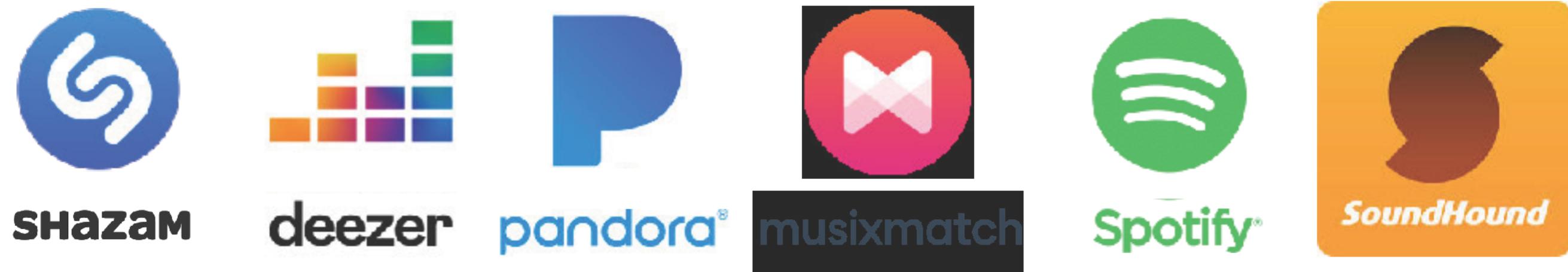


Music Literature (Text)



What is MIR?

Applications



Music fingerprinting/melody identification: *Shazam, Gracenote, SoundHound*

Music Recommendation and Playlist generation: *YouTube, Google, Last.fm, Pandora, Spotify*

Score Following: *SmartMusic, RockBand*

Music Interaction: *Chordify*

Pro Tools: *Melodyne, Autotune*

Others: *Smule, Native Instruments, ROLI, Steinberg*

What is MIR?

Uses

listen

allow browsing, discovery, recommendation

produce

enhance music production intelligently

educate

support music students interactively

create

generate new content

Multimedia Systems

MIR Components

Basics

- Sound and Music Fundamentals
- Music Representations
- Sound and Music Descriptors

Musical Elements

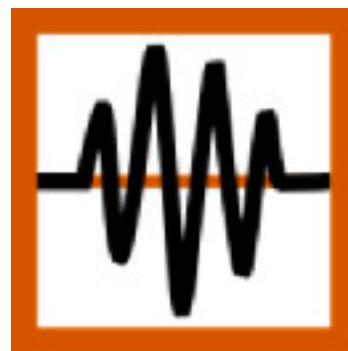
- Rhythm, Structure
- Melody, Harmony, Timbre

Applications

- Content-based audio retrieval
- Source Separation
- Music Identification
- Automatic Mashup

Multimedia Systems

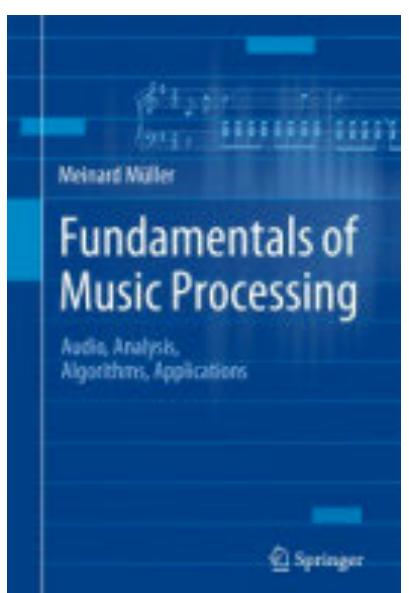
MIR Software Tools



Sonic Visualiser



Audacity



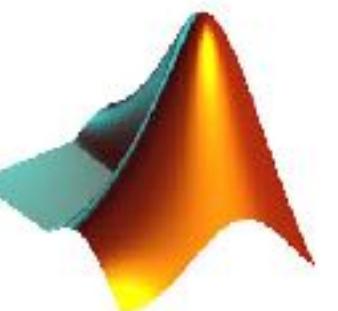
Meinard Müller

Fundamentals of Music Processing

Audio, Analysis, Algorithms, Applications

ISBN: 978-3-319-21944-8

Springer, 2015

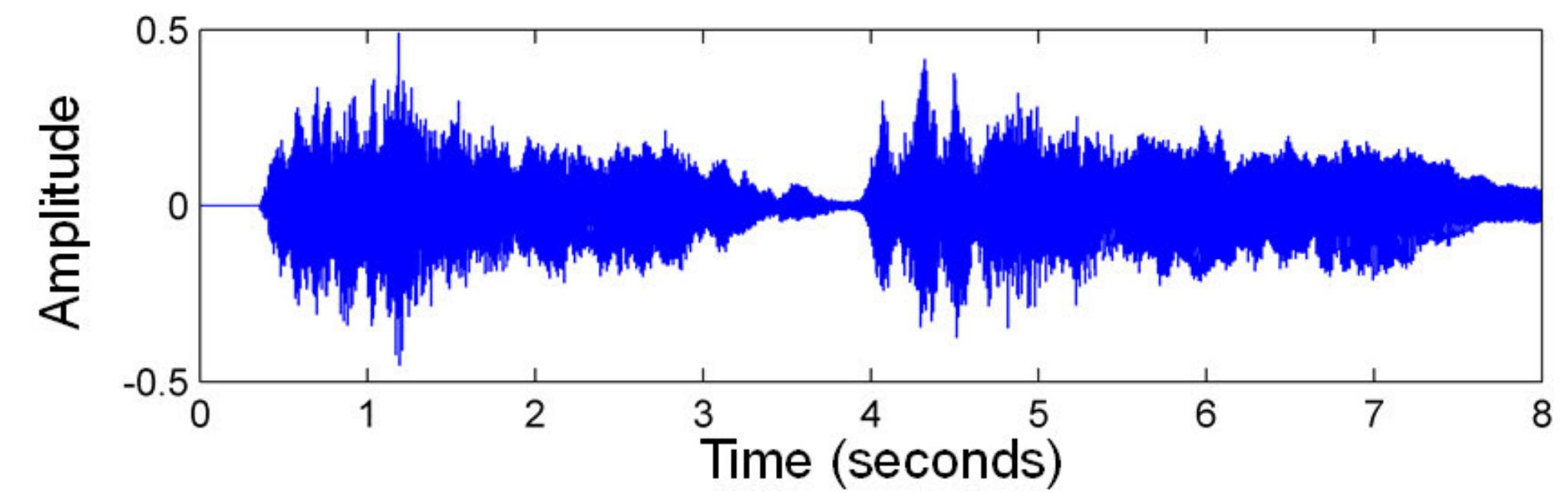
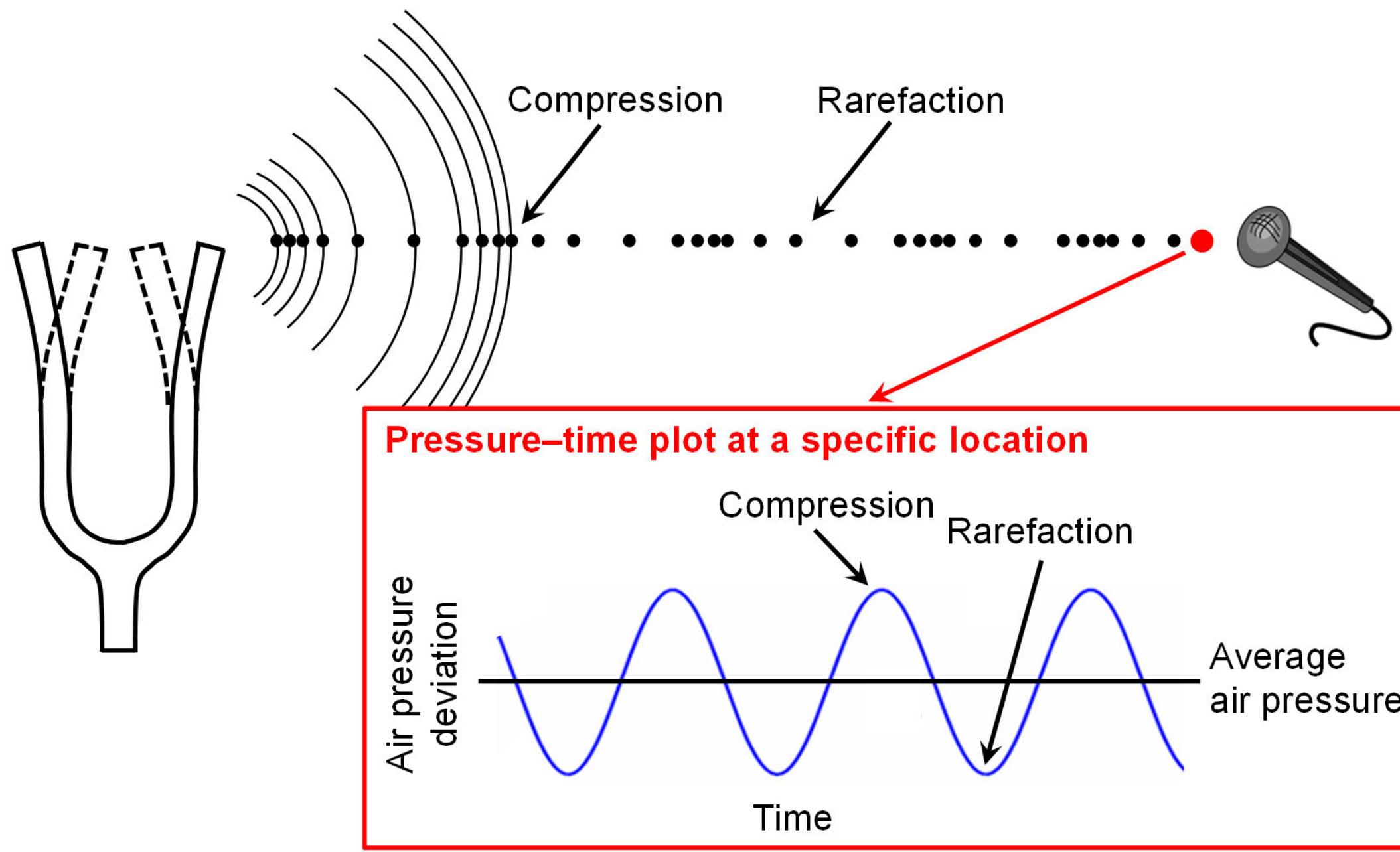


Matlab

Chapter	Title
B	Basics
0	Overview
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

Audio Signal Fundamentals

Physics of sound



from: (Müller) Fundamentals of Music Processing, Springer 2015

Audio Signal Fundamentals

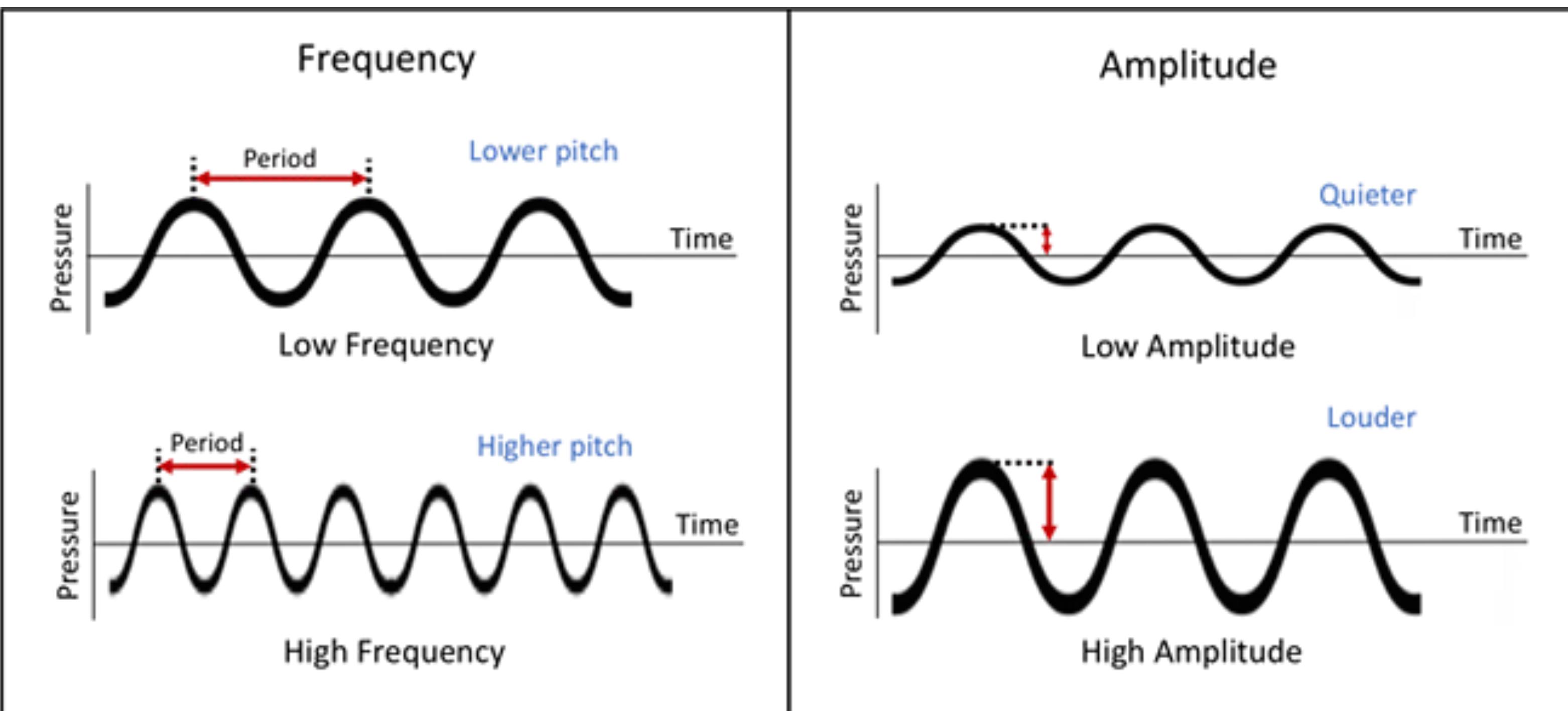
Physics of sound

$$v = f\lambda$$

v : speed ($m.s^{-1}$)
 λ : wavelength (m)
 f : frequency (Hz or s^{-1})

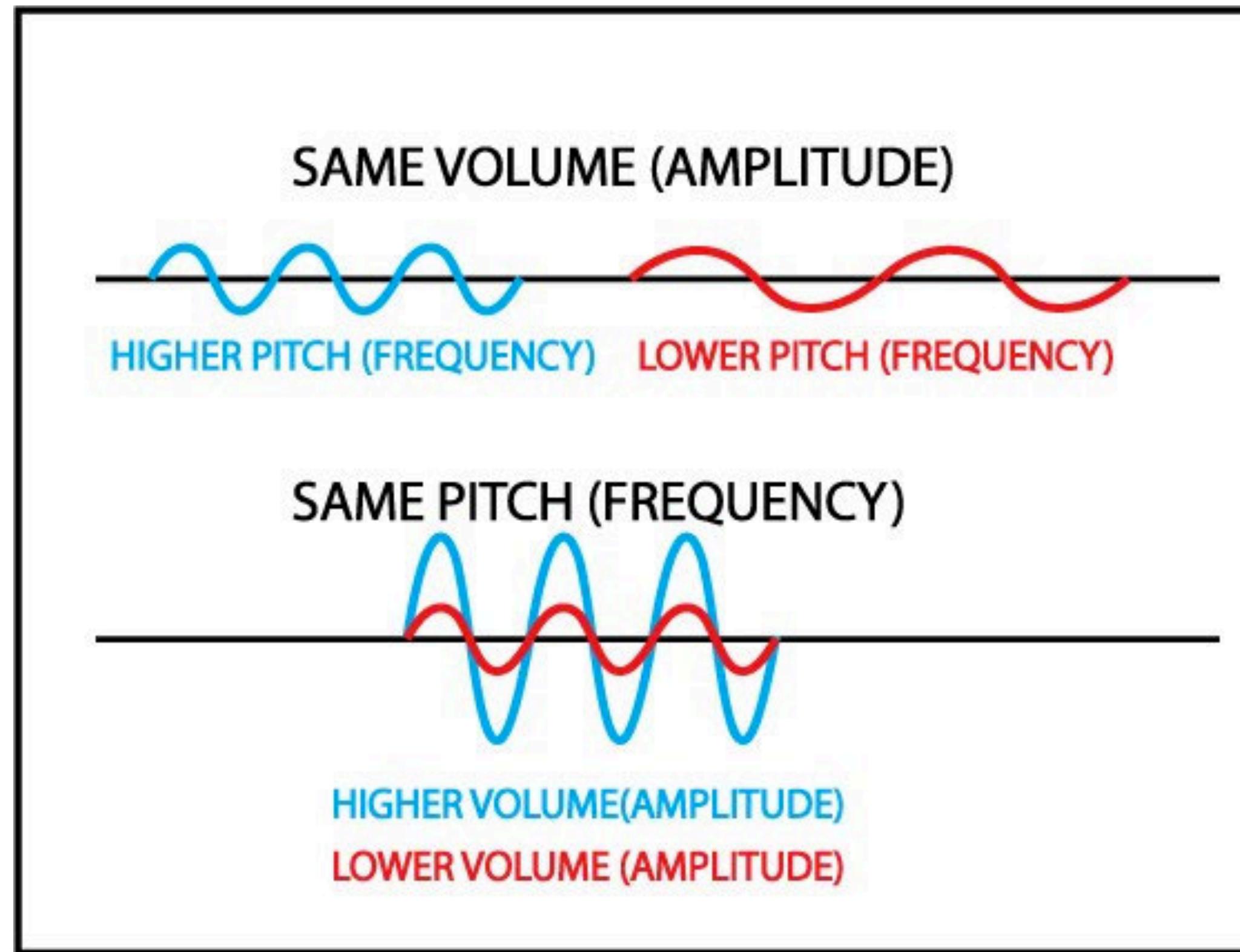
$$v = \frac{\lambda}{T}$$

v : speed ($m.s^{-1}$)
 λ : wavelength (m)
 T : period (s)



Audio Signal Fundamentals

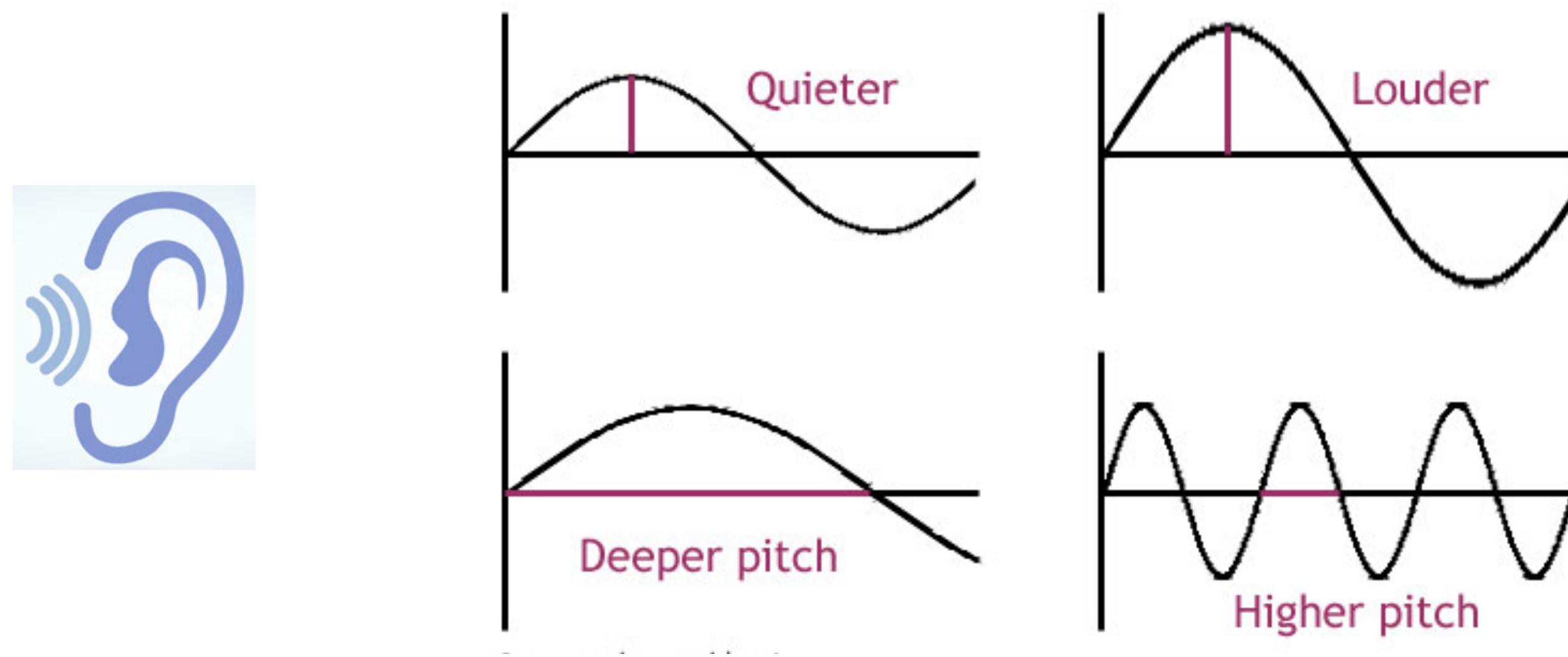
Physics of sound



!!

Audio Signal Fundamentals

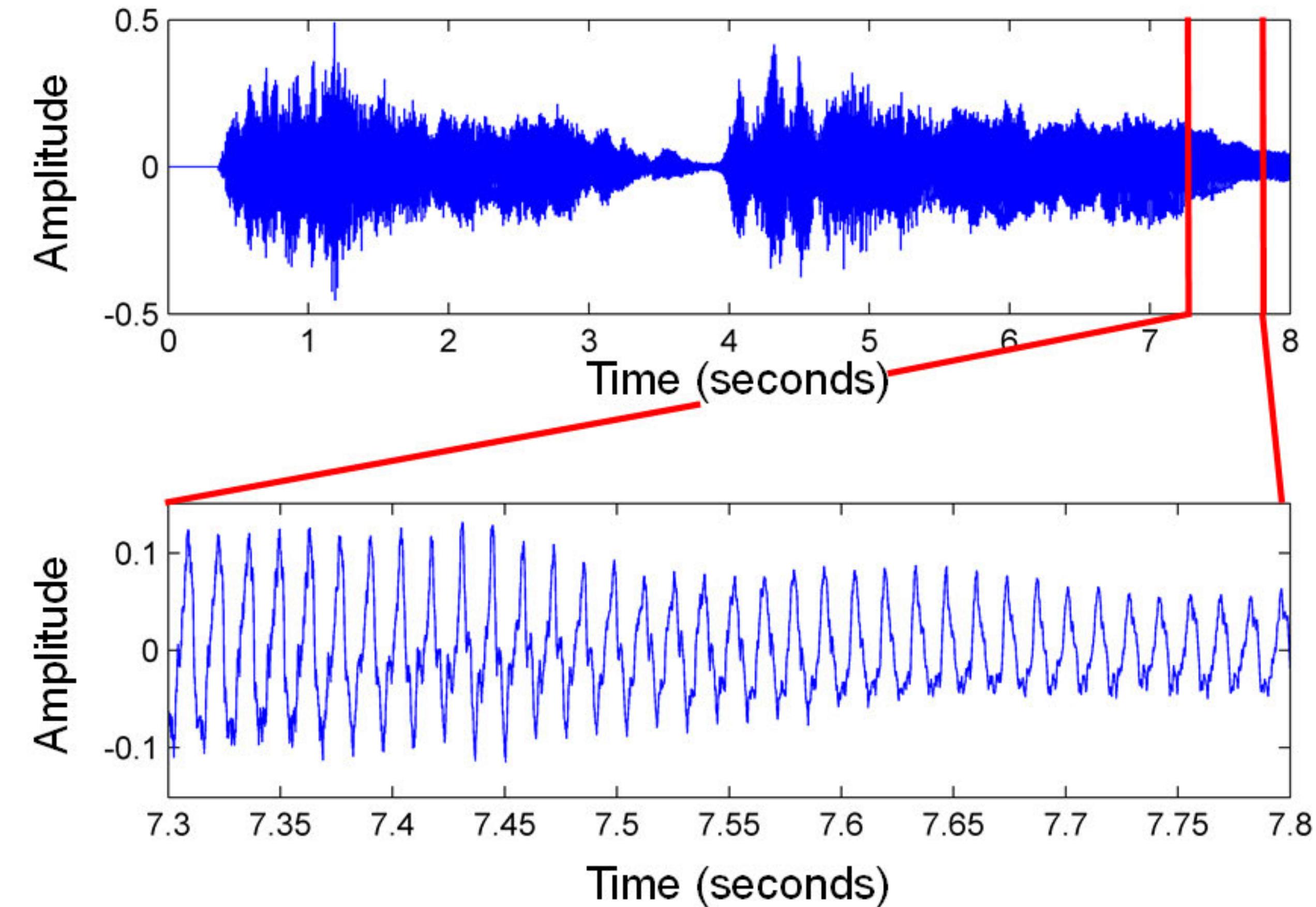
Physics of sound



What are the units/axis of interest?

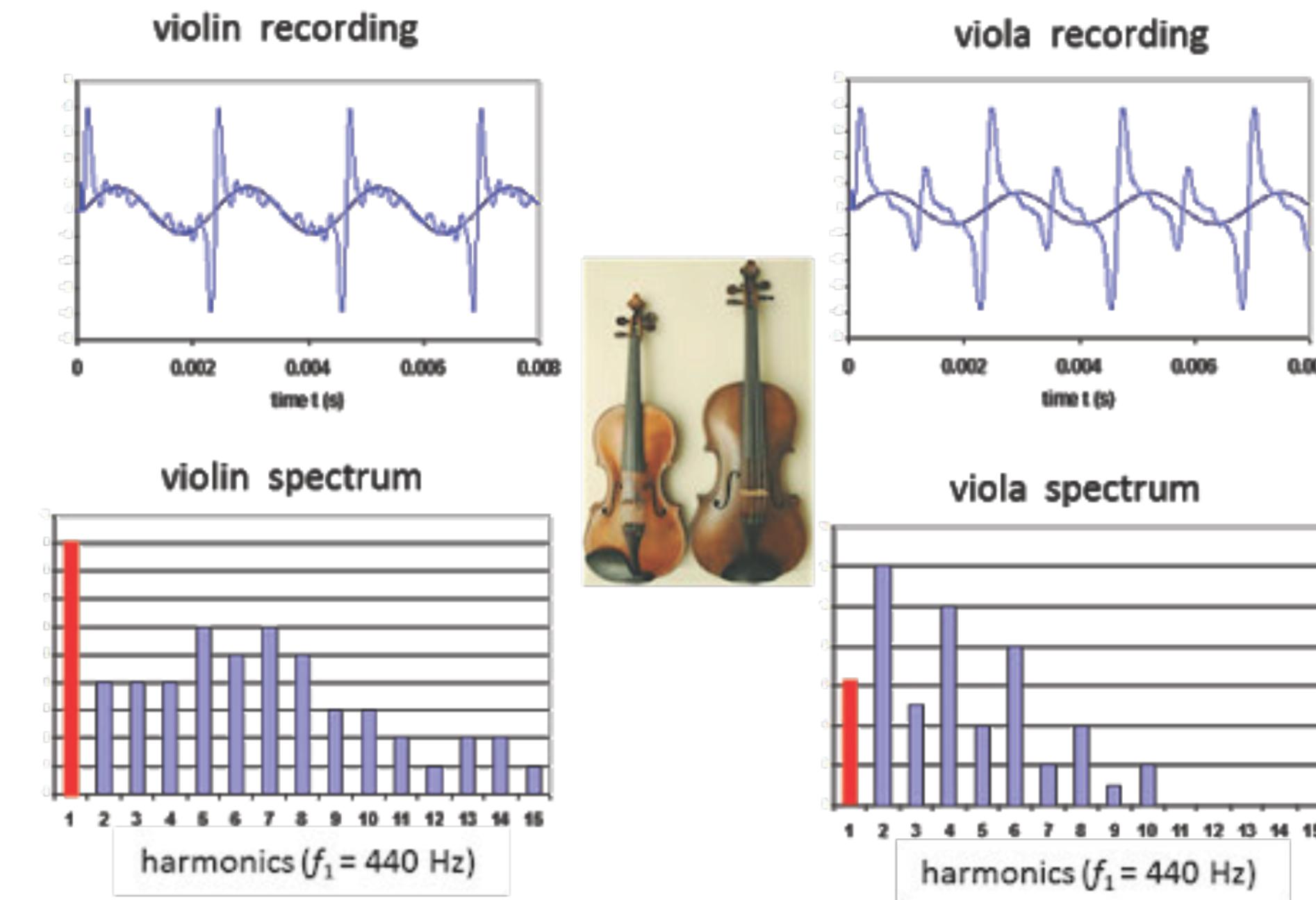
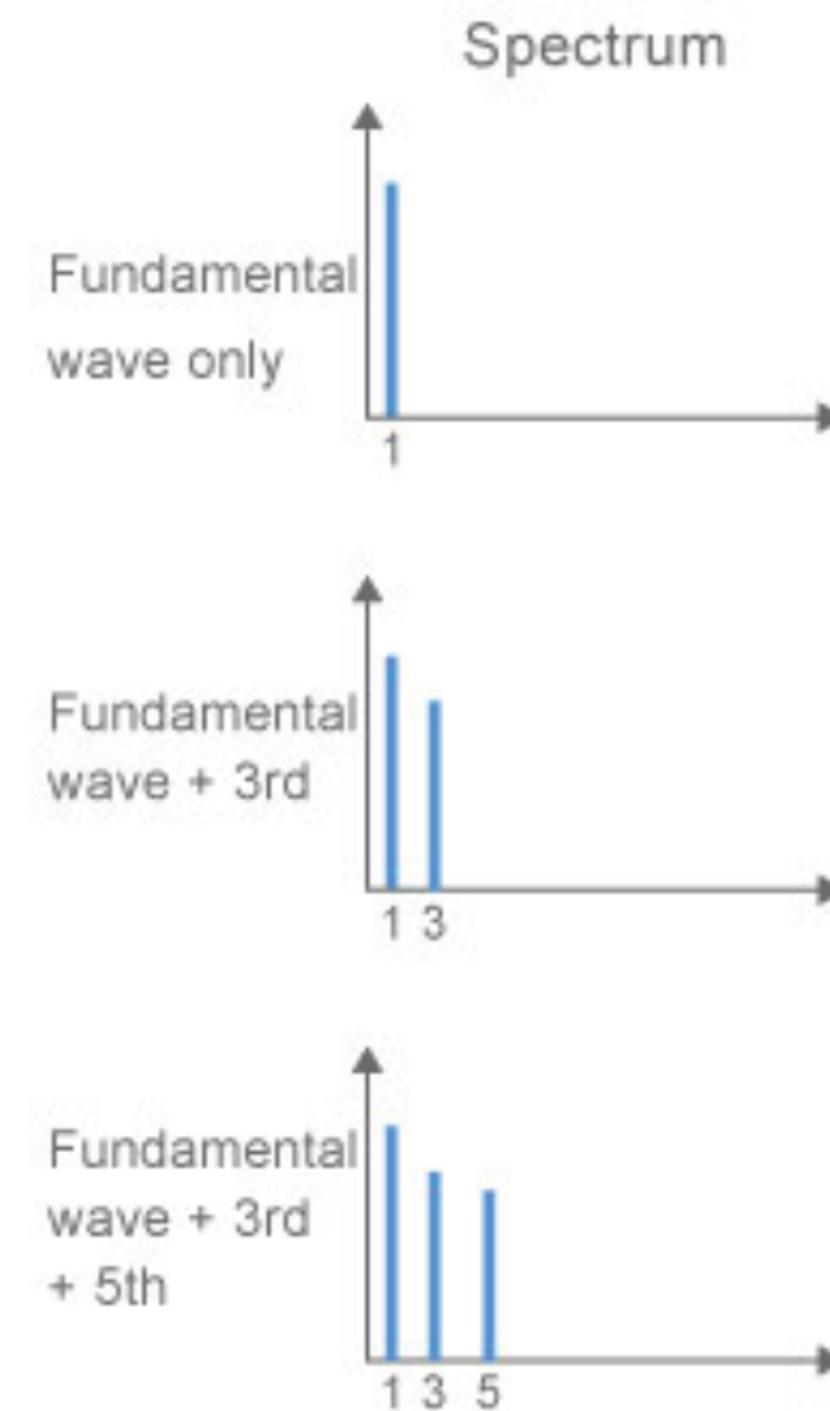
Audio Signal Fundamentals

Waveform



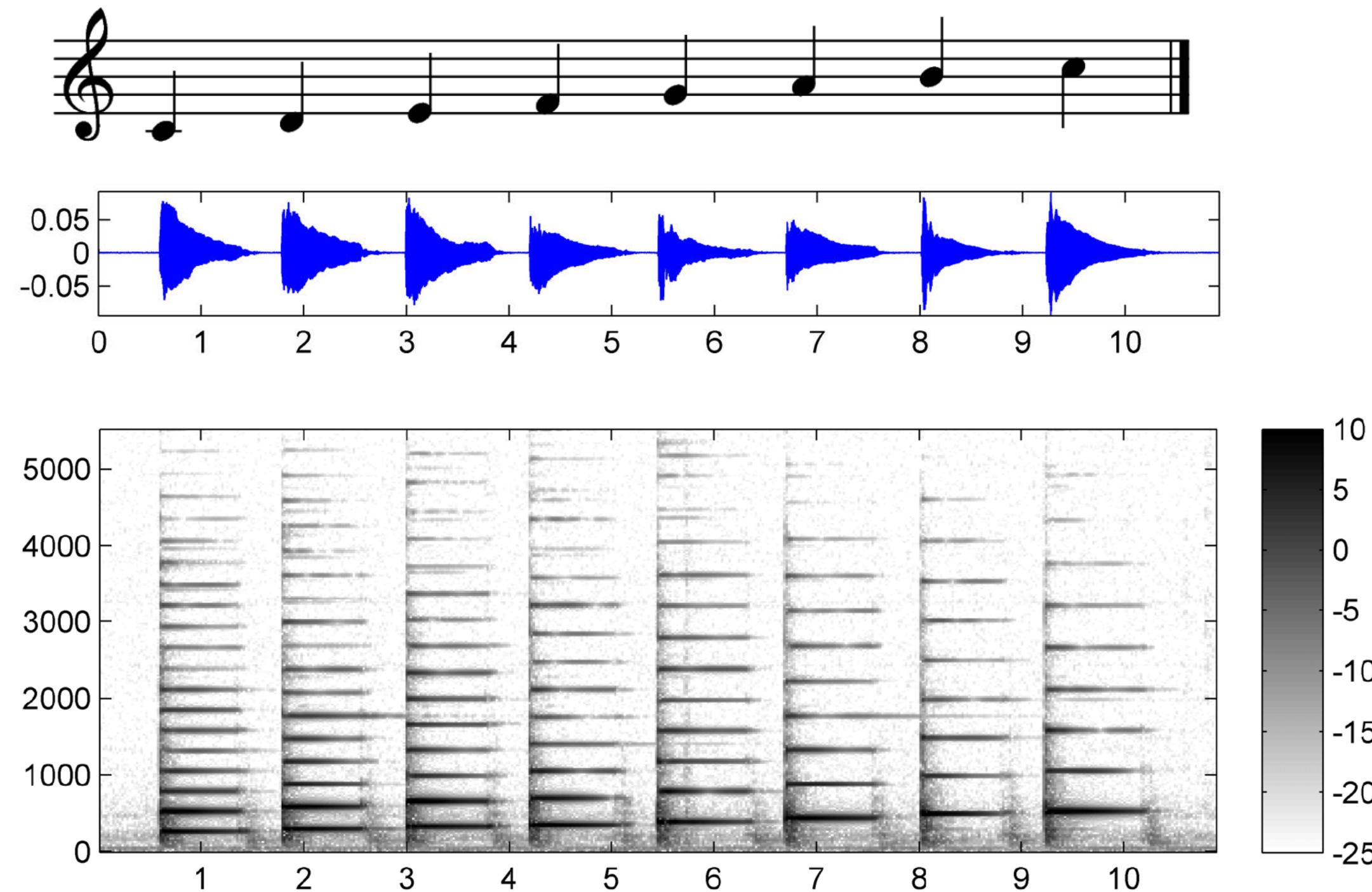
Audio Signal Fundamentals

Spectrum



Audio Signal Fundamentals

Spectrogram



from: (Müller) Fundamentals of Music Processing, Springer 2015

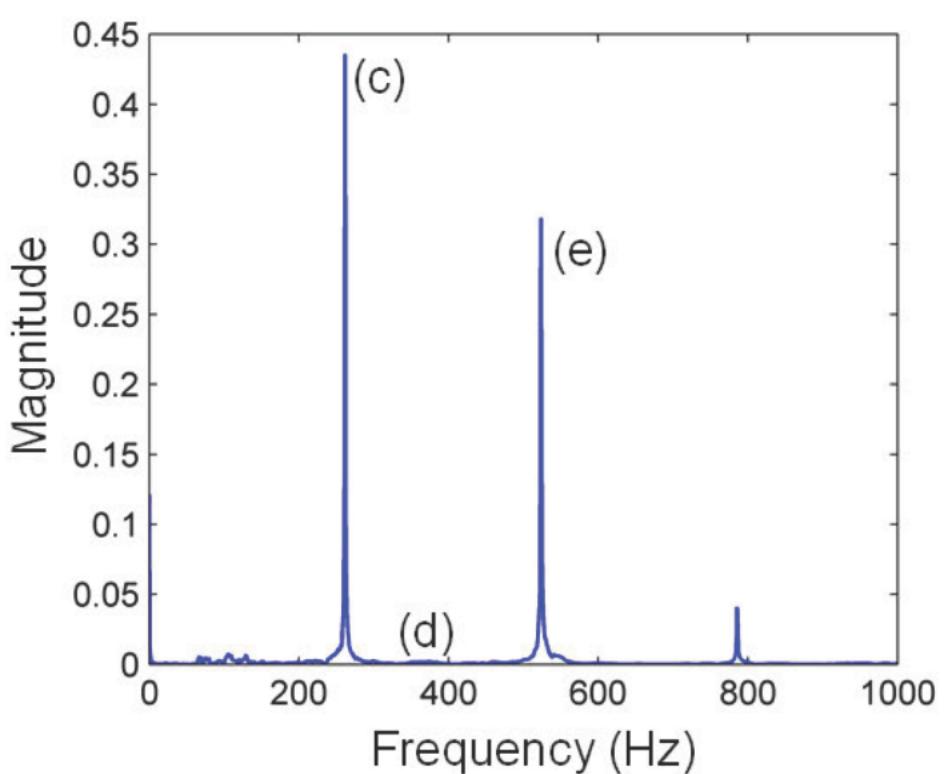
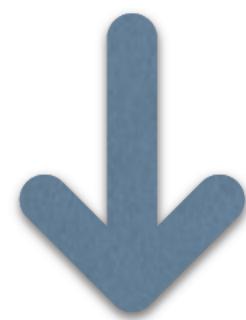
Audio Signal Fundamentals

The Fourier Transform

FFT (Analysis)

$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-i2\pi kn/N}$$

from **time** to **frequency**



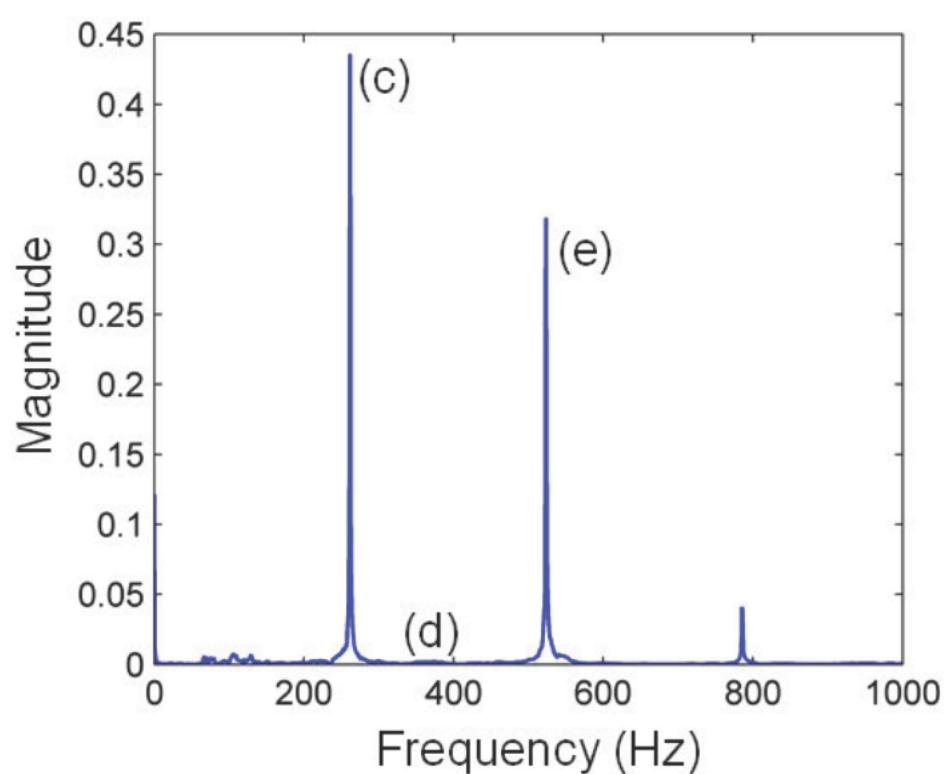
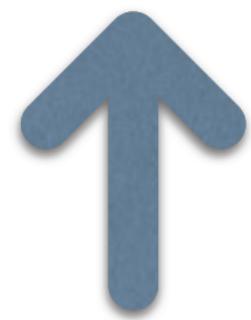
Audio Signal Fundamentals

The Fourier Transform

IFFT (Synthesis)

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k \cdot e^{i2\pi kn/N}$$

from **frequency** to **time**



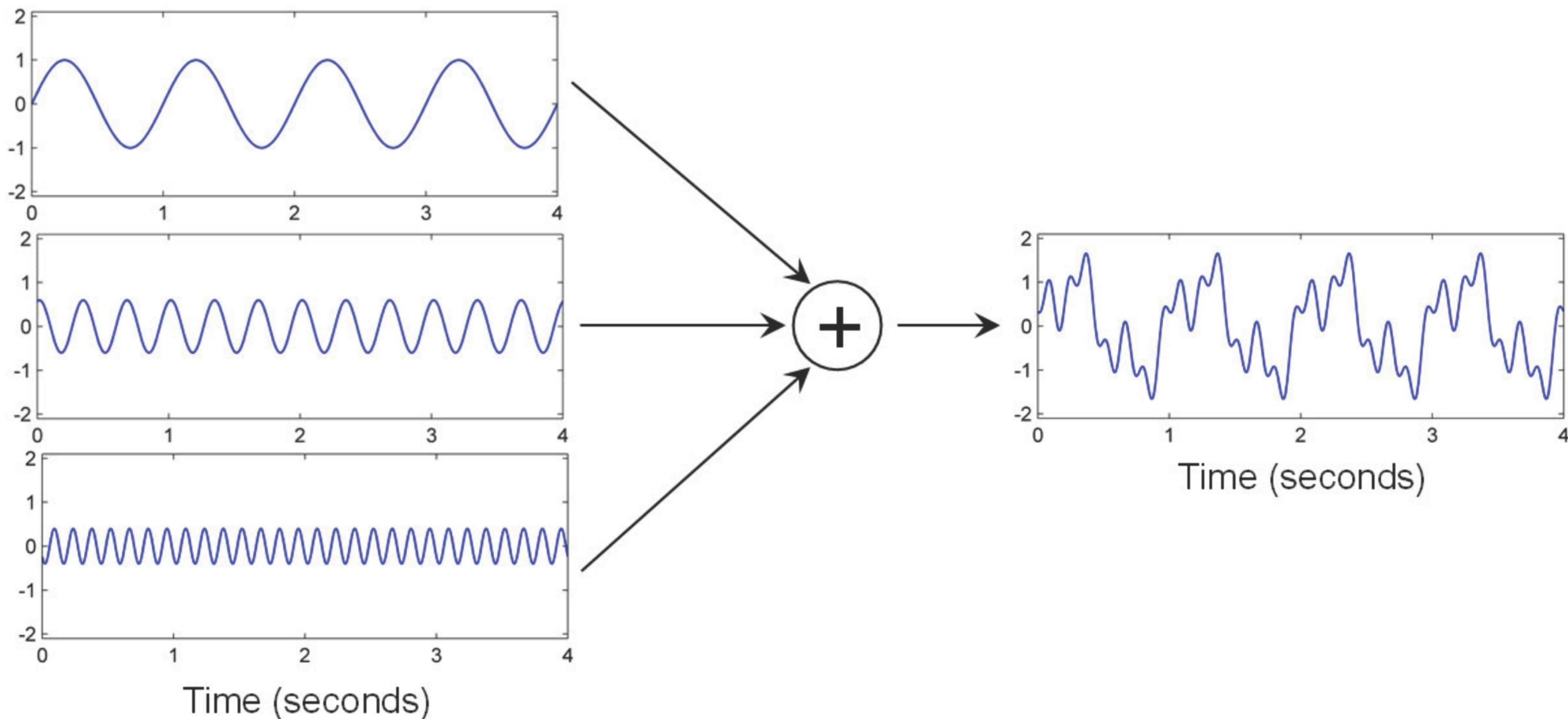
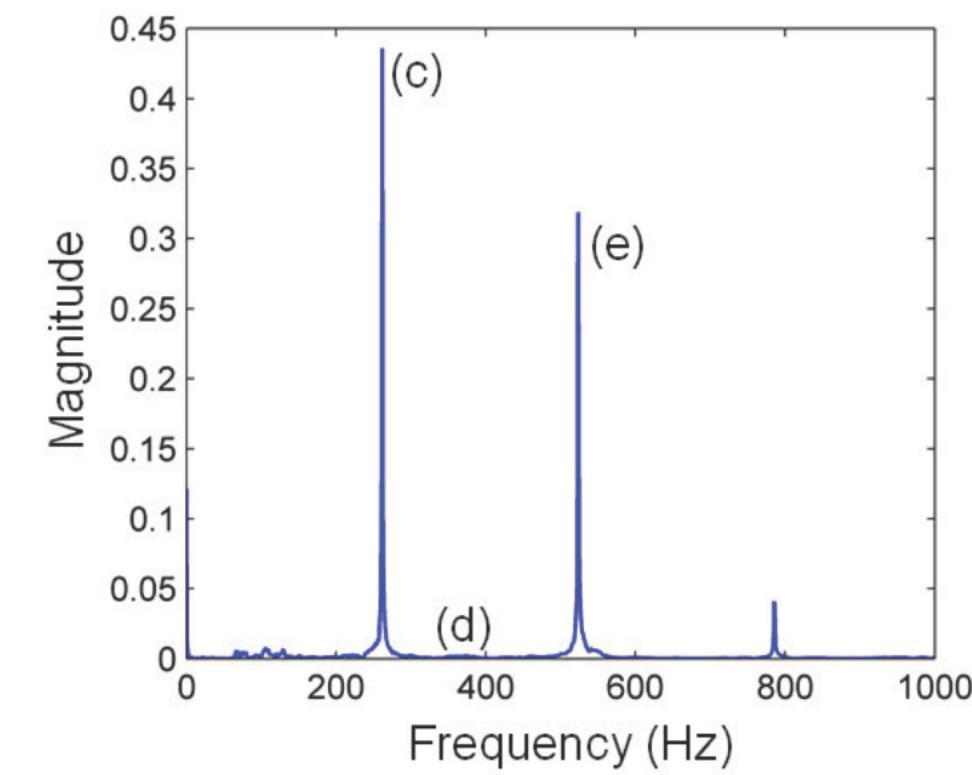
Audio Signal Fundamentals

The Fourier Transform

IFFT (Synthesis)

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k \cdot e^{i2\pi kn/N}$$

from **frequency** to **time**

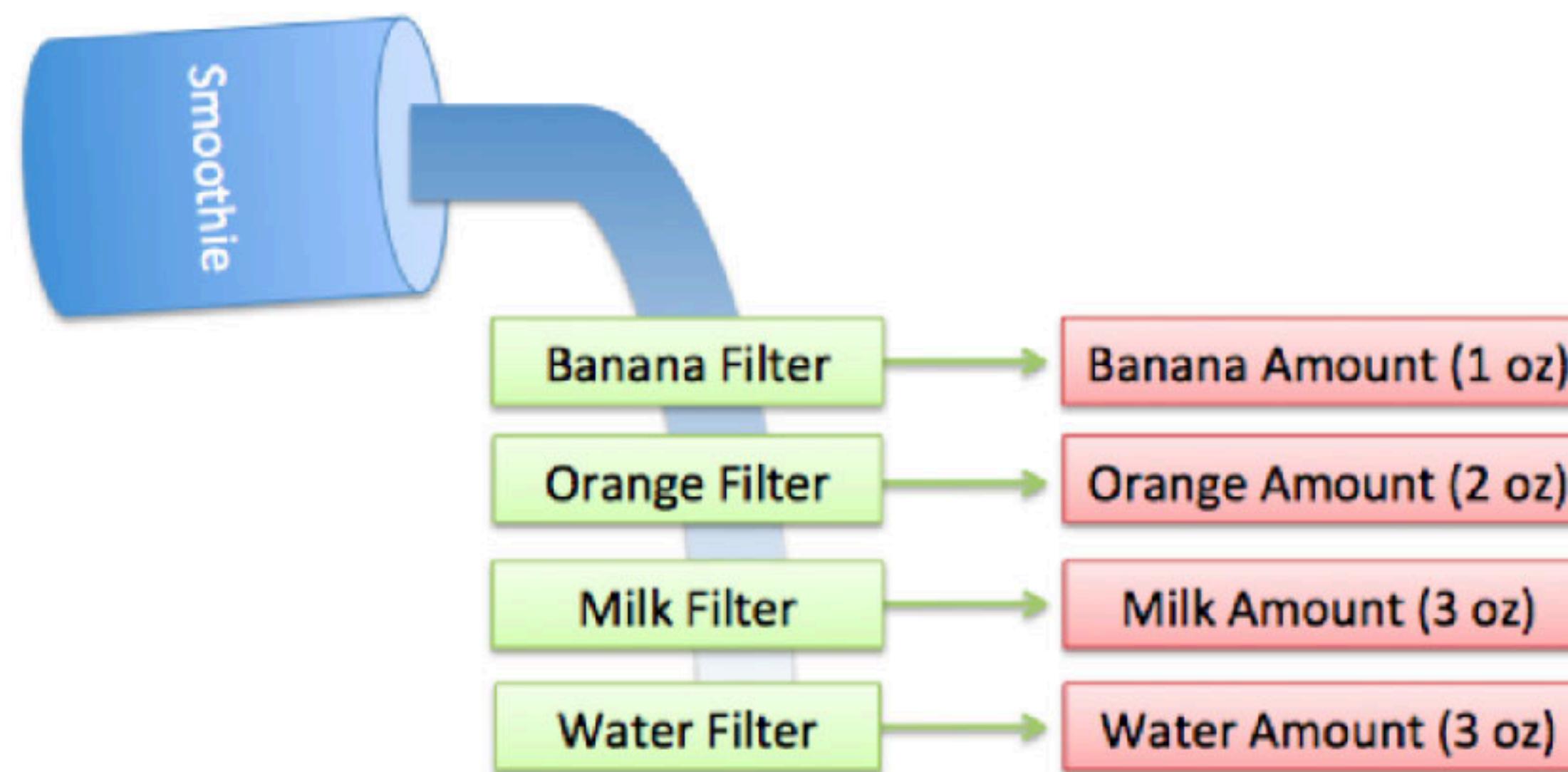


Audio Signal Fundamentals

The Fourier Transform

FFT (**Analysis**)

from **smoothie** to **recipe**

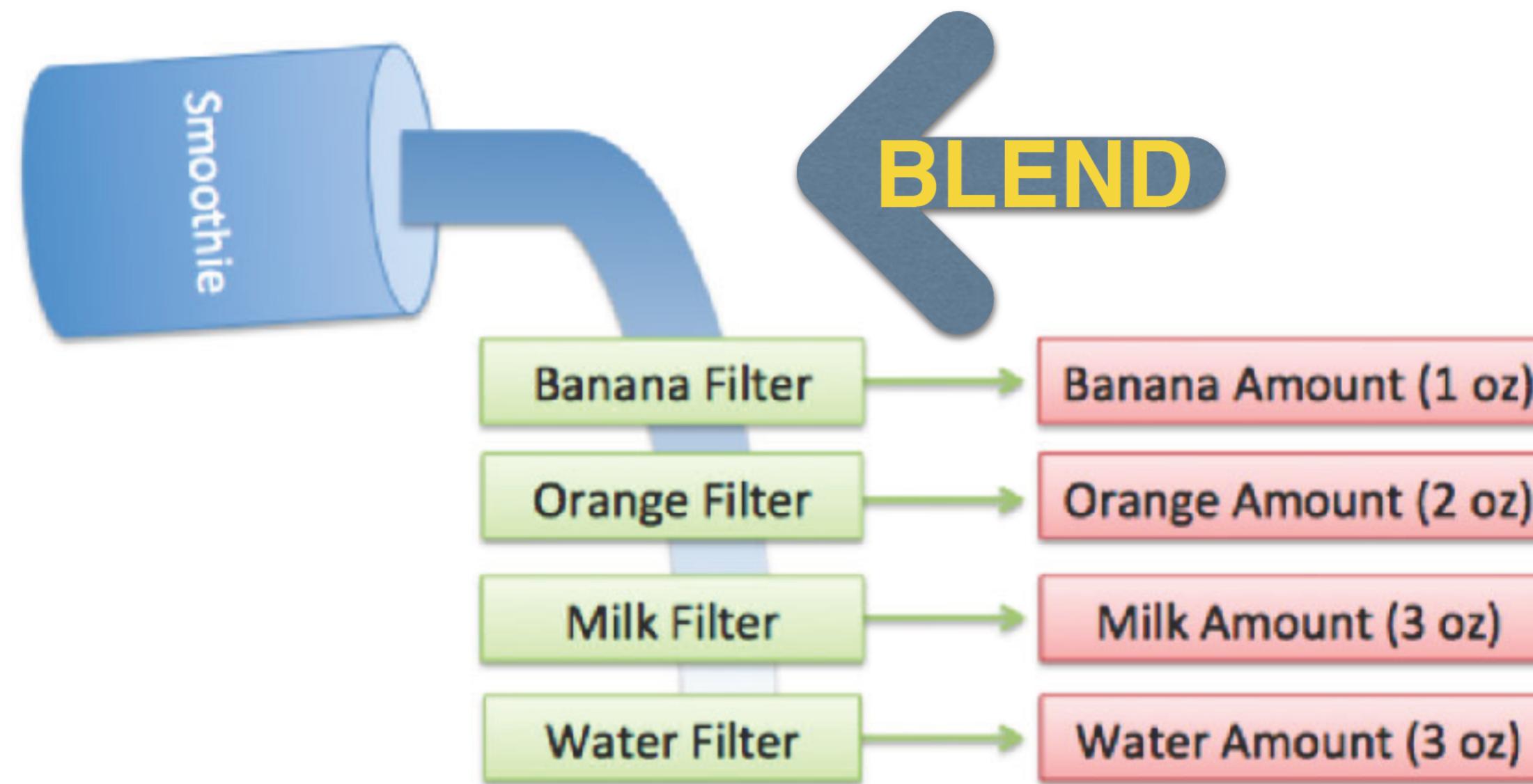


Audio Signal Fundamentals

The Fourier Transform

FFT (Synthesis)

from **recipe** to **smoothie**



Audio Signal Fundamentals

The Fourier Transform

FFT (Analysis)

$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-i2\pi kn/N}$$

from **time** to **frequency**

IFFT (Synthesis)

$$x_n = \frac{1}{N} \sum_{k=0}^{N-1} X_k \cdot e^{i2\pi kn/N}$$

from **frequency** to **time**

Recommended Resources

[BetterExplained: An interactive guide to the Fourier transform \(web\)](#)

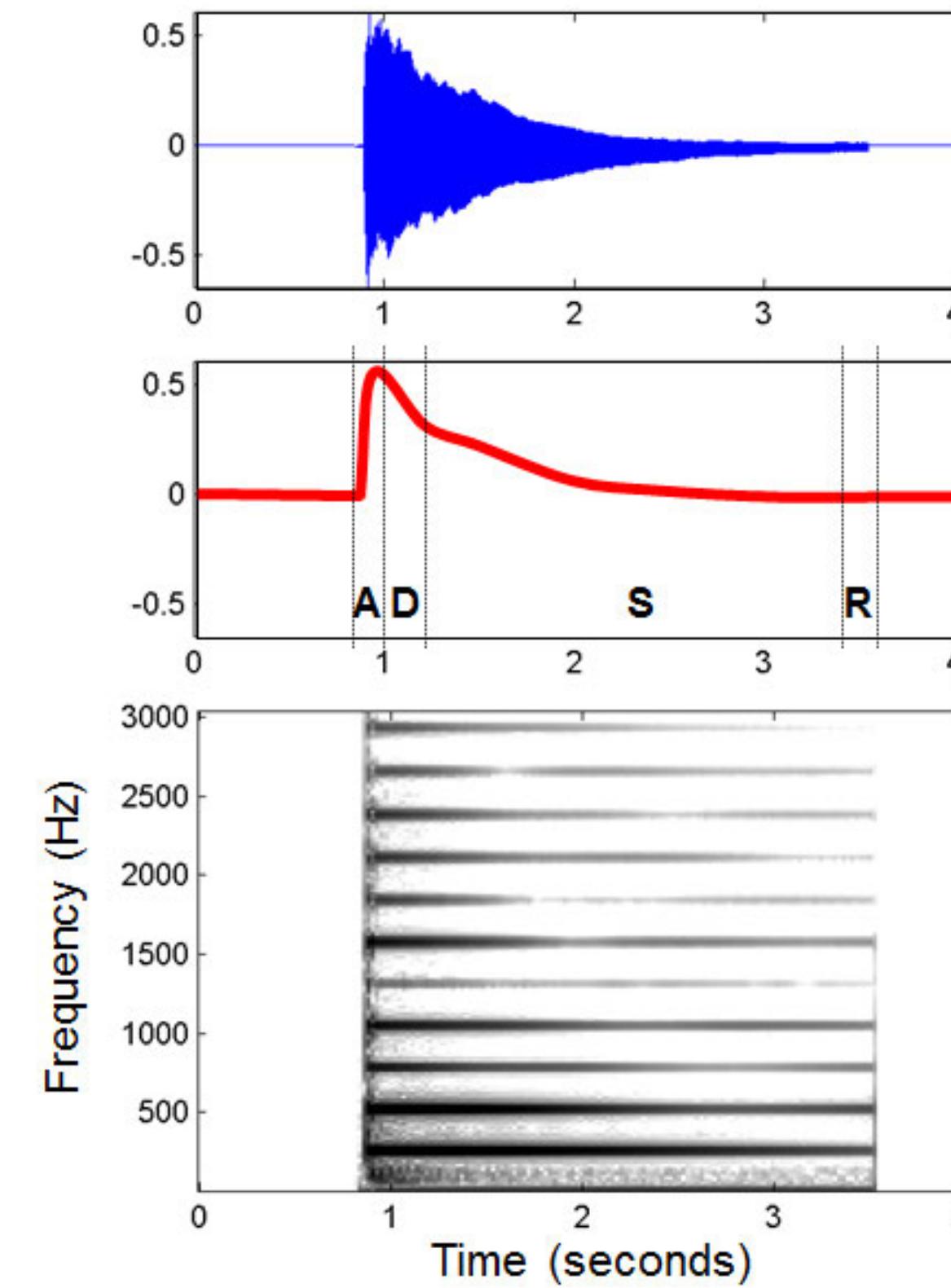
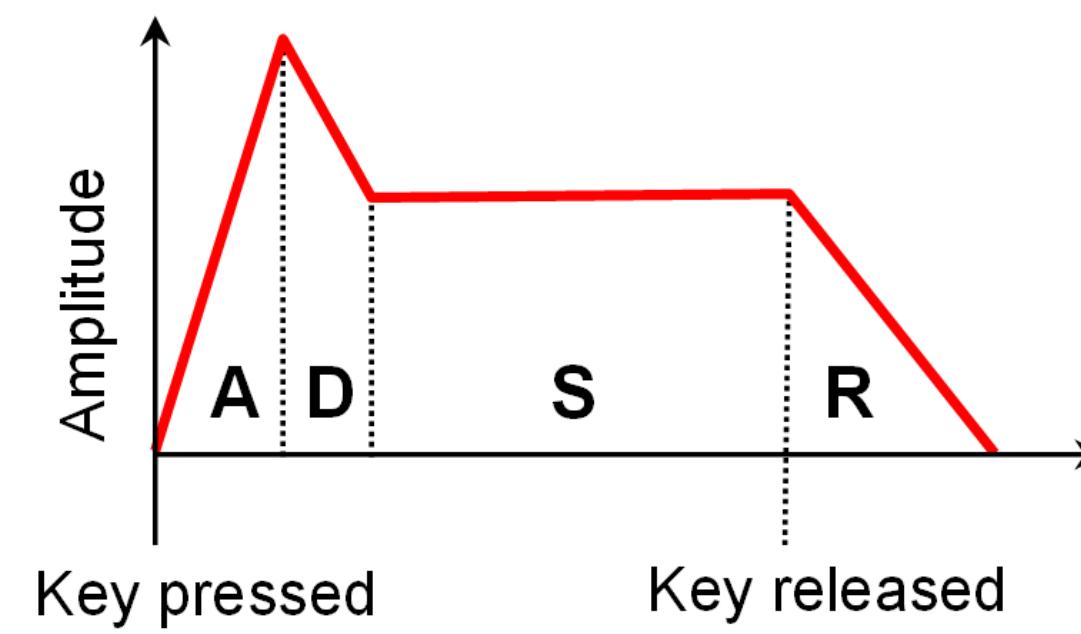
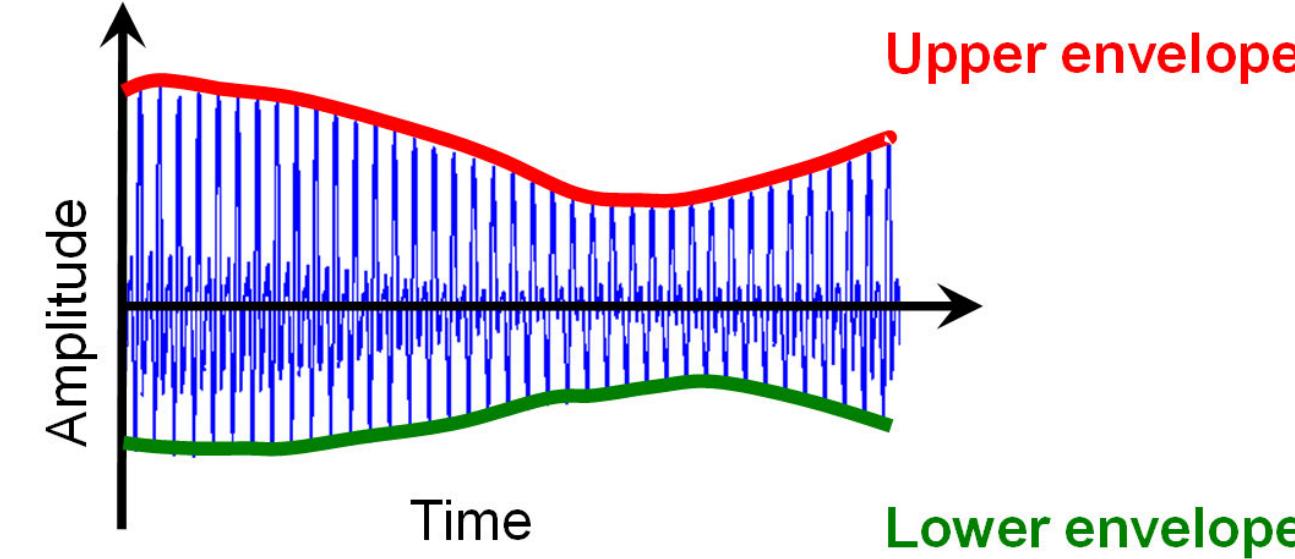
[An Interactive Introduction to Fourier Transforms \(web\)](#)

[A Visual Introduction to the Fourier Transform \(youtube\)](#)

[The Harmonic Analyzer \(Mechanical Machine\)](#)

Sound and Music Representations

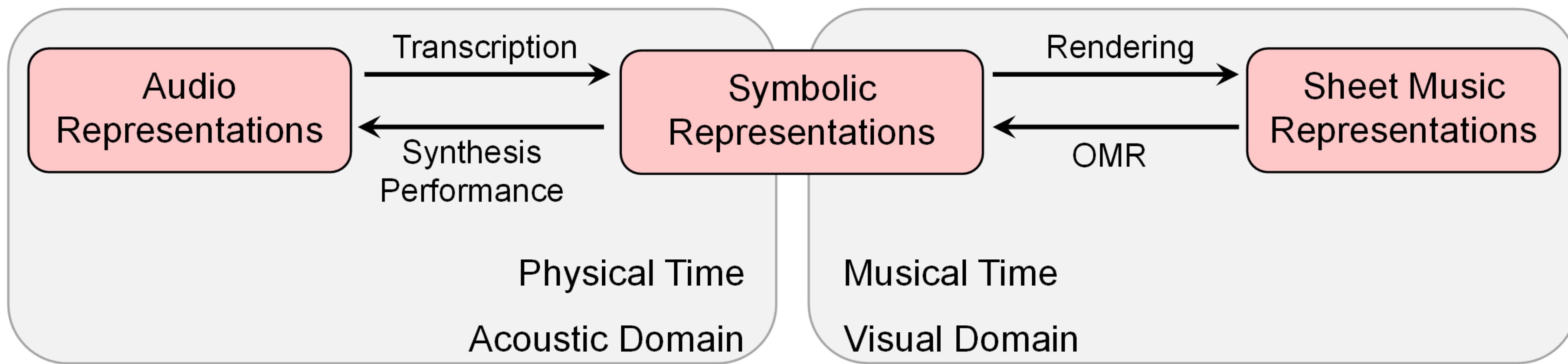
Audio



from: (Müller) Fundamentals of Music Processing, Springer 2015

Sound and Music Representations

Audio / Symbolic



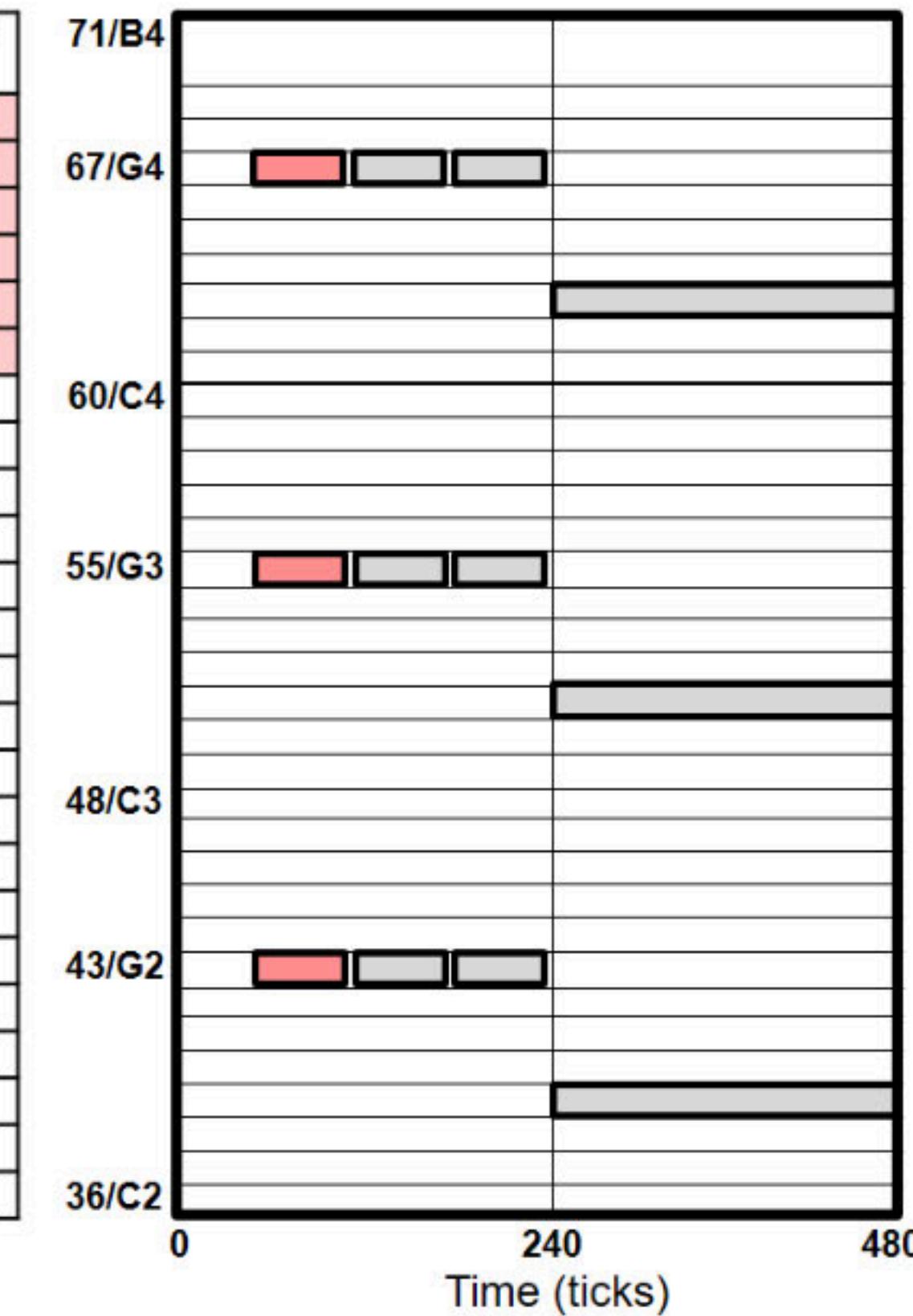
from: (Müller) Fundamentals of Music Processing, Springer 2015

Sound and Music Representations

Symbolic



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



Sound and Music Descriptors

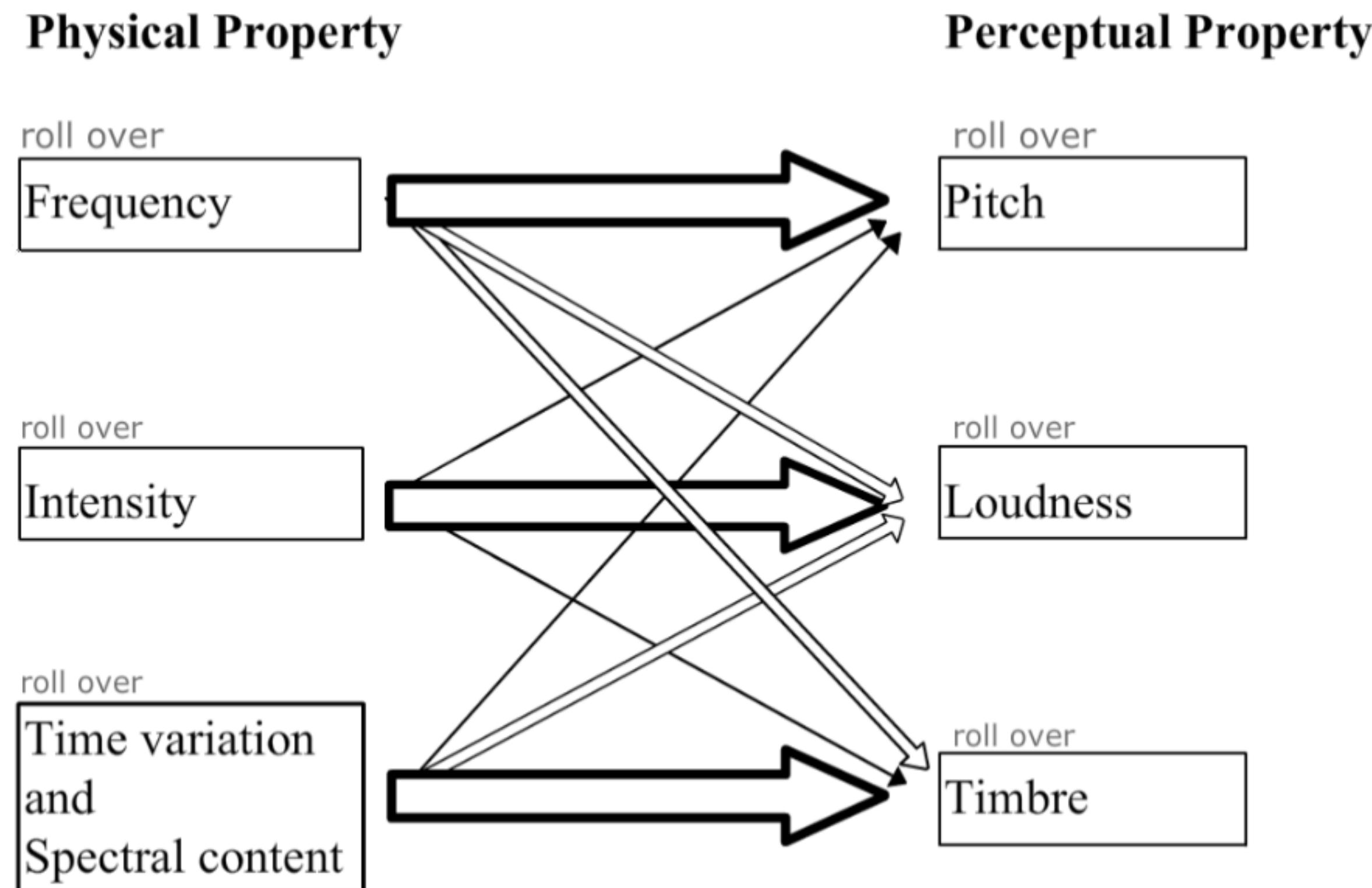
Audio/Symbolic-based

Audio-based			
Mid-level Features	Rhythm	Pitch	Harmony
	- Beat histogram (BH) - Beat-per-minute (BPM)	- Pitch histogram (PH) - Pitch class profile (PCP) - Harmonic pitch class profile (HPCP)	- Chord sequences (CS) - Chord histogram (CH)
Low-level Features	Timbre	Temporal	
	- Mel-frequency cepstrum coefficient (MFCC) - Zero crossing rate (ZRC) - Daubechies wavelet coefficient histogram (DWCH) - Spectral rolloff (SR) - Spectral centroid (SC)	<ul style="list-style-type: none">- Statistical moments (SM)- Auto-regressive modeling (ARM)- Frequency modulation (FM)- Amplitude modulation (AM)	

Symbolic-based	
High-level Features	Pitch
	<ul style="list-style-type: none">- pitch/pitch class histogram- sequence of pitch/pitch class- sequence of melodic intervals- absolute/relative pitch contour- melodic intervals histograms- pitch extension/range- number of pitch contour changes- dominant pitch/pitch class prevalence- pitch volume- pitch class variety- pitch bend fraction- number/ proportion of specific intervals (e.g., major seconds)- dominant melodic interval- quantity of ascending or descending intervals- number of notes- descriptors from the top of pitch or pitch class histograms
Rhythm	Timbre
	<ul style="list-style-type: none">- duration histogram- sequence of inter-onset interval (IOI)- sequence of duration ratio or IOIR- sequence of tempo/meter changes- sequence of time ratios- sequence of duration contour- time attacks- duration range- number of meter/tempo changes- number of duration contour changes- dominant time ratio, tempo, IOIR- number of syncopations- duty factor- number of distinct IOI- number of distinct note durations
Harmony	Timbre
	<ul style="list-style-type: none">- instrument patches- instrument classes- percussion sets- prevalent instrument- histogram of used instruments- number of used instruments- fraction of notes from unpitched instruments- fraction of notes from pitched instruments
Harmony	
<ul style="list-style-type: none">- chord names/degrees progression- used tonal scales- number of non-diatonic/diatonic notes- number of accidental notes- key changes- number of distinctive harmonic intervals- dominant harmonic interval- harmonic interval histogram- chord dictionary/histogram	

Sound and Music Descriptors

Physical/Perceptual



from: (UNES.EDU) Introduction to Sound

notes

Some of the ideas on these slides were inspired in the following recommended resources:

(Müller) Fundamentals of Music Processing - *Springer, 2015*

(Bello) MIR Lecture Notes - *New York University*

(Serra and Smith) Audio Signal Processing for Music Applications - *Coursera*

(Davies) AGPCM Lecture Notes - *FEUP*

(Tjoa) Notes on Music Information Retrieval - musicinformationretrieval.com

End.