# Theoretical Backgrounds of Audio & Graphics

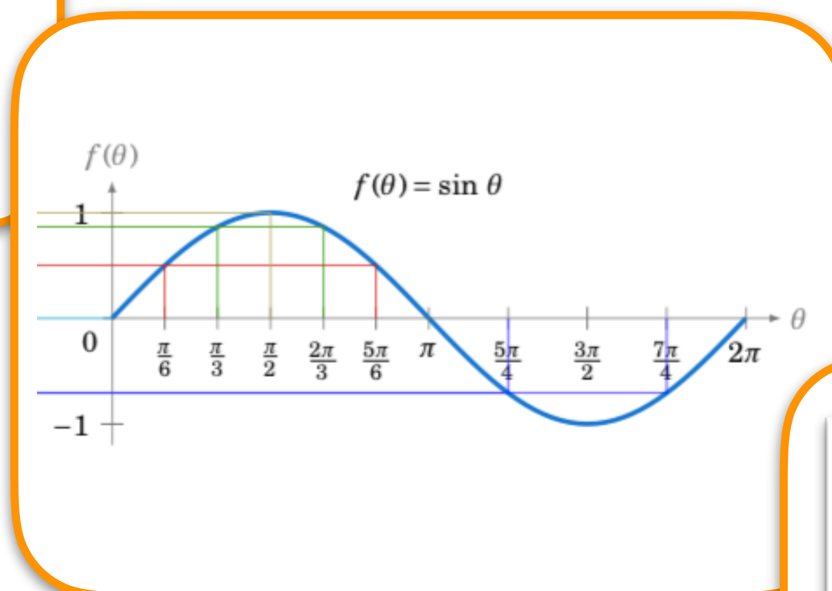
## Digital Audio Fundamentals


Till Bovermann | Dr.-Ing.
Audio & Interactive Media Technologies

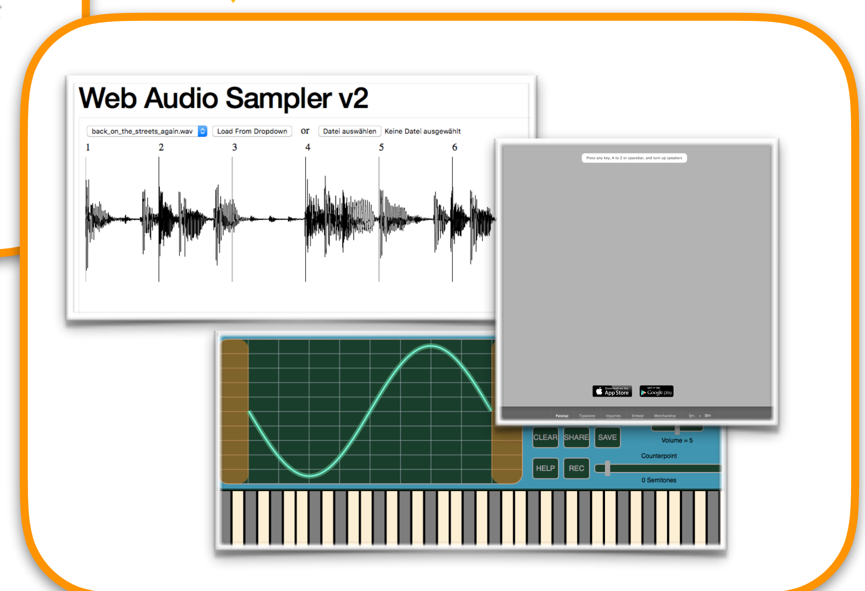
Filmuniversität Babelsberg
*KONRAD WOLF*


Winter term 2019/2020

# Today …



$f(\theta) = \sin\theta$
$f(\theta)$
$\theta$
1
0
−1
$\frac{\pi}{6}$ $\frac{\pi}{3}$ $\frac{\pi}{2}$ $\frac{2\pi}{3}$ $\frac{5\pi}{6}$ $\pi$ $\frac{5\pi}{4}$ $\frac{3\pi}{2}$ $\frac{7\pi}{4}$ $2\pi$

Web Audio Sampler v2

# Top LOTs (learning objectives today)

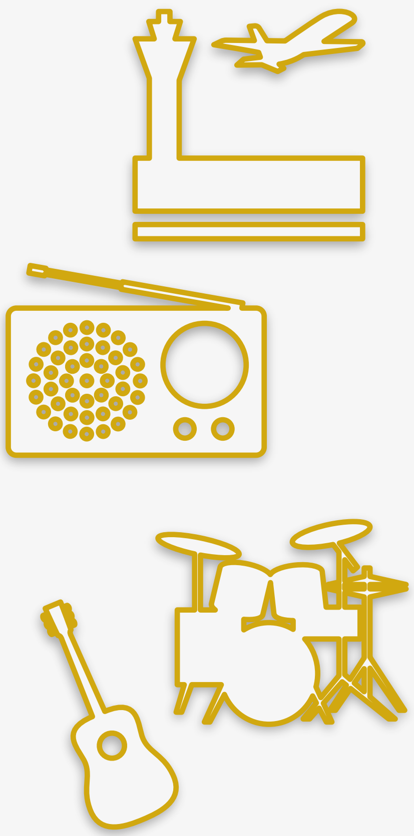
1. What is sound, actually?
2. What are its fundamental physical aspects?
3. How do we describe sound formally?
4. What is the digital representation of sound?
5. How do we work with „digital sound“?

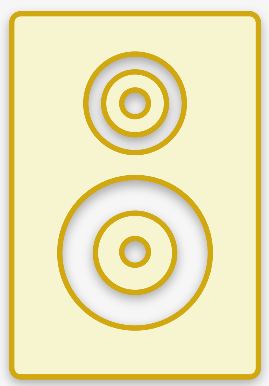
# Contents

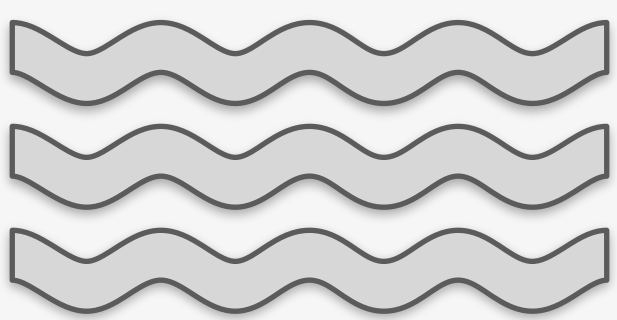
- Sound & Audio
- Properties of Sound
- Mathematical Representation
- Digital Representation
- Audio Programming

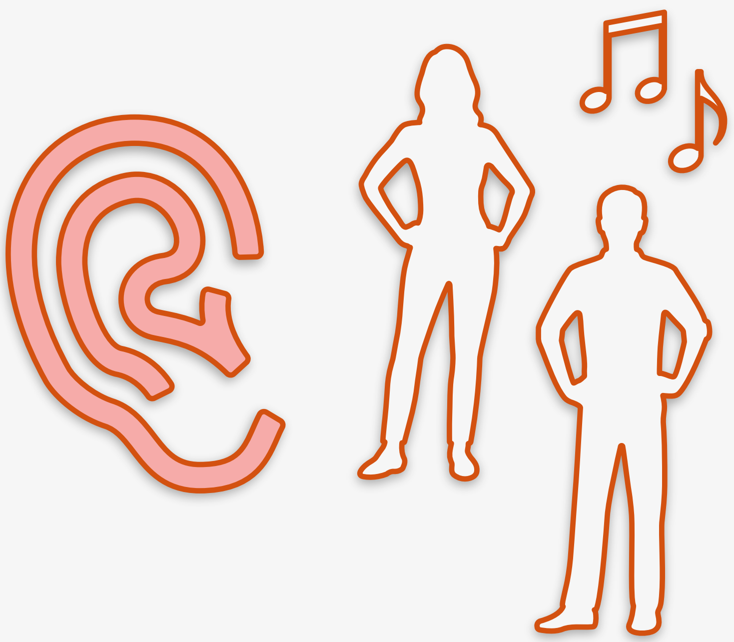
# Sound & Audio

# Simple Acoustic Scene









**Sources**
that create sound

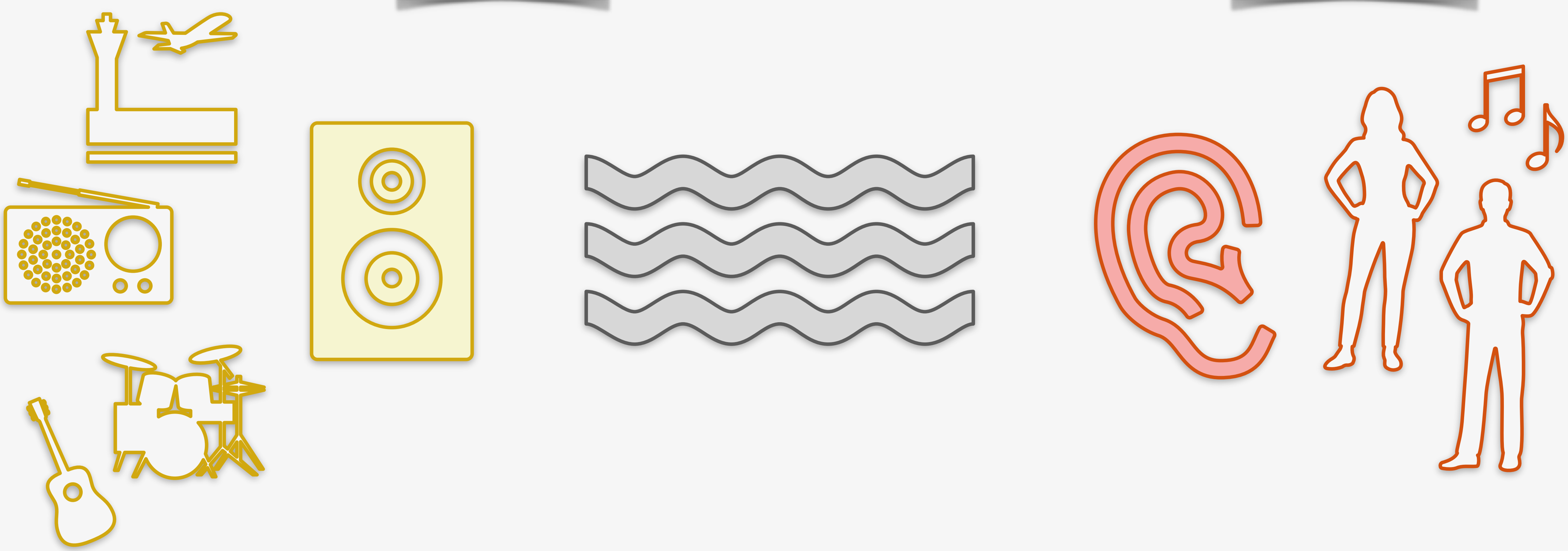
**Medium**
through which
sound propagates

**Listeners**
who perceive sound

# Simple Acoustic Scene

(Physical)
**Sound event**

(Cognitive)
**Auditory event**



**Sources**
that create sound

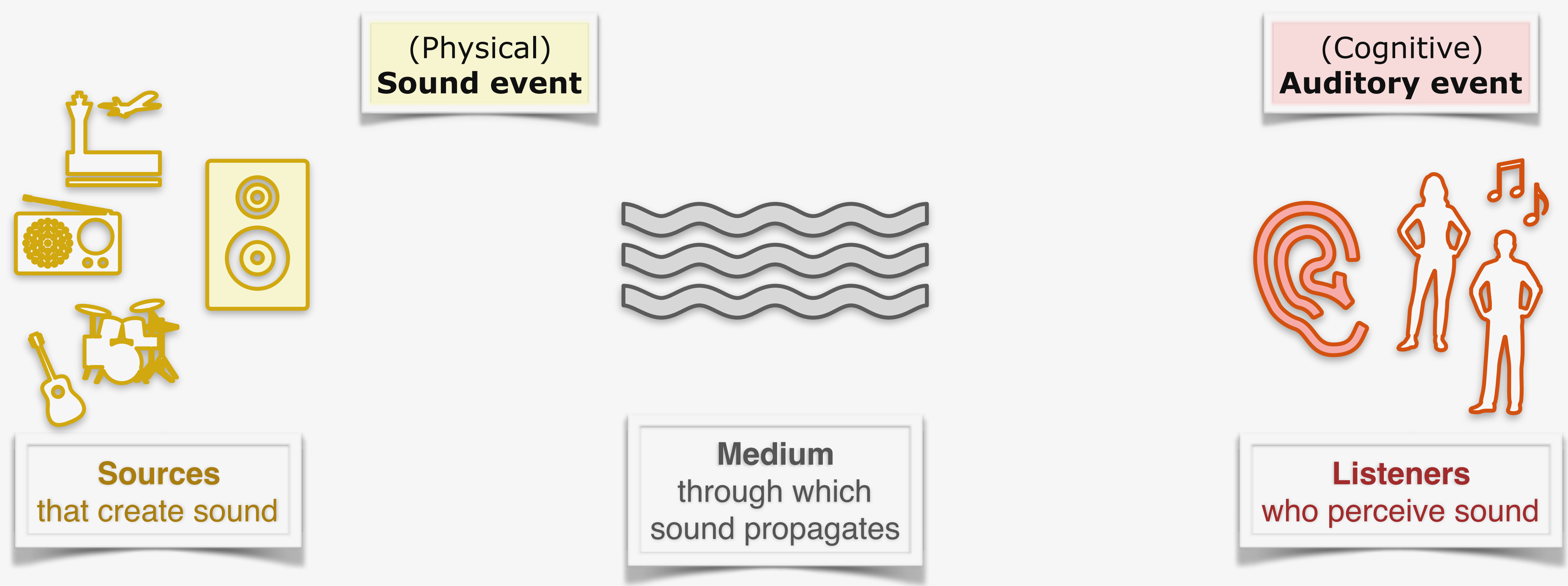
**Medium**
through which
sound propagates

**Listeners**
who perceive sound

# Simple Acoustic Scene

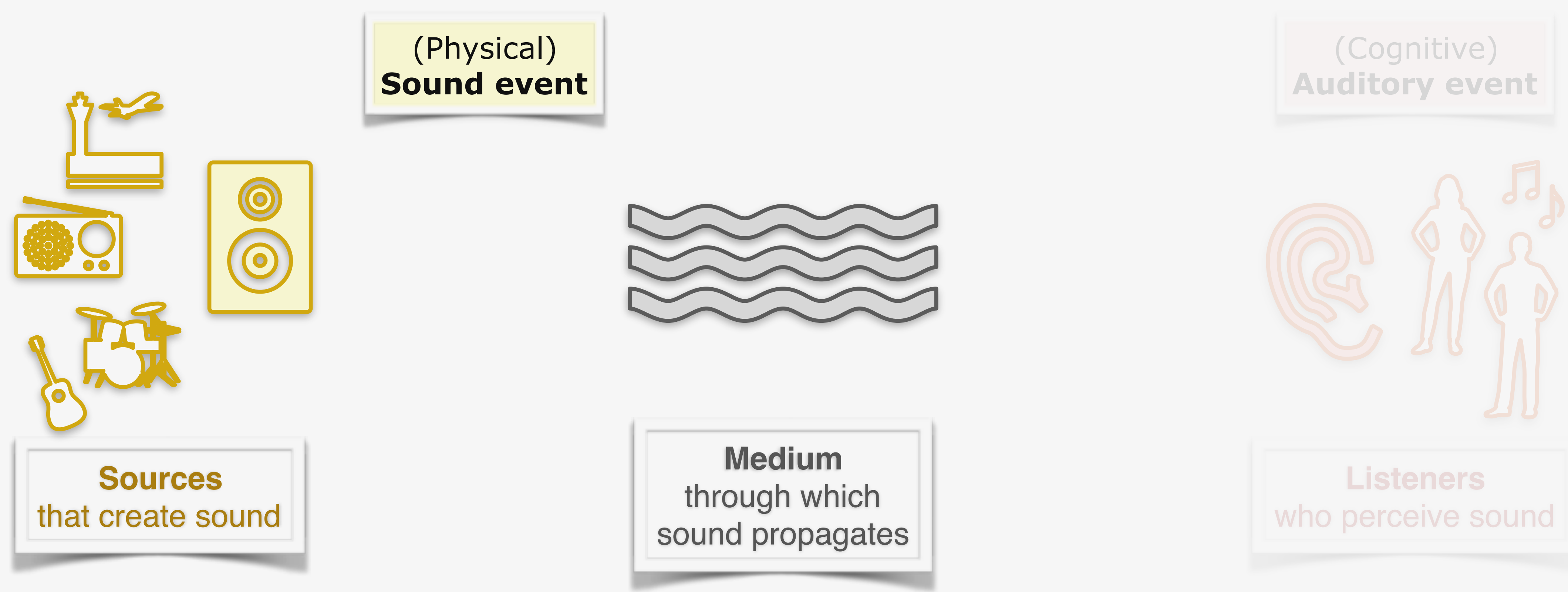
Psychophysical or cognitive is anything that we perceive about the physical world

- Sound is a complex physical & psychophysical (cognitive) phenomenon
- Audio usually refers to sound that is within the human hearing range



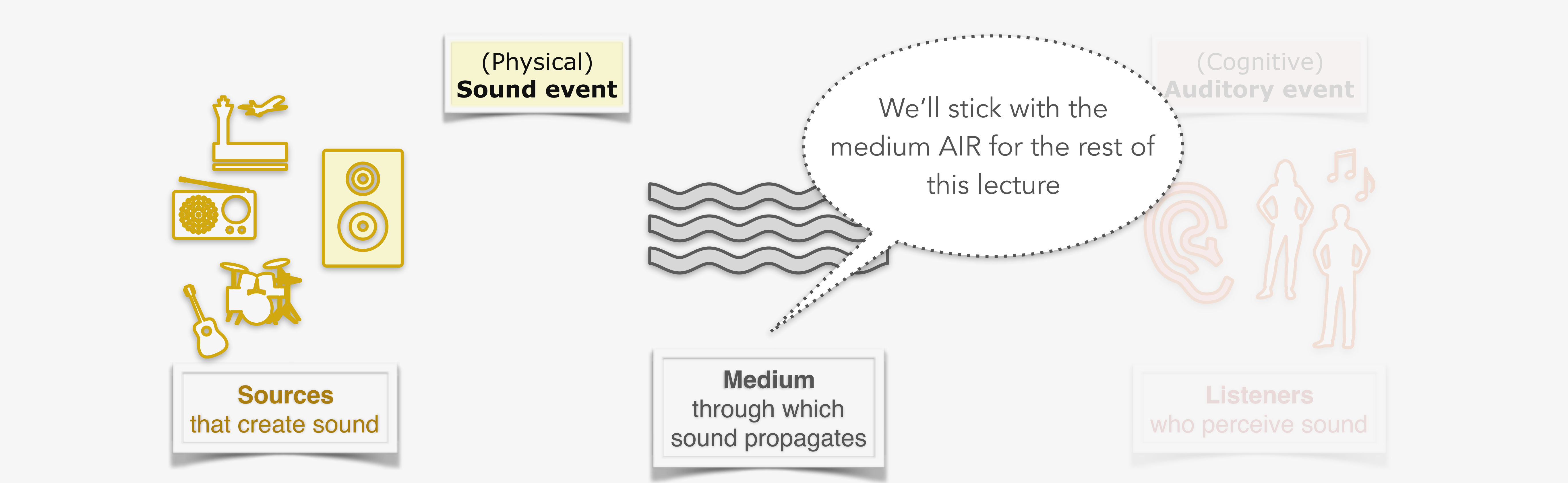
# Simple Acoustic Scene

- Sound in a physical sense
  - Mechanical pressure wave transmits energy through elastic medium
  - Wave & medium have measurable properties that allow to clearly describe sound field or wave propagation (i.e., frequency, amplitude, medium density)



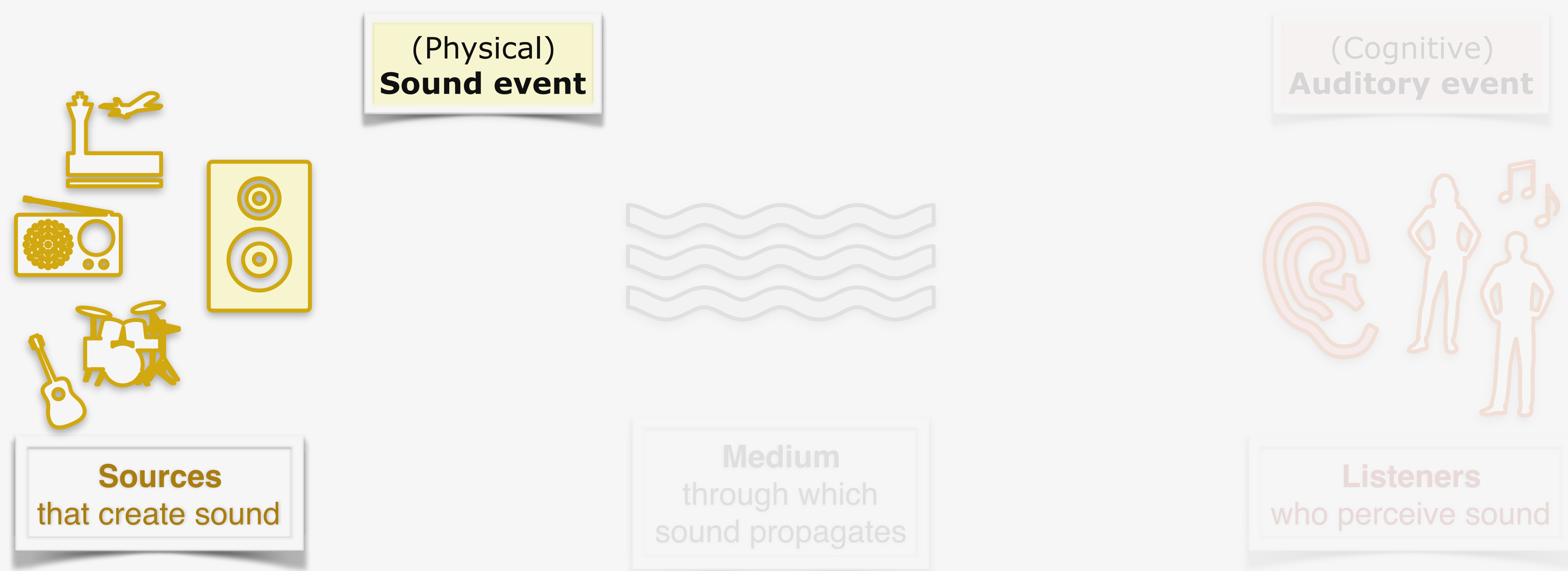
# Simple Acoustic Scene

- Sound in a physical sense
  - Mechanical pressure wave transmits energy through elastic medium
  - Wave & medium have measurable properties that allow to clearly describe sound field or wave propagation (i.e., frequency, amplitude, medium density)



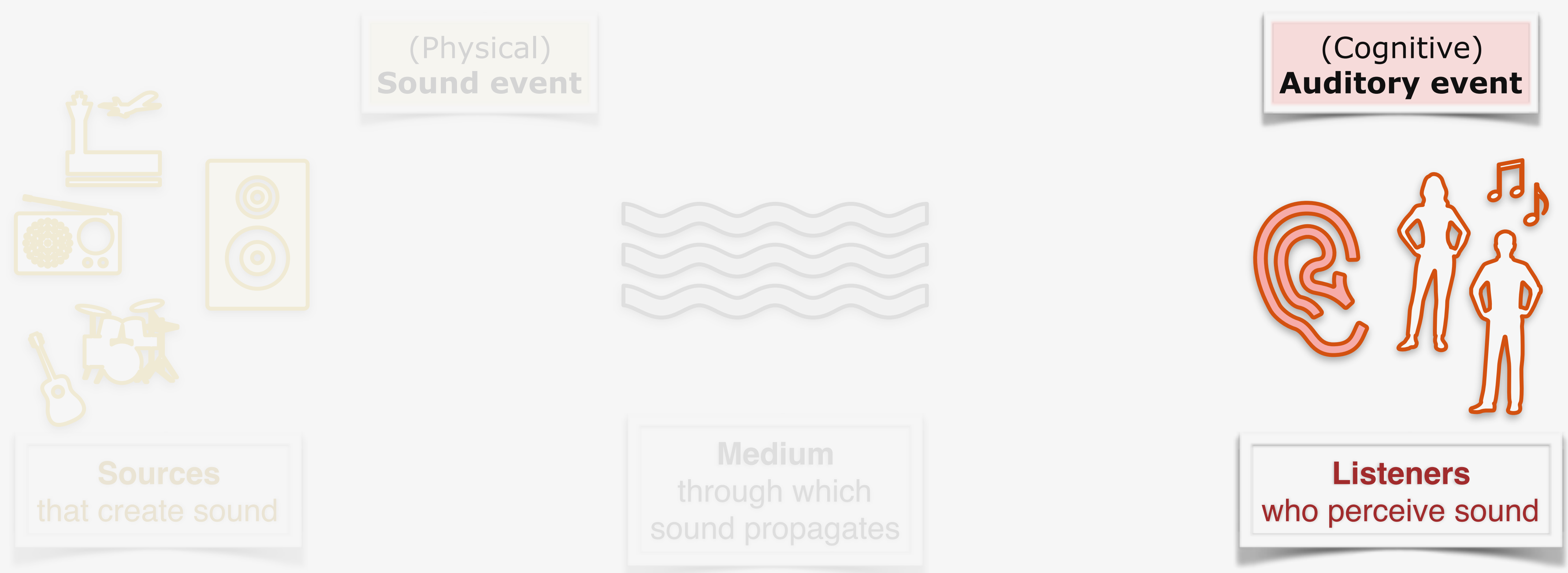
# Simple Acoustic Scene

- Sound in a physical sense
  - Mechanical pressure wave transmits energy through elastic medium
  - Wave & medium have measurable properties that allow to clearly describe sound field or wave propagation (i.e., frequency, amplitude, medium density)



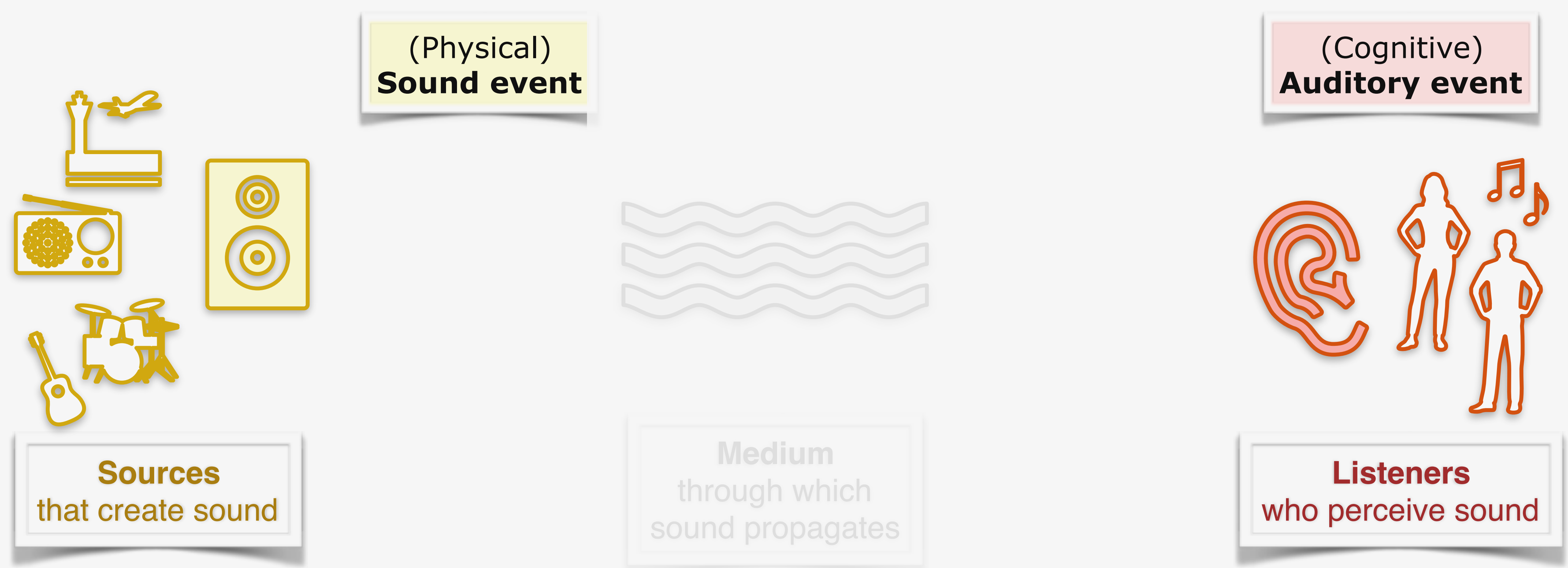
# Simple Acoustic Scene

- Sound in a psychophysical sense:
  - A perception (loudness, pitch, etc.) inside the mind of the listener
  - A sensation stimulated in the organs of hearing caused by the vibrating eardrum & interpreted by the brain as an audible sound



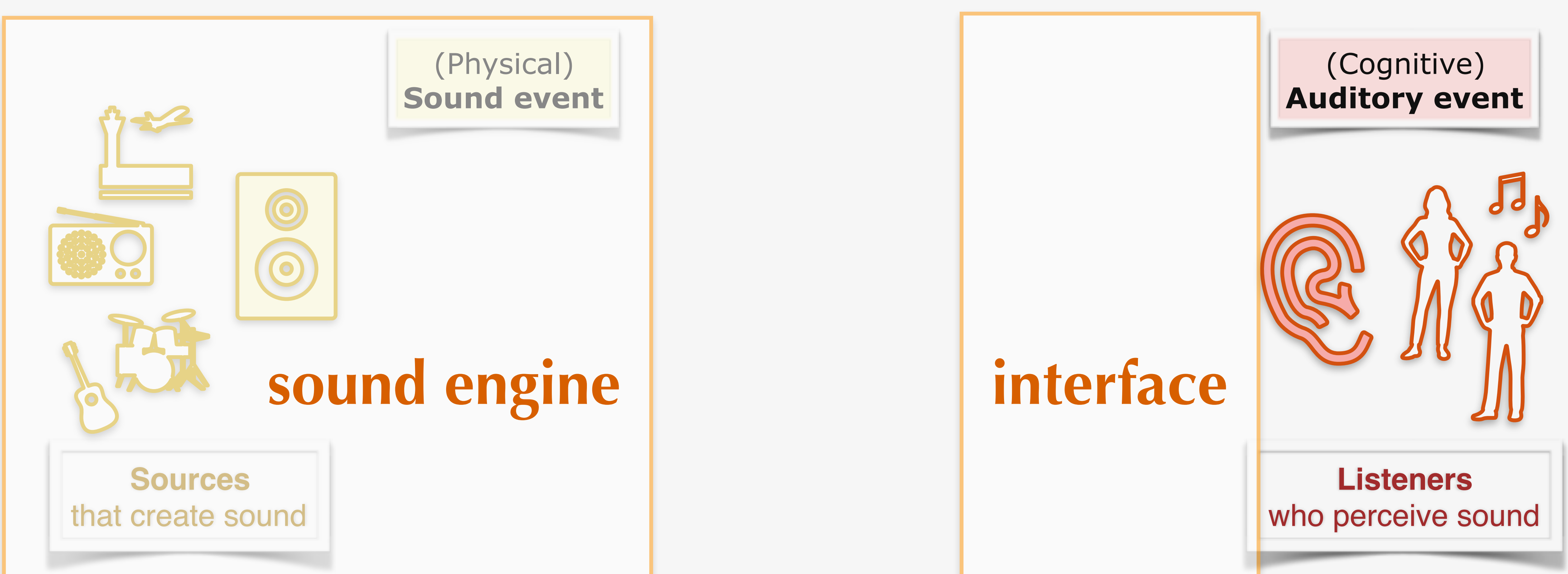
# Simple Acoustic Scene

- What is sound, actually?
  - Sound consists of physical and psychophysical aspects
  - Audio systems address the human auditory system



# Simple Acoustic Scene

- Intuitive approach to an audio software system
  - A sound engine that generates & processes sound
  - An interface to interact with & parameterize sound



# Properties of Sound

# Sound Source

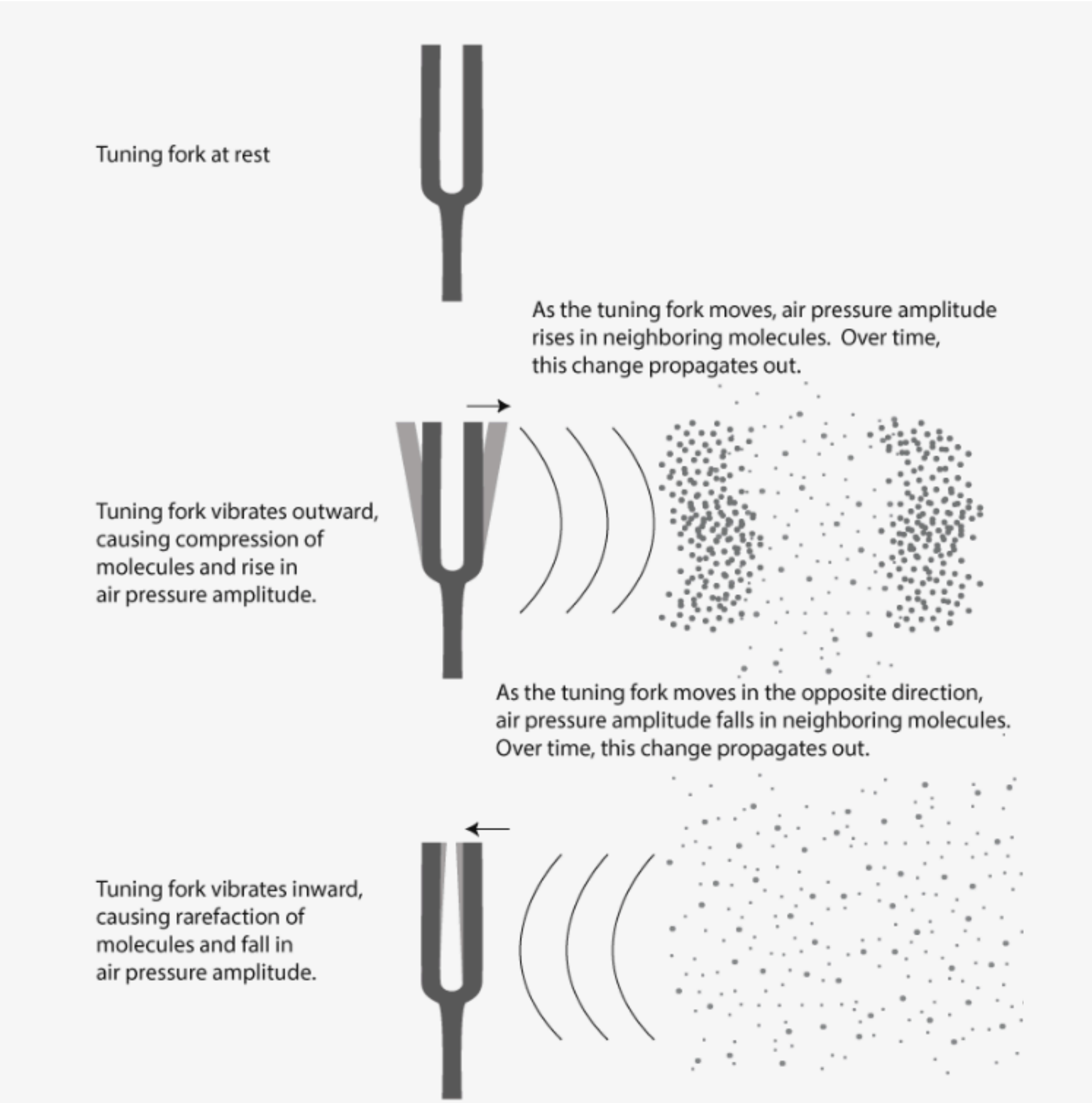



Image credit: http://digitalsoundandmusic.com/chapters/ch2/#prettyPhoto[gallery600]/1/

- A vibrating object causes a mechanical disturbance in a medium (i.e., air)
- If the disturbance or oscillation is strong enough, it starts propagating through the medium, i.e., neighboring particles are excited and start oscillating
- A wave is a sequence of oscillating particles of the medium that transfer energy through the medium

# Sound Perception

- When the wave reaches our ear, the eardrum oscillates
- Nerve impulses are sent to the brain & interpreted
- When the wave is audible, i.e., within the range of hearing, we are talking about a sound wave

Image credit: Sethares, W.A. (2005): Tuning, Timbre, Spectrum, Scale. London: Springer-Verlag.

# Sound Wave

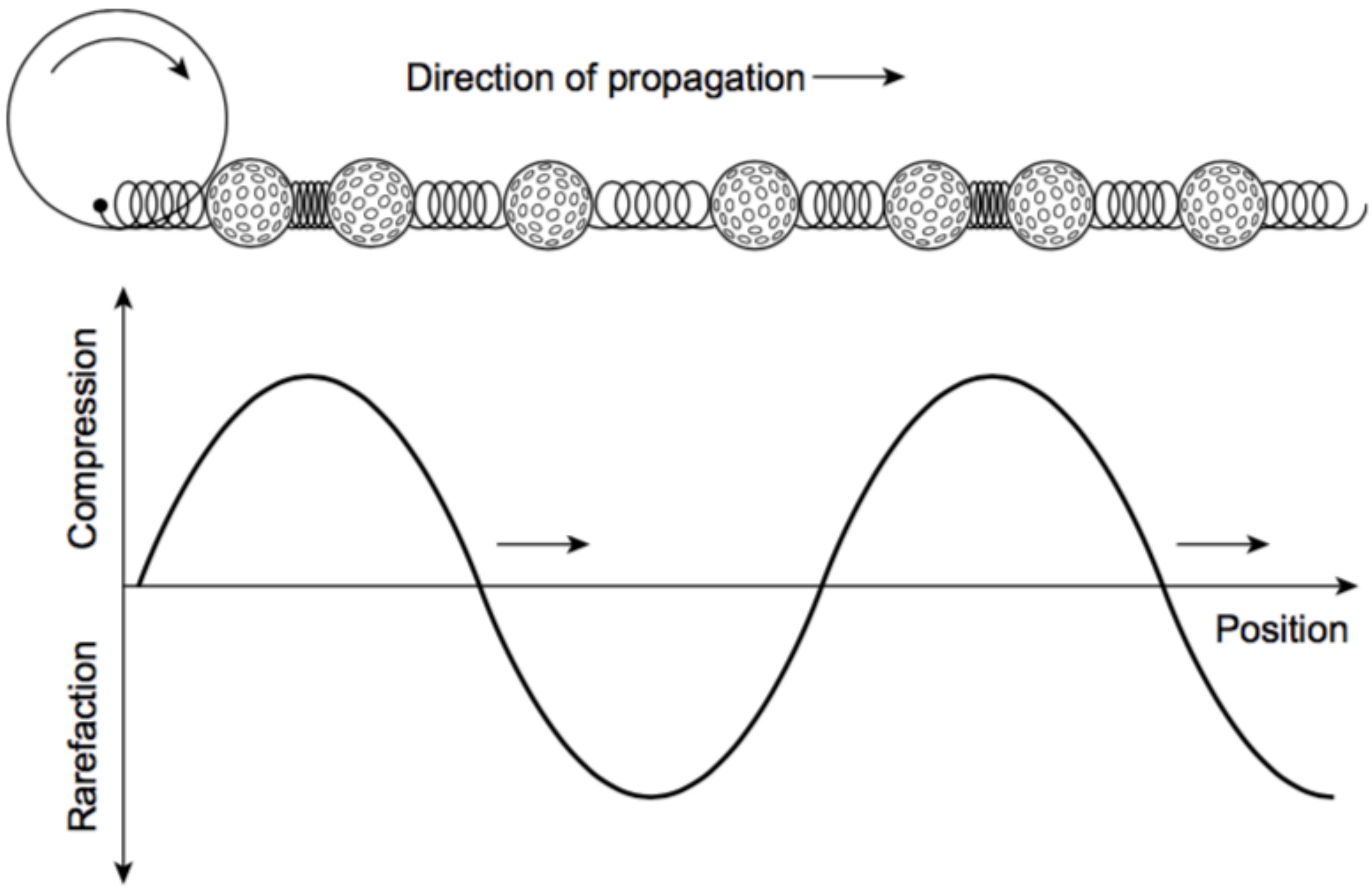
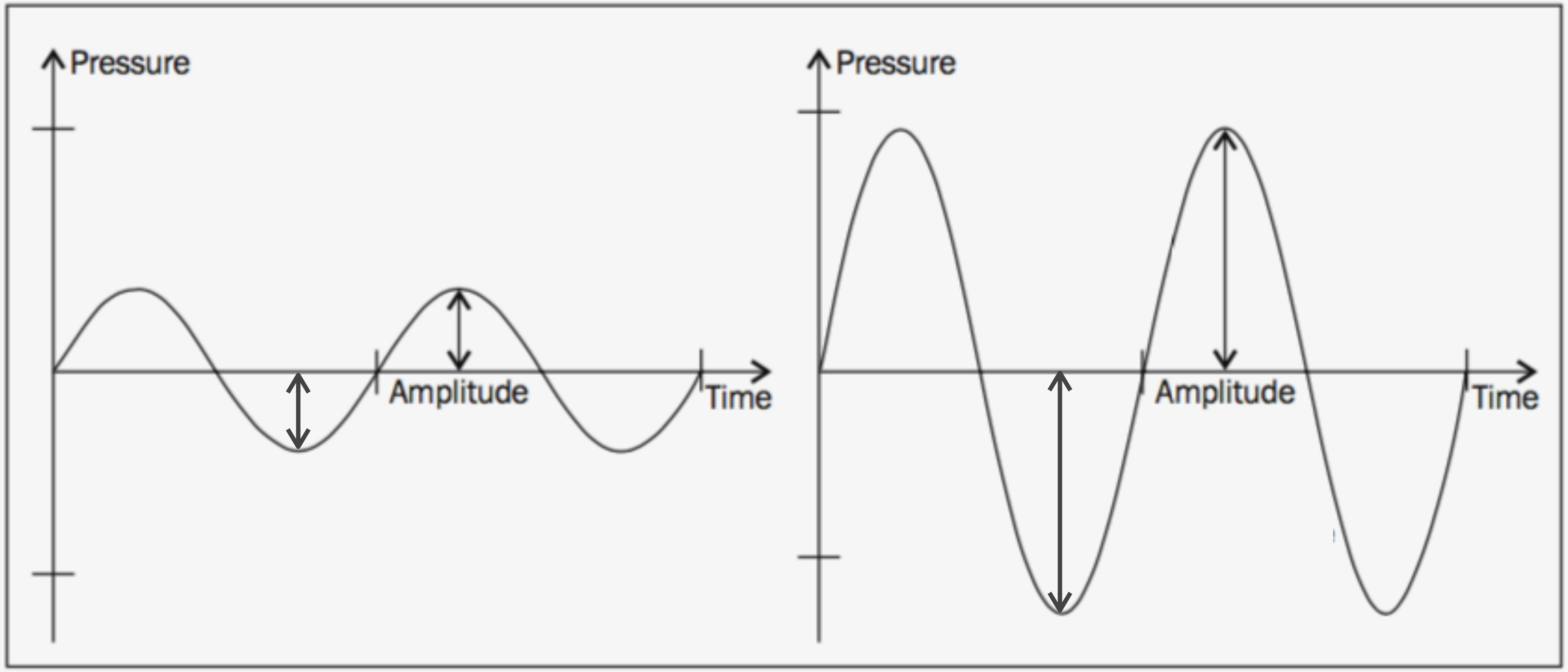



Image credit Howard, D.M. & Angus, J.A.S. (2009): Acoustics and Psychoacoustics. 4th Edition. Oxford, UK: Focal Press.

# Sound Wave Properties

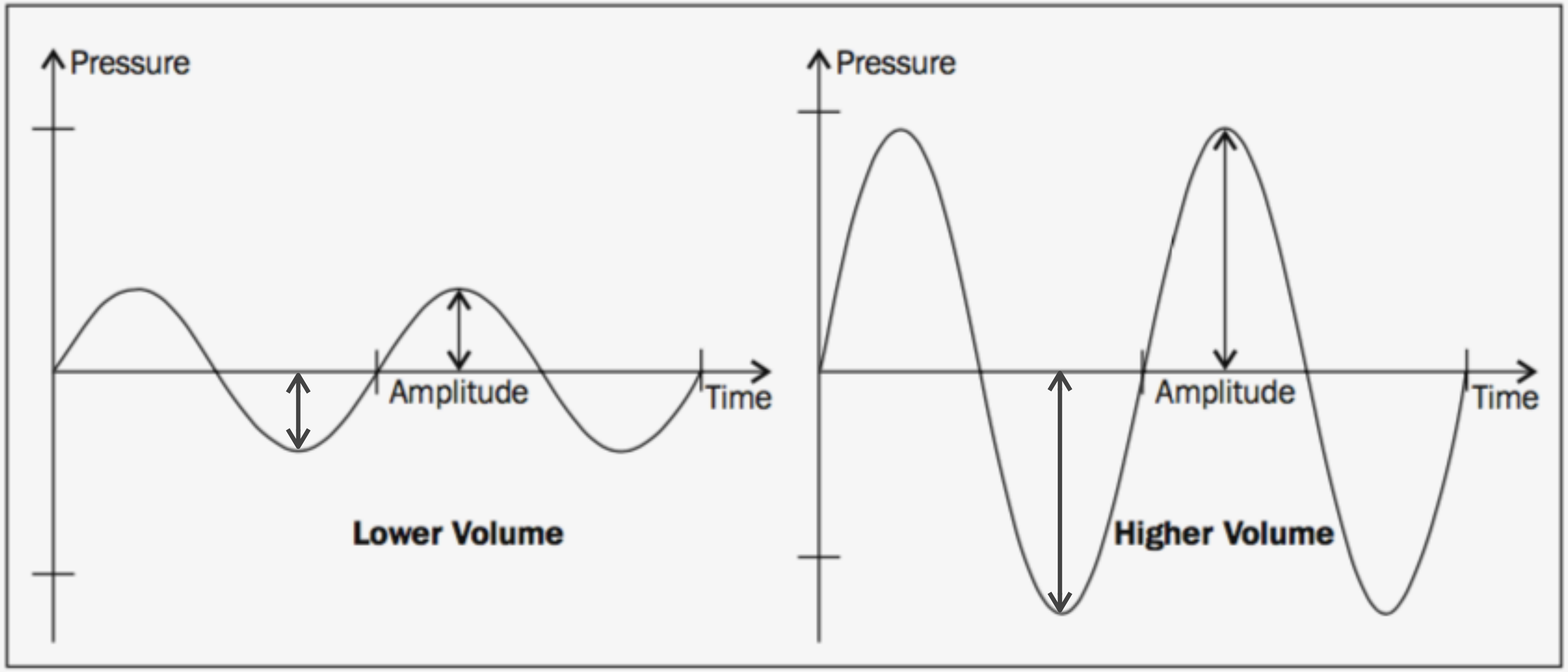
- **Amplitude *A*** describes the strengths of the air pressure changes, the maximum magnitude or distance of displacement from equilibrium to maximal compression or rarefaction.




Images credit: Gouveia, D. (2013): Getting Started with C++ Audio Programming for Game Development. Birmingham, UK: Packt Publishing Ltd.

# Sound Wave Properties

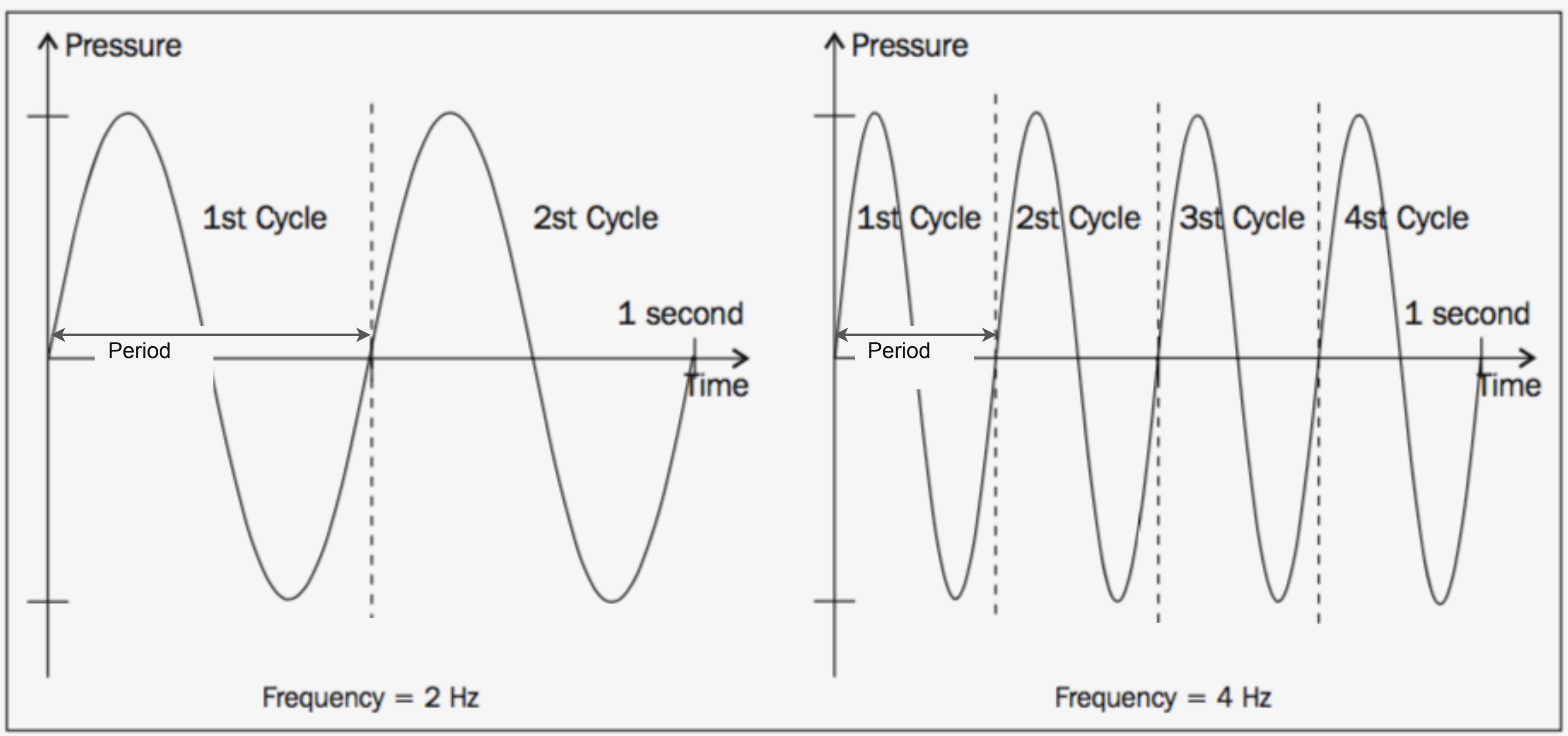
- **Amplitude *A*** is (roughly) perceived as **loudness**




Images credit: Gouveia, D. (2013): Getting Started with C++ Audio Programming for Game Development. Birmingham, UK: Packt Publishing Ltd.

# Sound Wave Properties

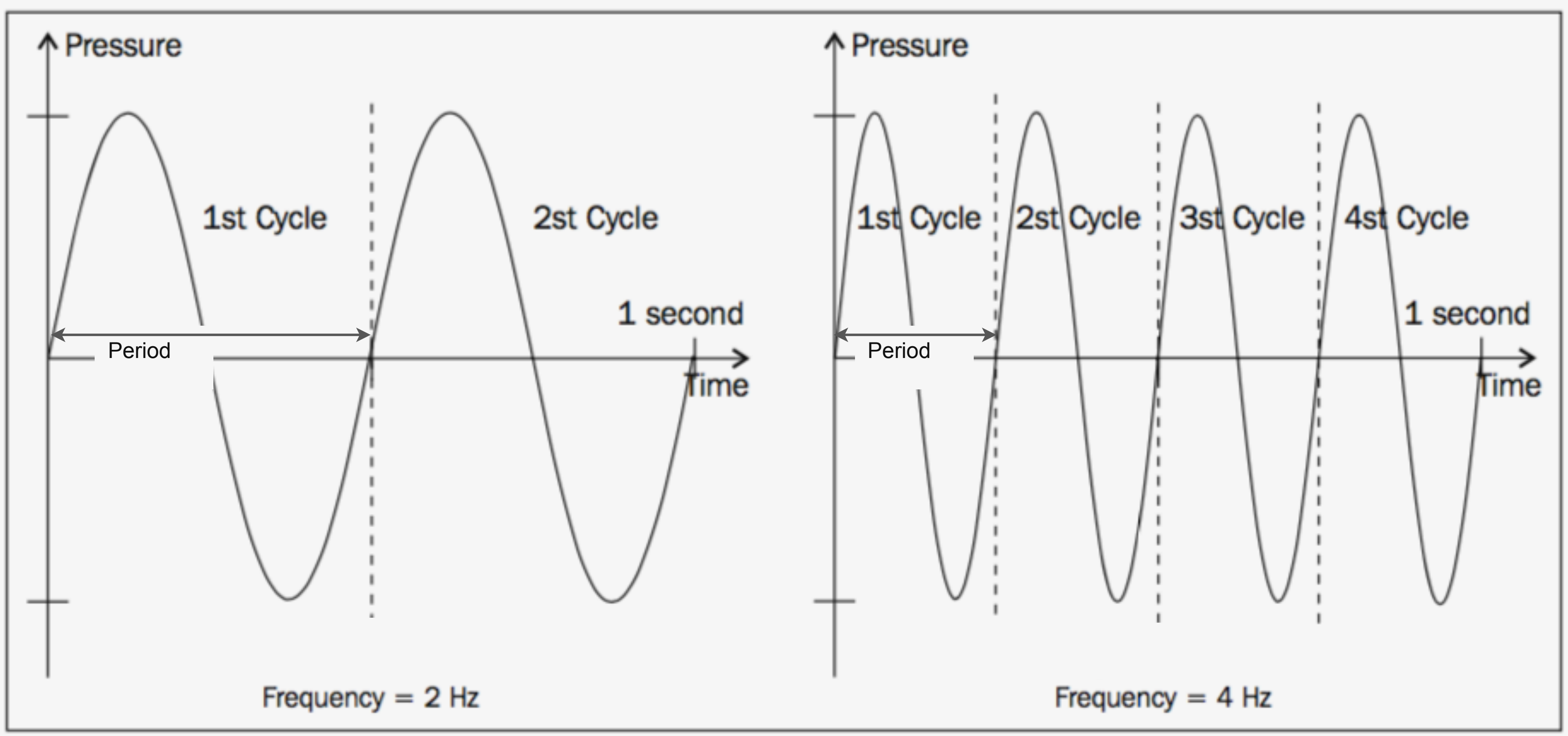
- **Period *T*** is the time in seconds (**s**) required for an entire cycle
- A **cycle** is one sequence of an oscillation from pressure compression to rarefaction & back to equilibrium




Images credit: Gouveia, D. (2013): Getting Started with C++ Audio Programming for Game Development. Birmingham, UK: Packt Publishing Ltd.

# Sound Wave Properties

- Frequency $f$ describes the rate at which the air pressure changes from compression to rarefaction per second
- It is measured as number of cycles per second in Hertz (Hz)




Period & frequency are inversely related by

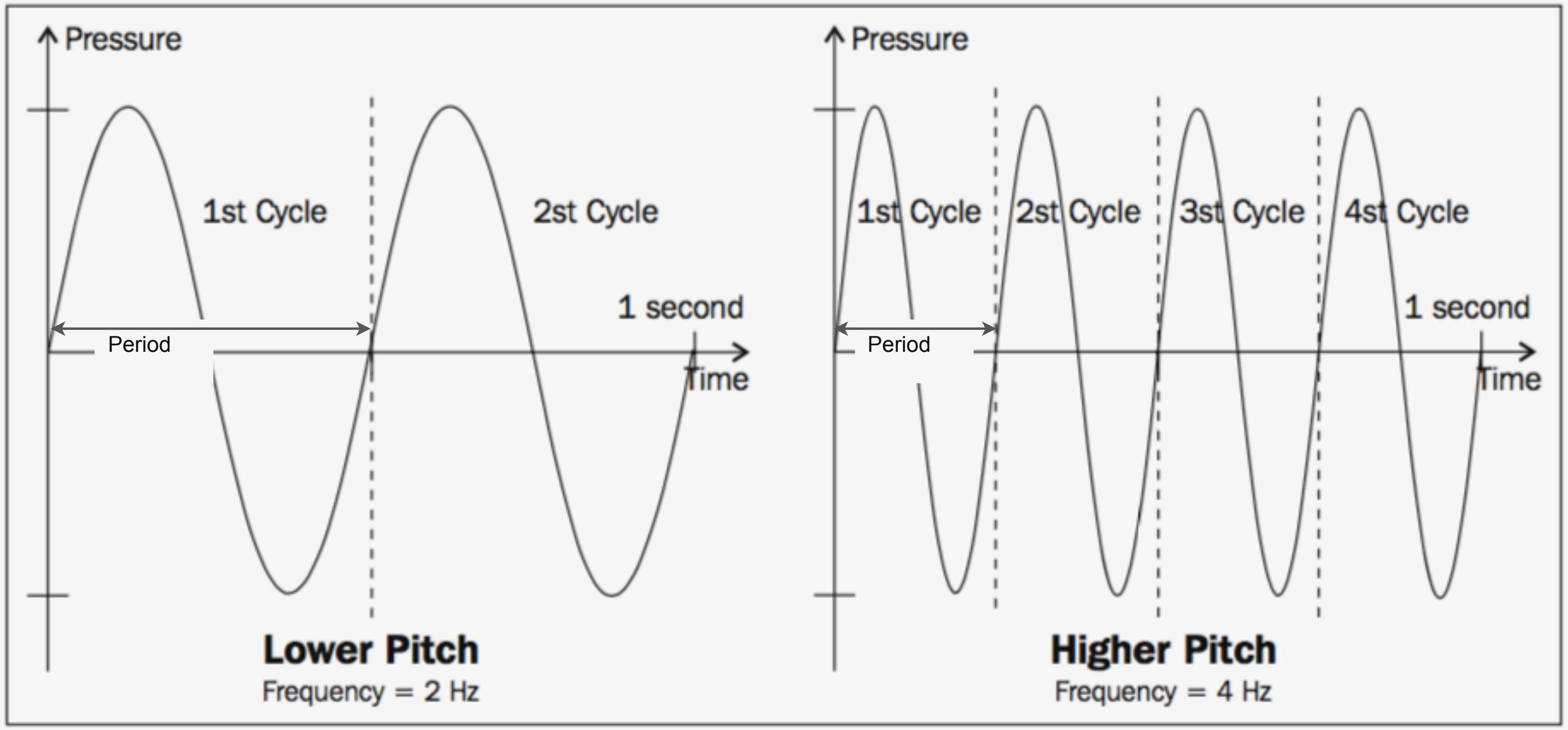
$$f = 1 / T$$

$$T = 1 / f$$

Images credit: Gouveia, D. (2013): Getting Started with C++ Audio Programming for Game Development. Birmingham, UK: Packt Publishing Ltd.

# Sound Wave Properties

- Frequency $f$ is roughly perceived as pitch.
  - the higher the frequency >> the higher the perceived pitch
- typical human hearing ranges between around 20Hz and 20kHz



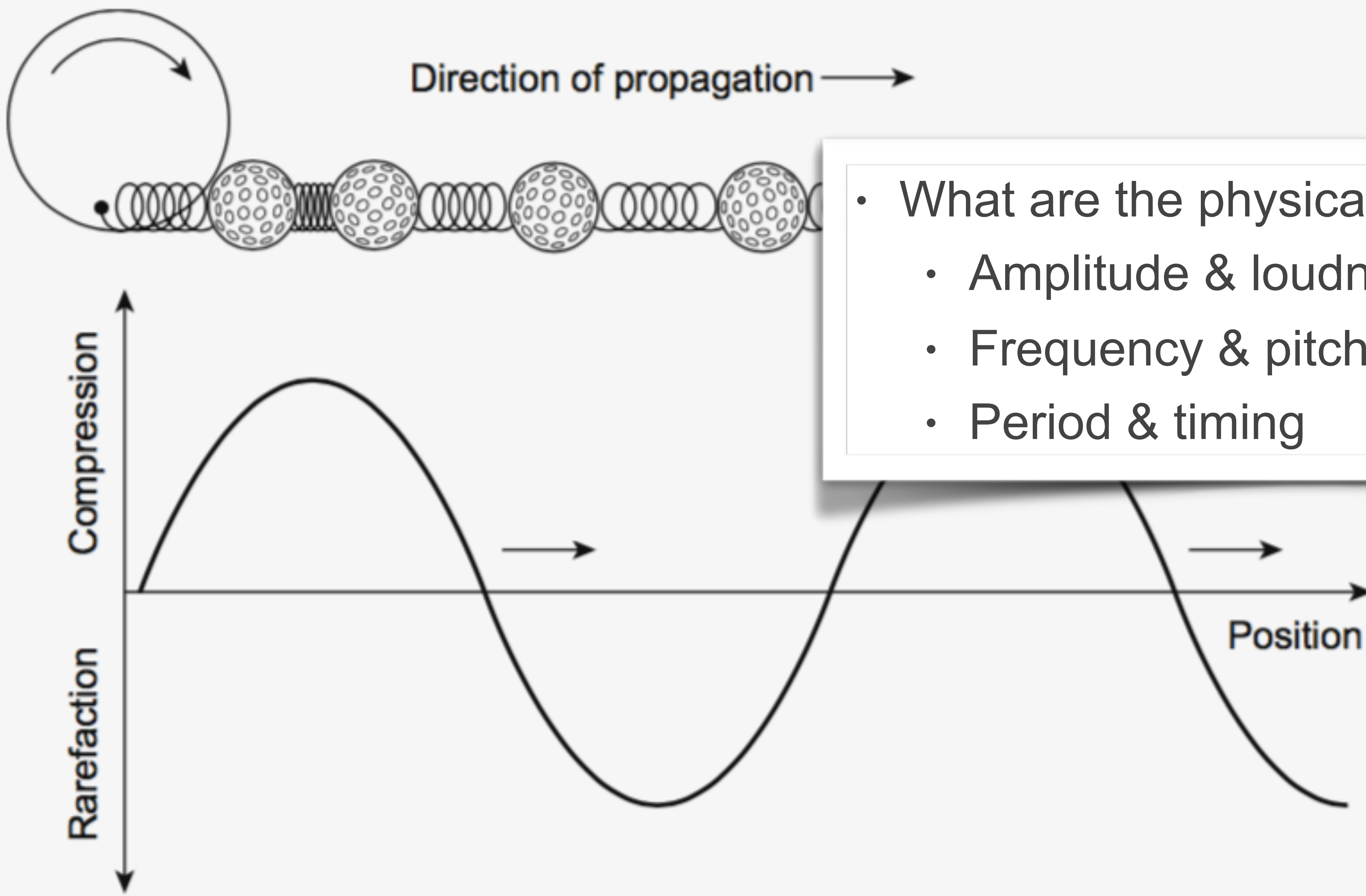

Period & frequency are inversely related by

$$f = 1 / T$$

$$T = 1 / f$$

Images credit: Gouveia, D. (2013): Getting Started with C++ Audio Programming for Game Development. Birmingham, UK: Packt Publishing Ltd.

# Sound Wave Properties




- What are the physical aspects of sound?
  - Amplitude & loudness
  - Frequency & pitch
  - Period & timing

Image credit Howard, D.M. & Angus, J.A.S. (2009): Acoustics and Psychoacoustics. 4th Edition. Oxford, UK: Focal Press.

# Types of Sound Waves

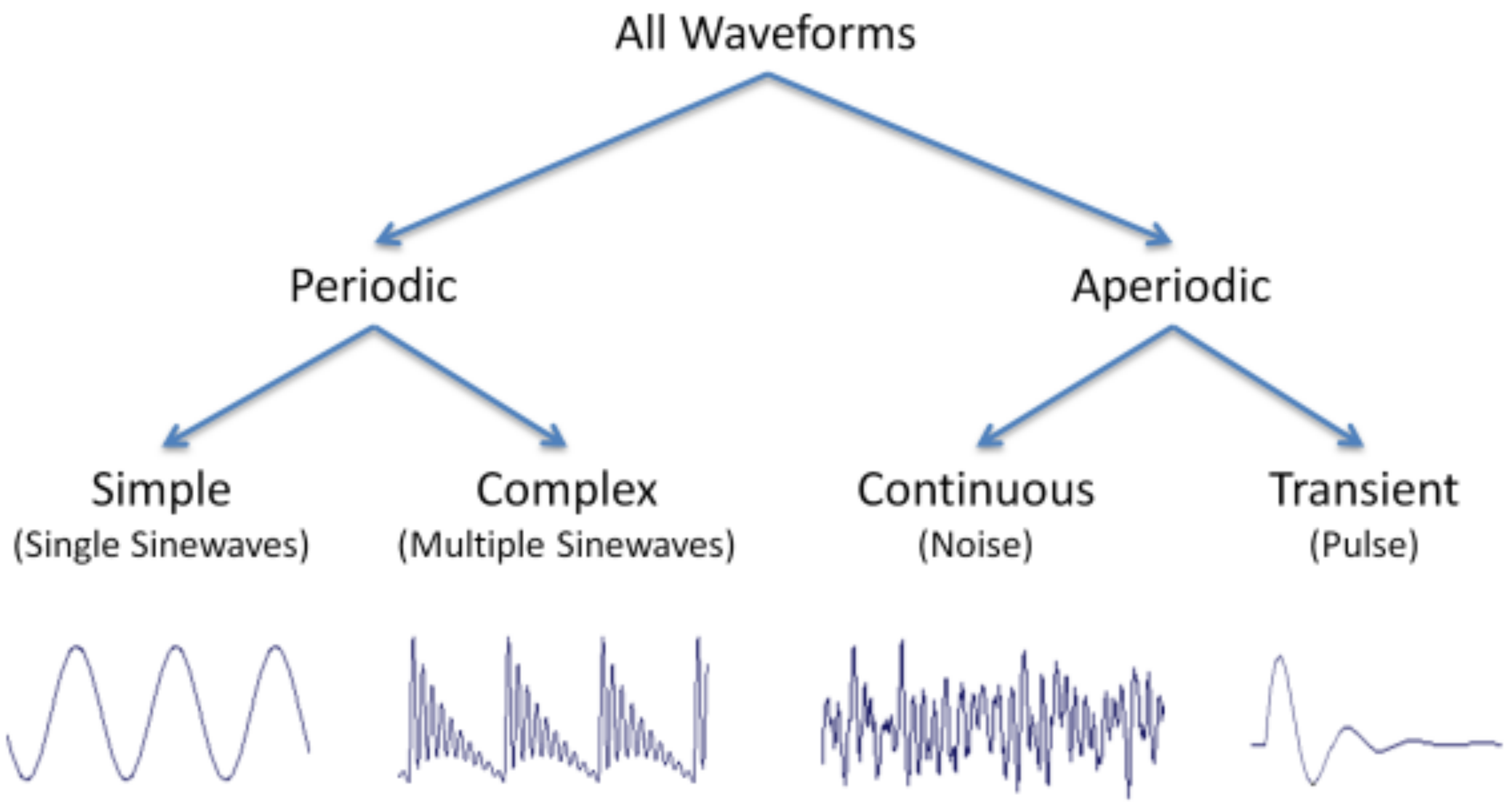



Image credit: http://www.phon.ucl.ac.uk/courses/spsci/iss/

# Types of Sound Waves

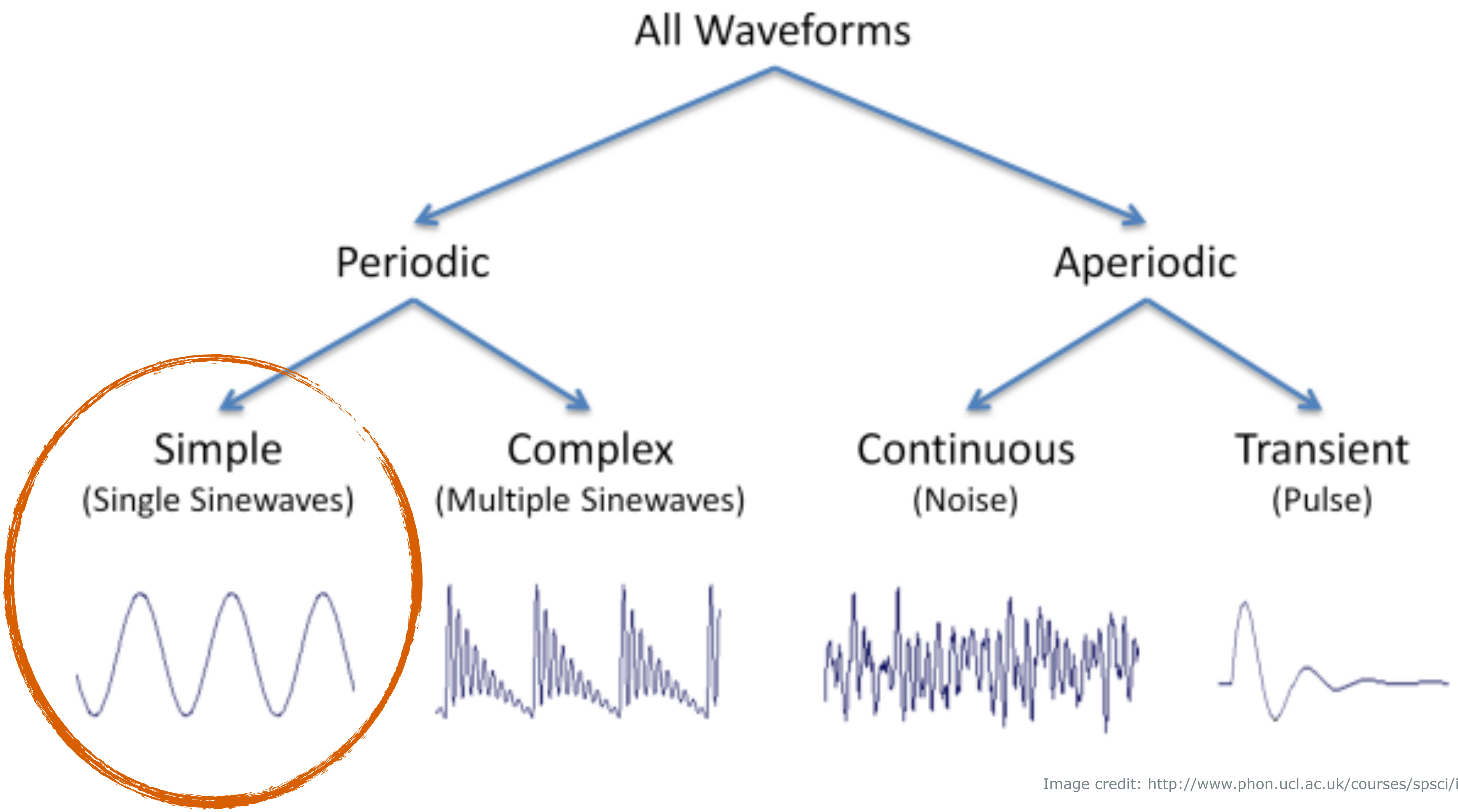



Image credit: http://www.phon.ucl.ac.uk/courses/spsci/iss/

# Mathematical Representation

# Sound & Sine Waves

- The oscillation caused by a vibrating object can be expressed mathematically by a periodic function of **time** (x-axis) against **amplitude** (y-axis)

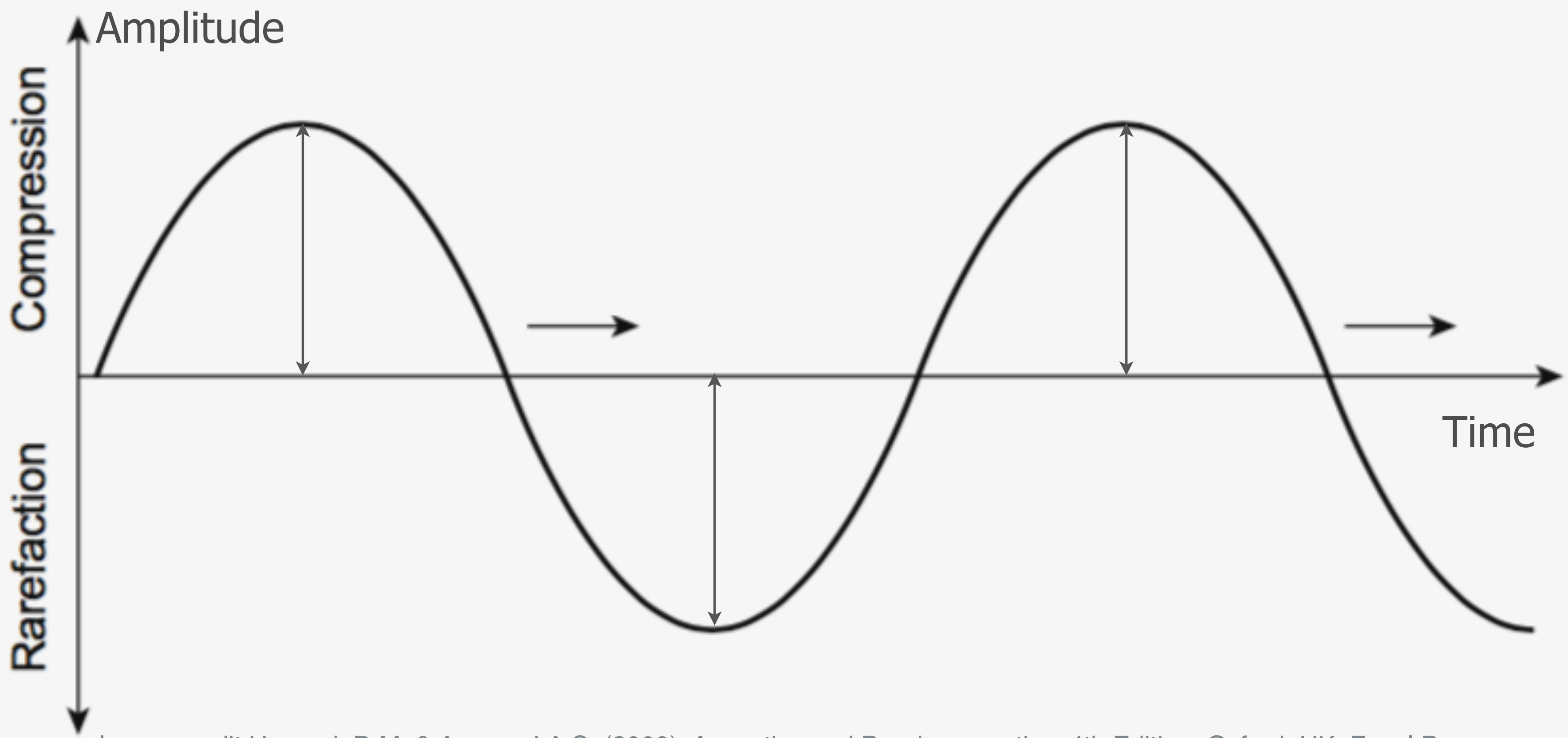



Image credit Howard, D.M. & Angus, J.A.S. (2009): Acoustics and Psychoacoustics. 4th Edition. Oxford, UK: Focal Press.

# Sound & Sine Waves

- Trigonometric functions **sine** & **cosine** describe oscillations well
- They have a **period of 2π** and are phased shifted by 2/π

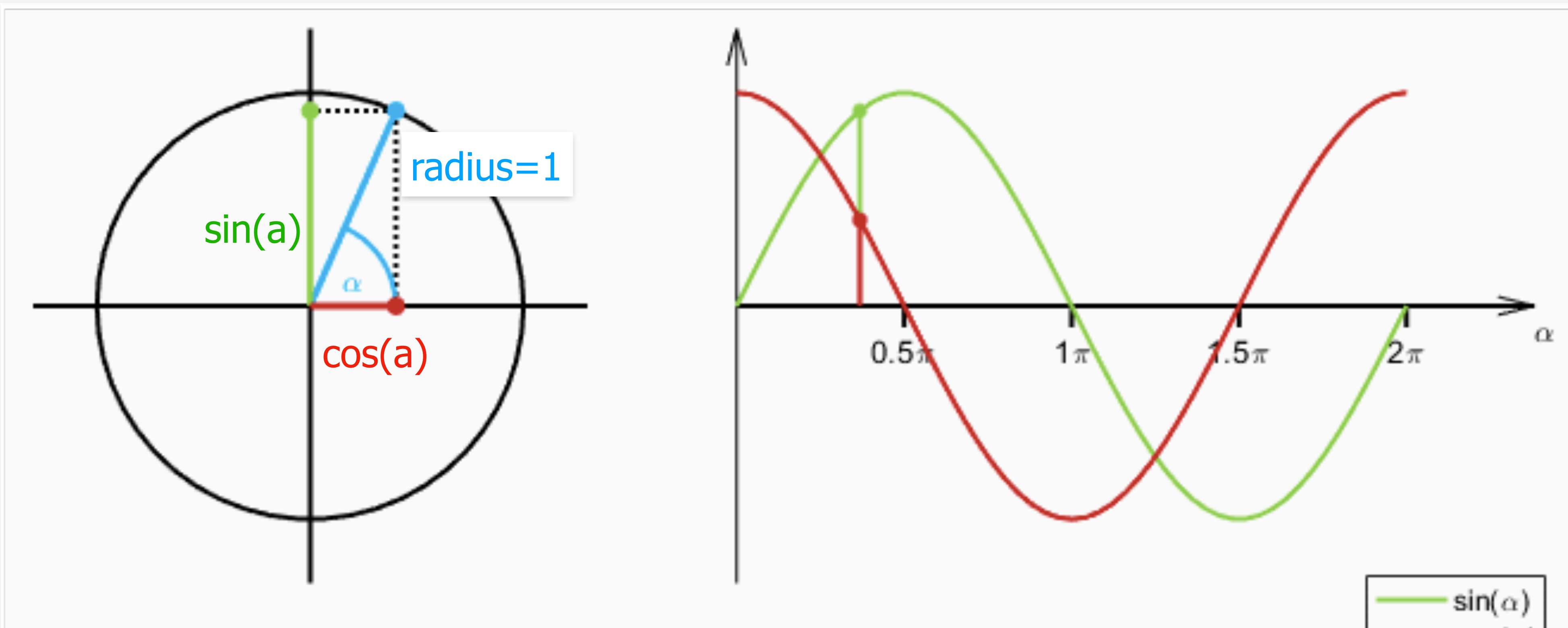



Image credit: https://upload.wikimedia.org/wikipedia/commons/f/f3/Sinus_und_Cosinus_am_Einheitskreis.gif

$$\sin x = \cos\left(x - \frac{\pi}{2}\right)$$

$$\cos x = \sin\left(x + \frac{\pi}{2}\right)$$

# Sound & Sine Waves

- In geometric terms, sine and cosine can be expressed as a vector p = (cos(ø), sine(ø)) that moves around the unit circle

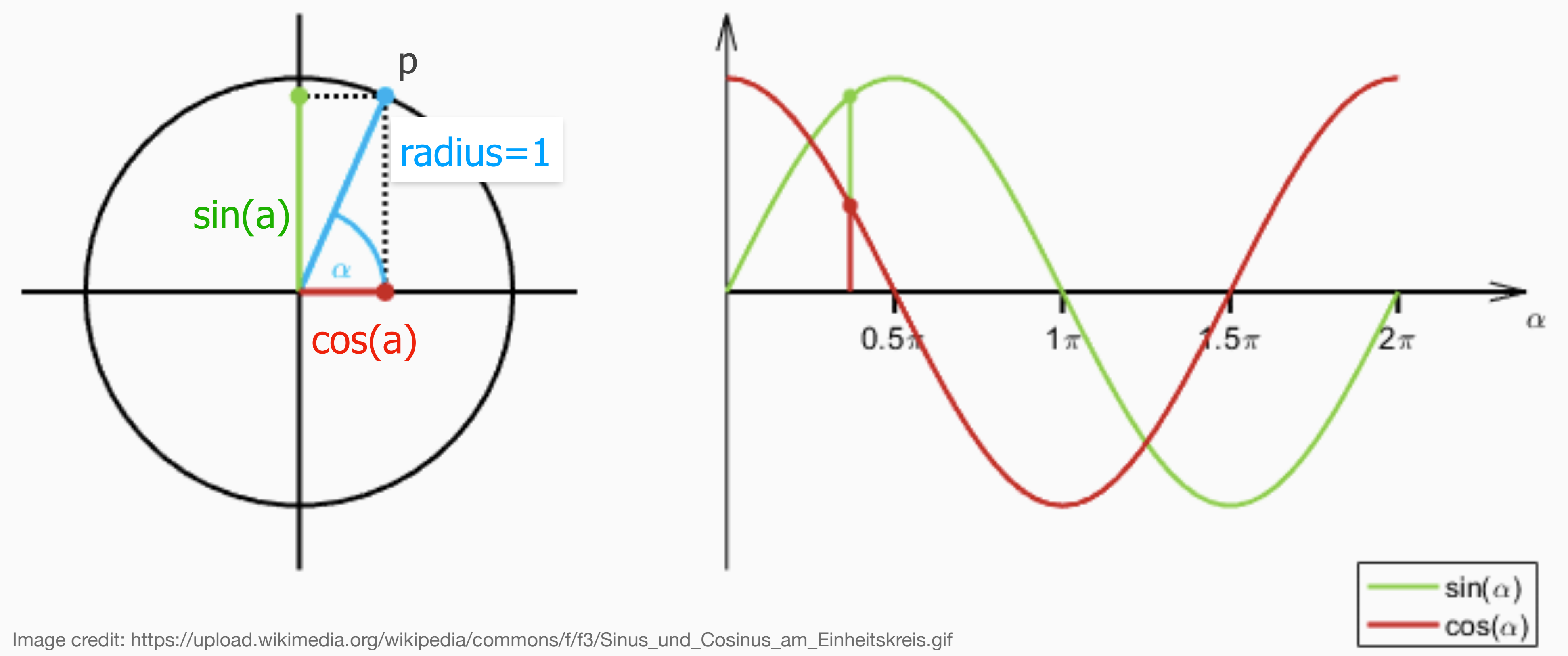



Image credit: https://upload.wikimedia.org/wikipedia/commons/f/f3/Sinus_und_Cosinus_am_Einheitskreis.gif

# Sound & Sine Waves

- A continuous process that grows the angle and lets vector p move around the circle is called **oscillator**
- An oscillator generates a periodic signal, i.e., a sine wave

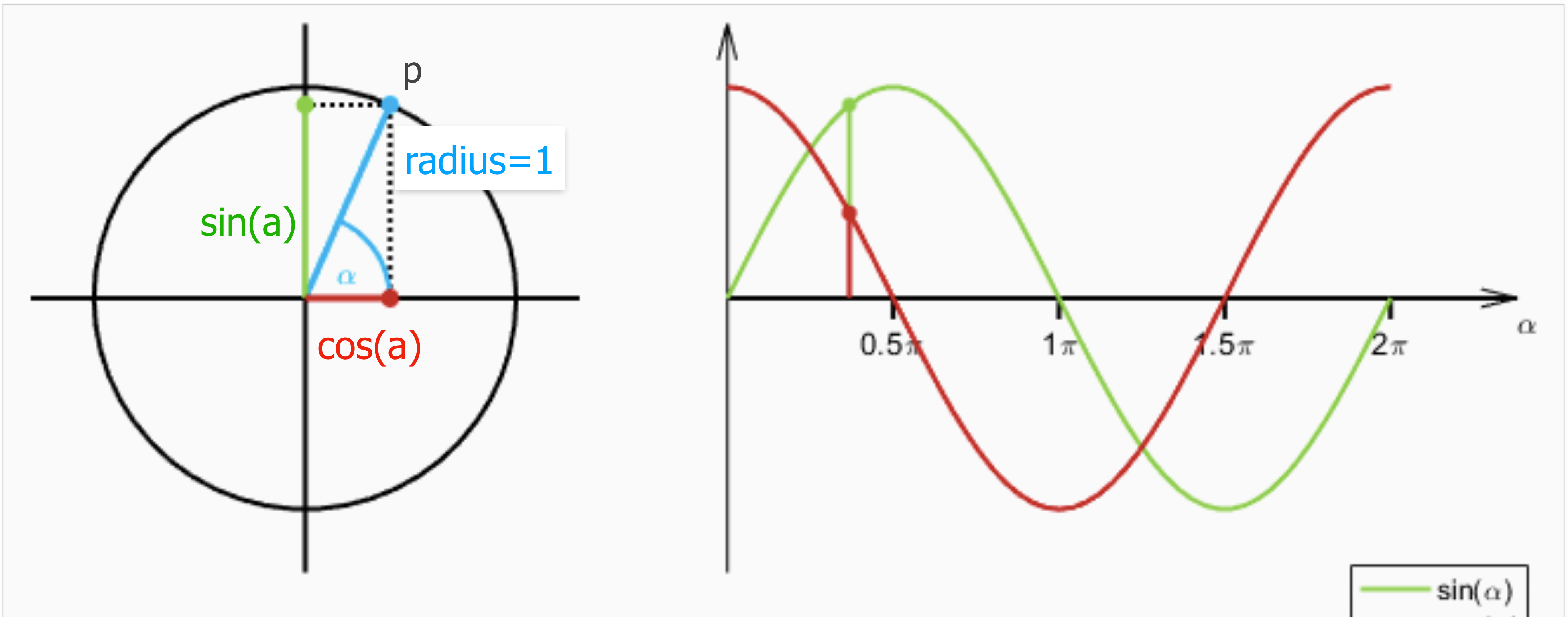



Image credit: https://upload.wikimedia.org/wikipedia/commons/f/f3/Sinus_und_Cosinus_am_Einheitskreis.gif

# Sound & Sine Waves

- The general form to create a sine wave, usually called **sinusoid**, is expressed by $\mathbf{y = A \sin(2\pi f t + \varnothing)}$

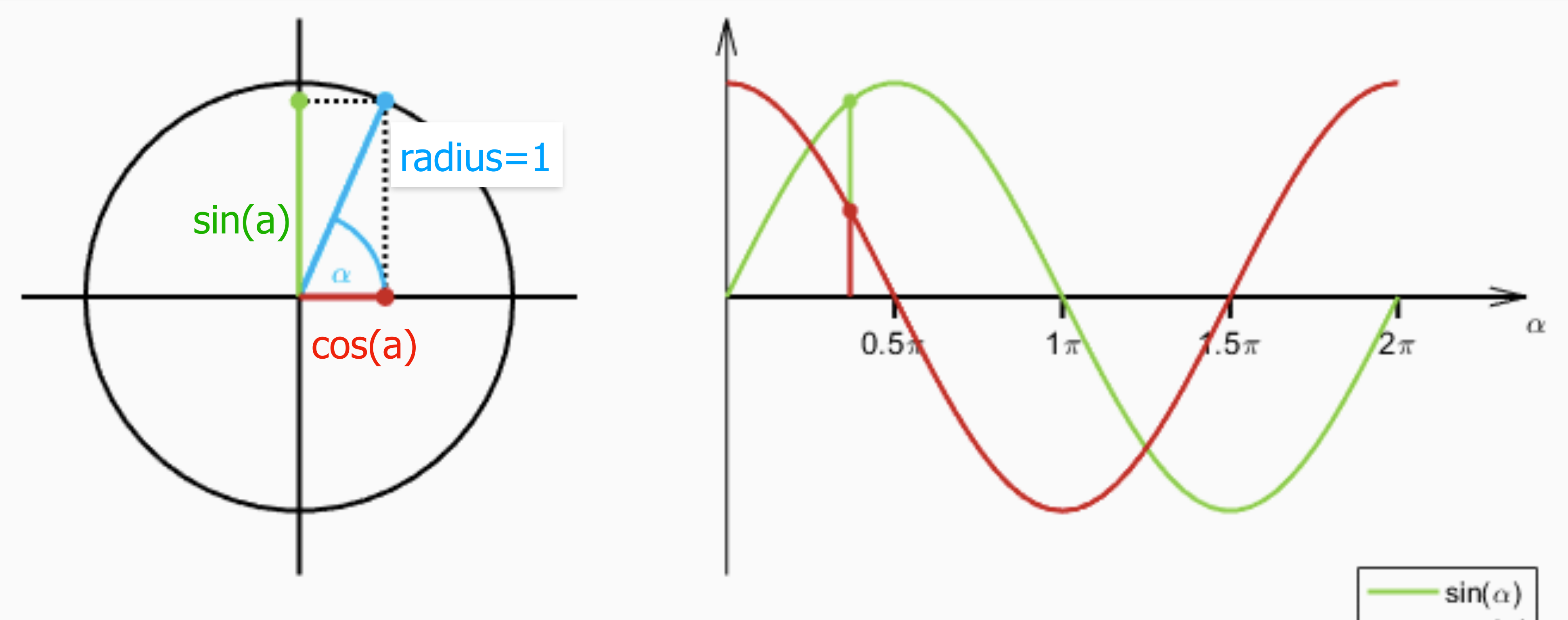



Image credit: https://upload.wikimedia.org/wikipedia/commons/f/f3/Sinus_und_Cosinus_am_Einheitskreis.gif

# Sound & Sine Waves

$y = A \sin(2\pi f t + ø)$

- ***A*** expresses the Amplitude
- **2π*f*** = 2π/T expresses the angular frequency *omega*
- **ø** expresses the phase shift, i.e., where the sinusoid starts at t=0

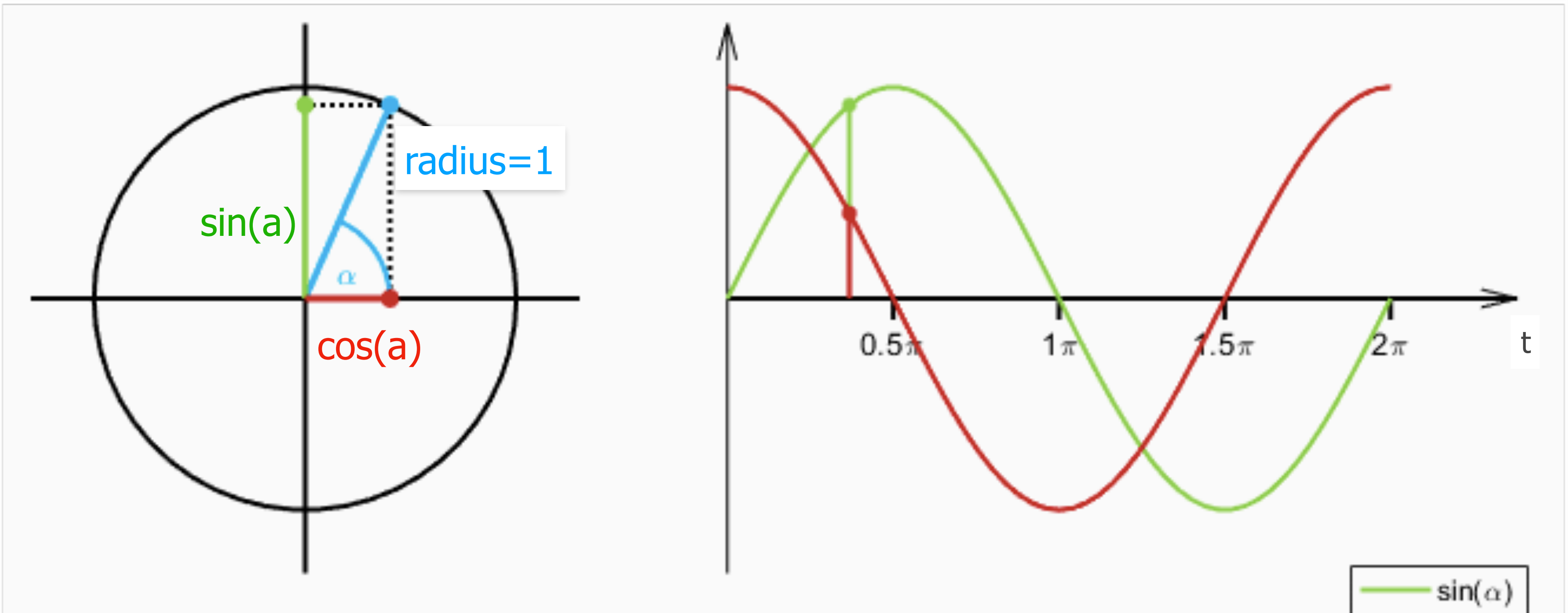



Image credit: https://upload.wikimedia.org/wikipedia/commons/f/f3/Sinus_und_Cosinus_am_Einheitskreis.gif

# Sound & Sine Waves

$$y = A \sin(2\pi f t + \varnothing)$$

- With these information we can now define a periodic signal at any future time t and at any audible frequency 20Hz < $f$ < 20kHz

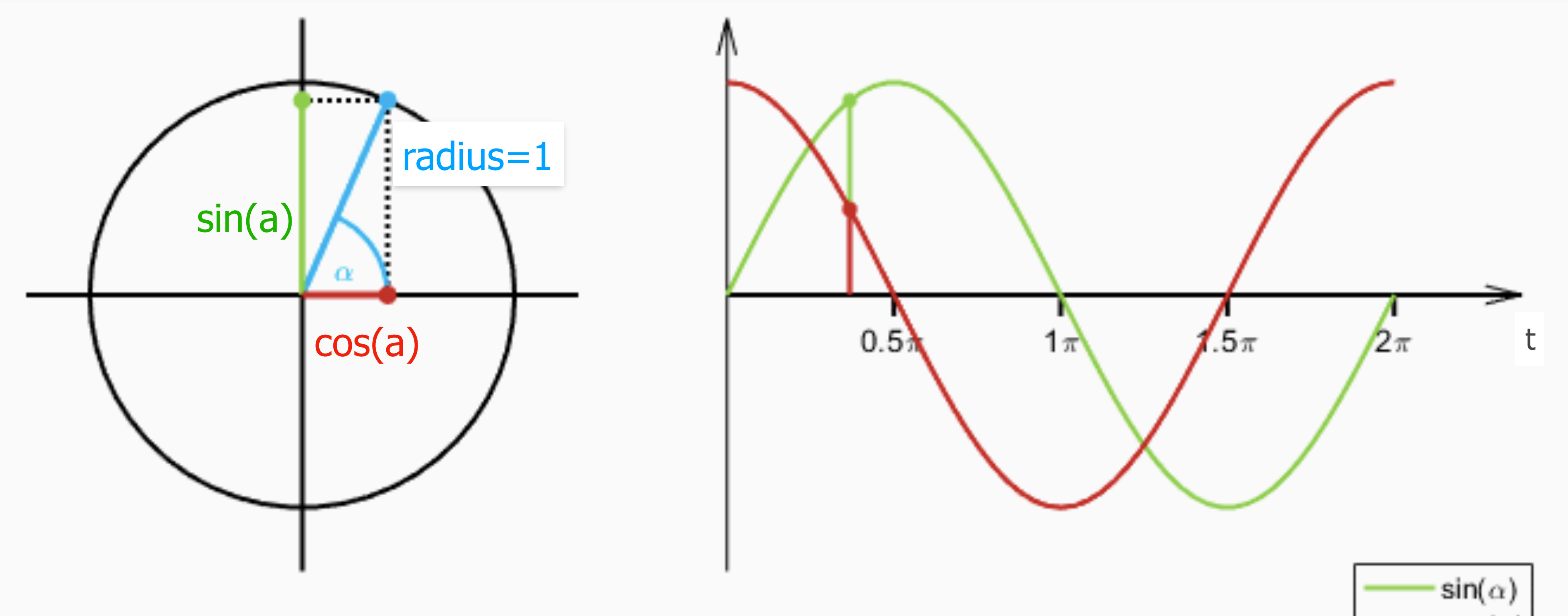



Image credit: https://upload.wikimedia.org/wikipedia/commons/f/f3/Sinus_und_Cosinus_am_Einheitskreis.gif

# Sound & Sine Waves

$y = A \sin(2\pi f t + ø)$

- For example:

  $y = 0.5 \sin(2\pi * 440 * t + ø)$ |=> for t = 0 to 120

  - *What is the result of this calculation???*

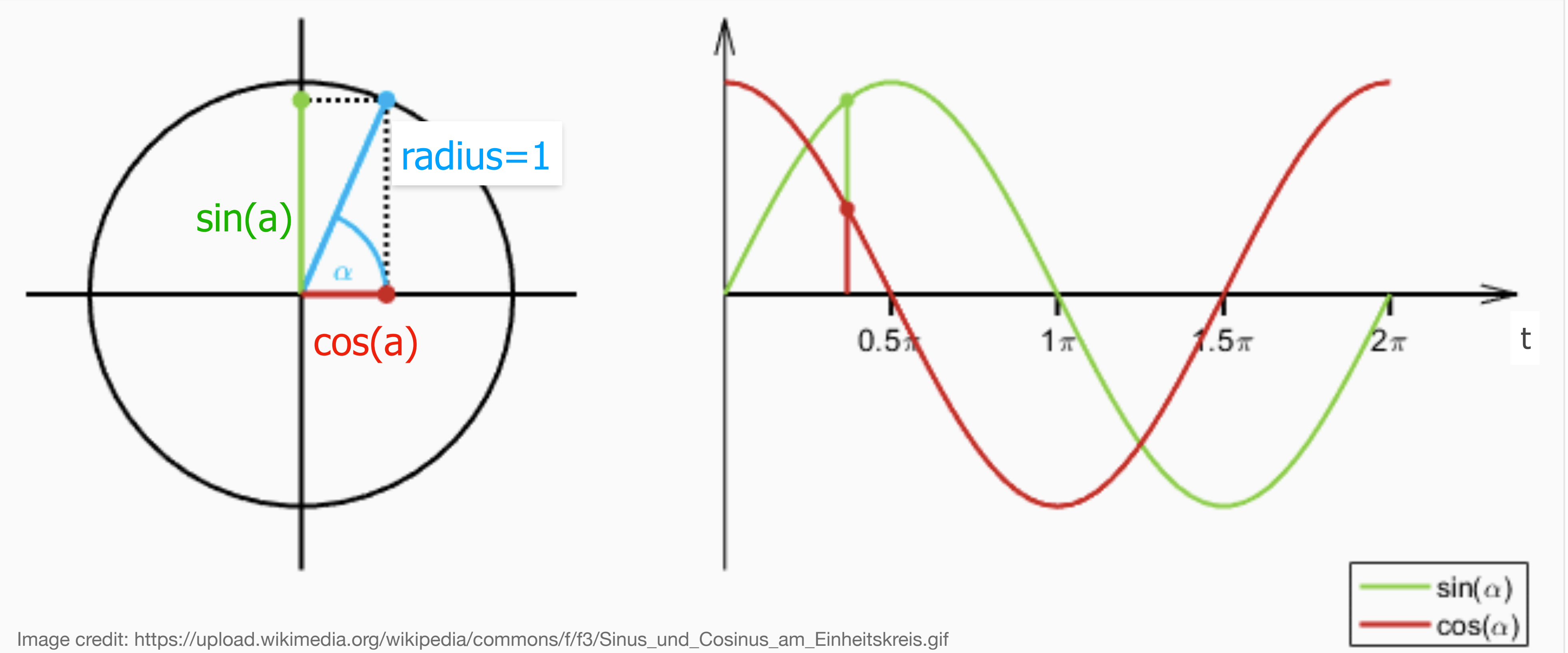



Image credit: https://upload.wikimedia.org/wikipedia/commons/f/f3/Sinus_und_Cosinus_am_Einheitskreis.gif

# Sound & Sine Waves

$$y = A \sin(2\pi f t + ø)$$

- For example:

  y = 0.5 sin(2π * 440 * t + ø) |=> for t = 0 to 120

  - *The result is a 2 min sound wave at frequency 440 Hz*

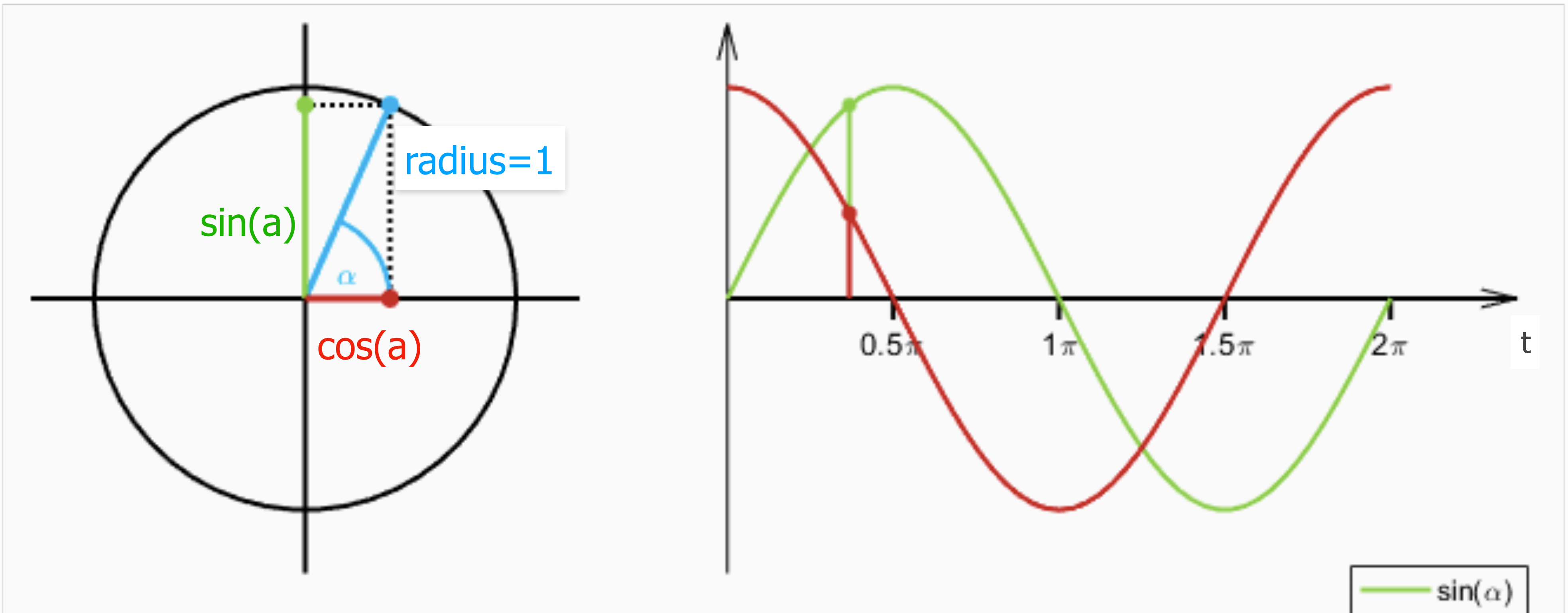



Image credit: https://upload.wikimedia.org/wikipedia/commons/f/f3/Sinus_und_Cosinus_am_Einheitskreis.gif

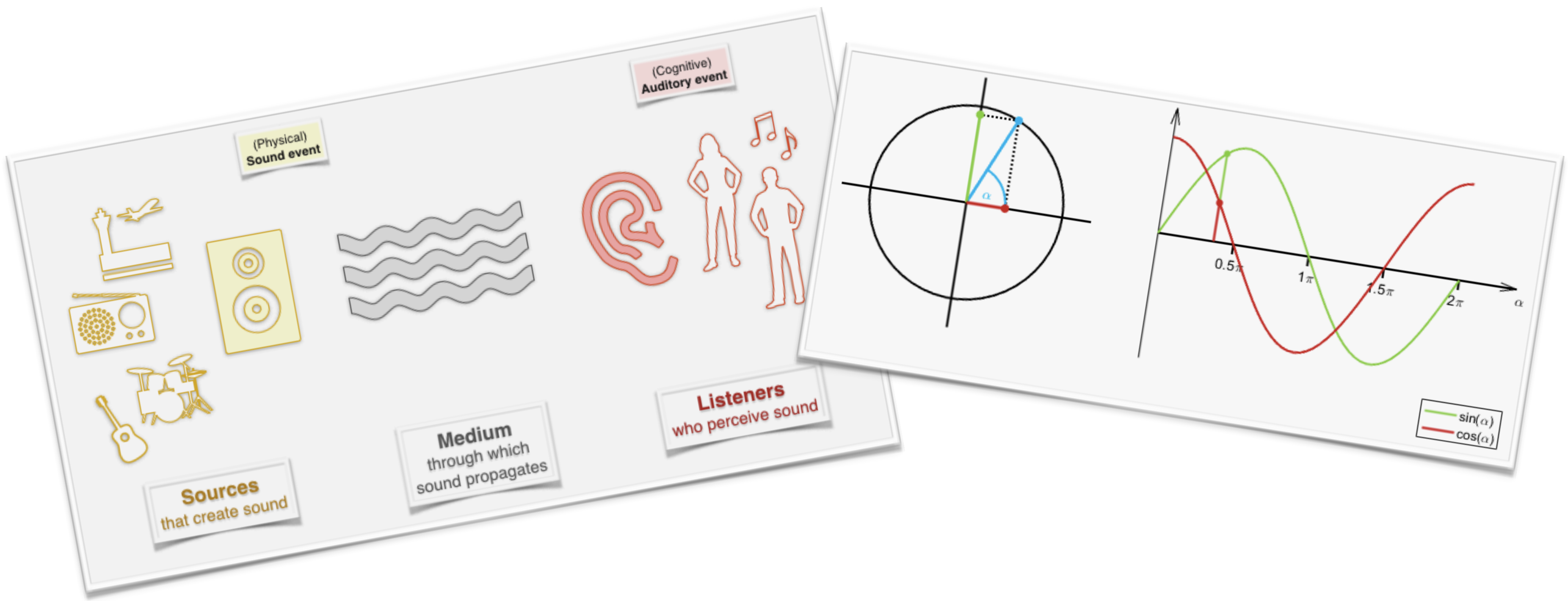
# Sound & Sine Waves

- How can we describe sound formally?
  - Sine waves are periodic functions that serve well to describe single frequency wave forms
  - The general form of a sine wave is referred to as sinusoid
  - Sinusoids can be used to create more complex multi-frequency wave forms — as we will see in the course of this lecture

# Digital Representation

# Digital Audio

Now that we have an idea of sound and how it can be described formally to generate sound waves —




how do we put this into *digital* action?

# Digital Audio

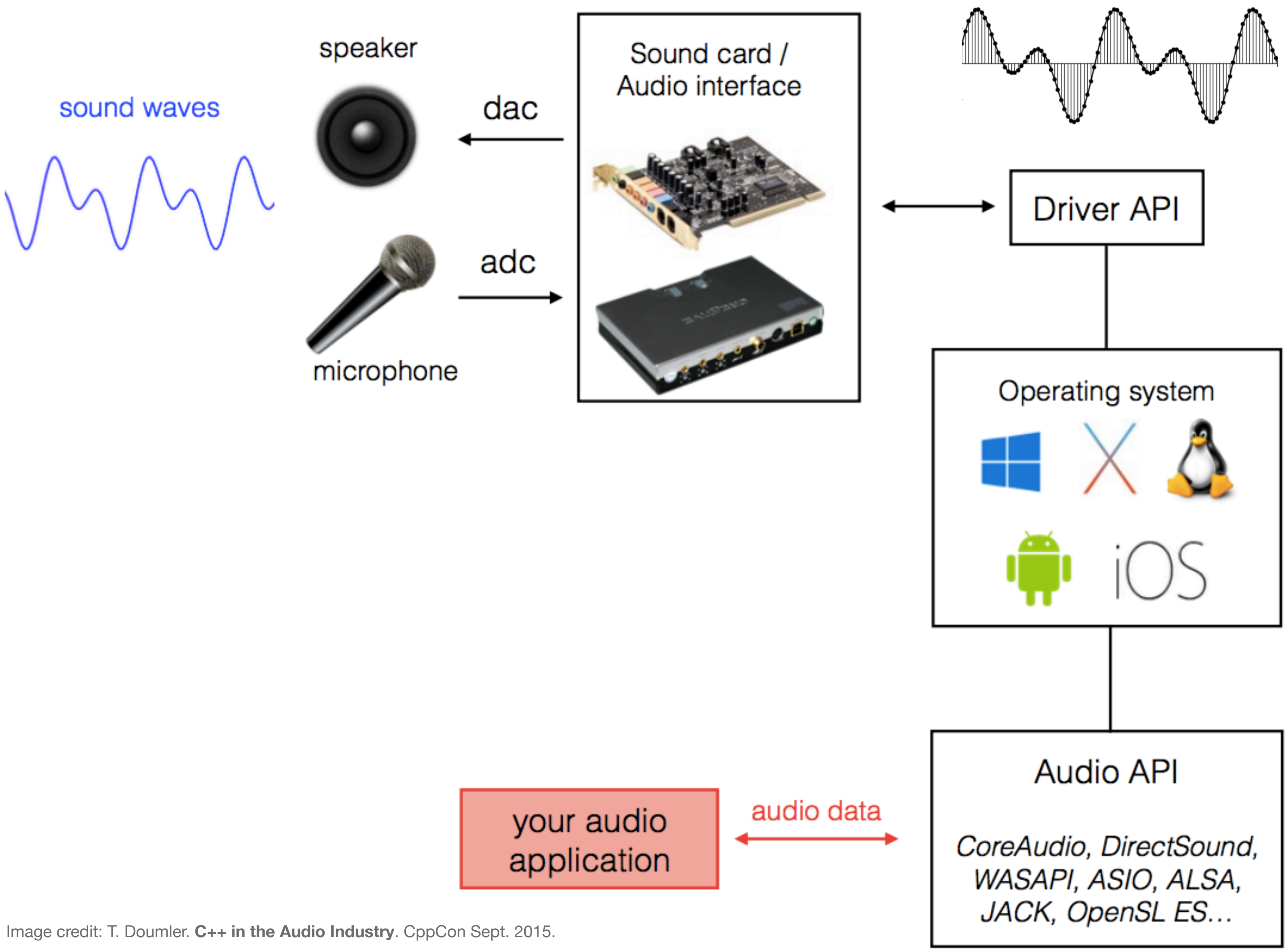
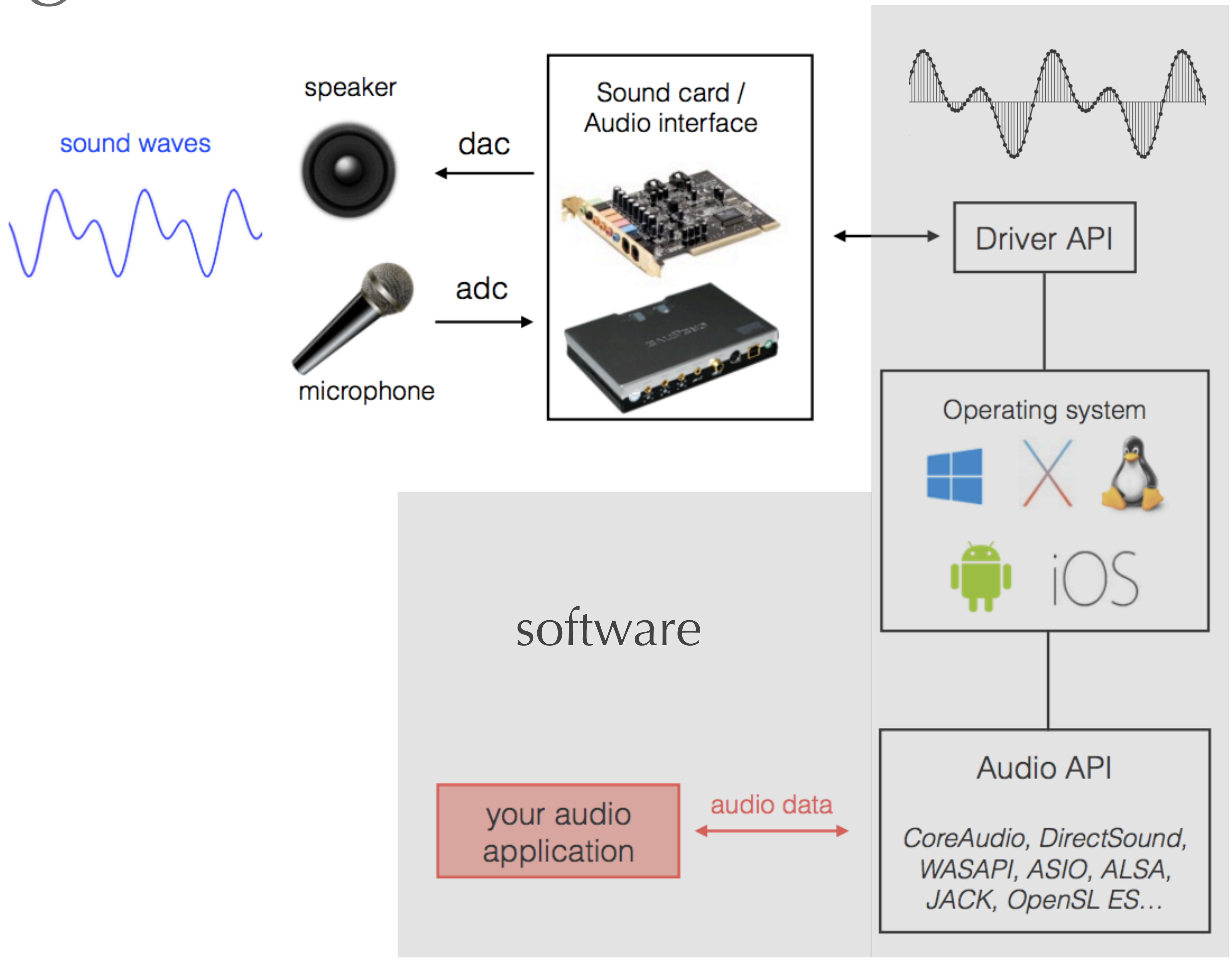
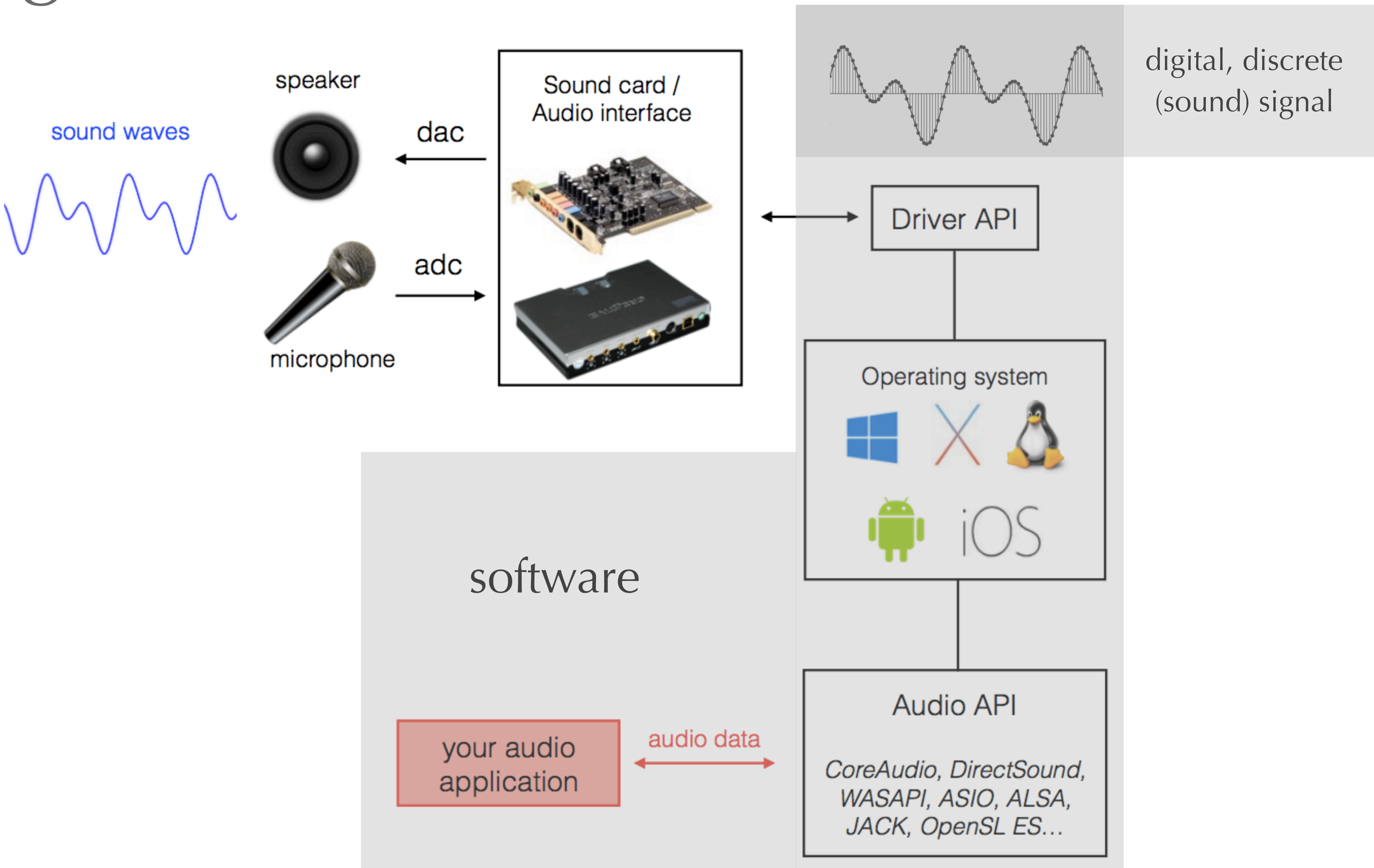



Image credit: T. Doumler. **C++ in the Audio Industry**. CppCon Sept. 2015.

# Digital Audio



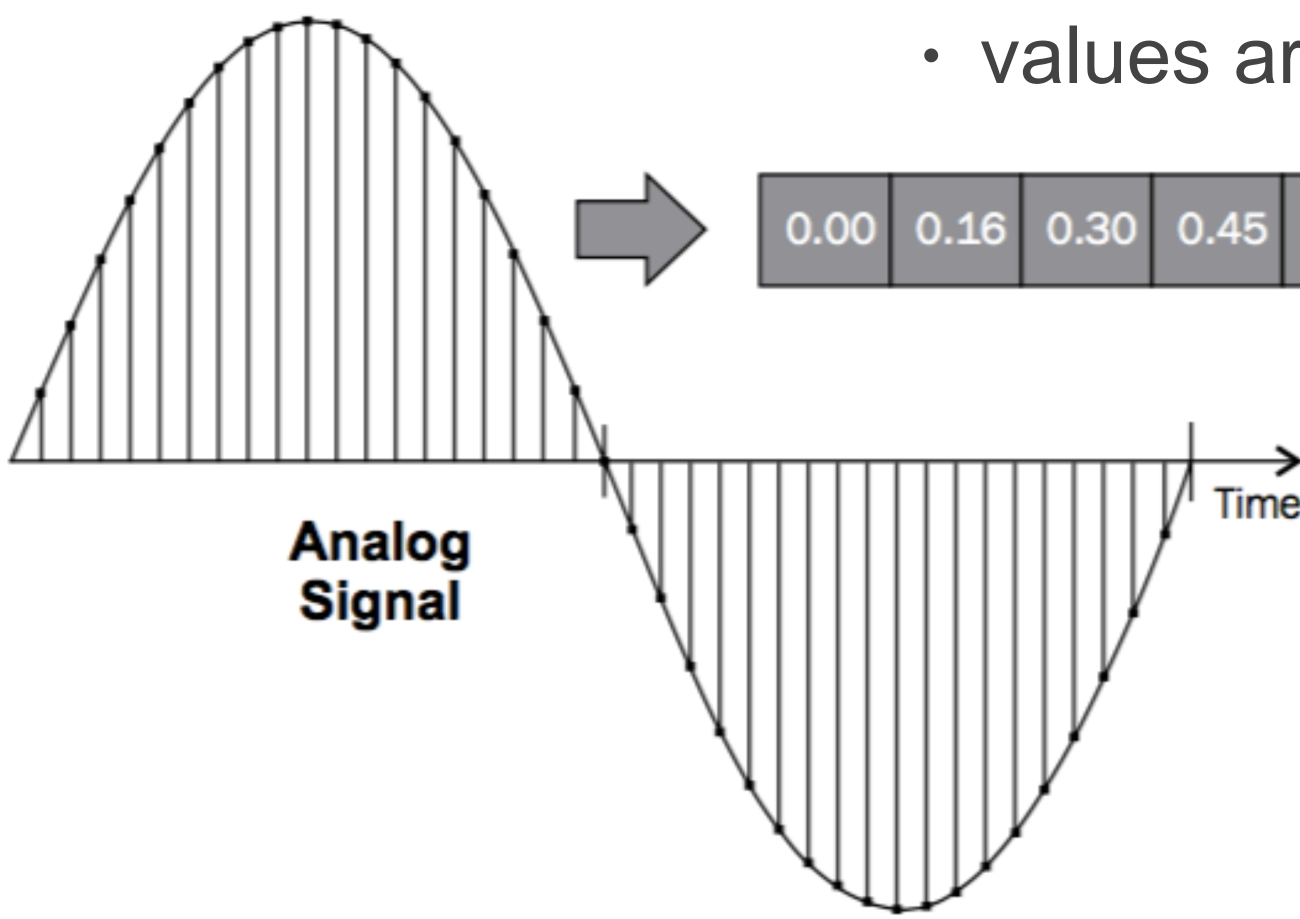
# Digital Audio

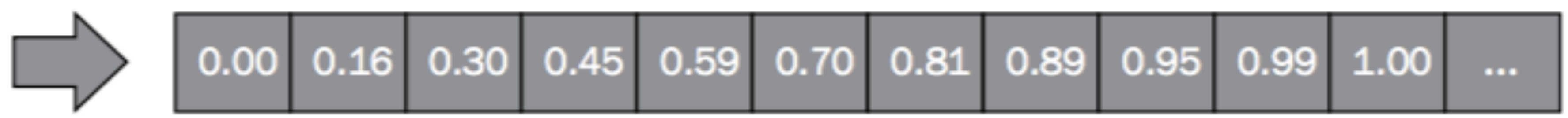


# Digital Audio

A digital (sound) signal is a list of discrete numbers in an array (buffer)

- each value represents an *amplitude*
- each position is a *discrete time event*
- values are *equidistantly spaced*




"audio buffer"

Image credit: Gouveia, D. (2013): Getting Started with C++ Audio Programming for Game Development. Birmingham, UK: Packt Publishing Ltd.

# Digital Audio

audio buffer is

- filled when sound is synthesised, or recorded & digitised
- read when digital sound is played back

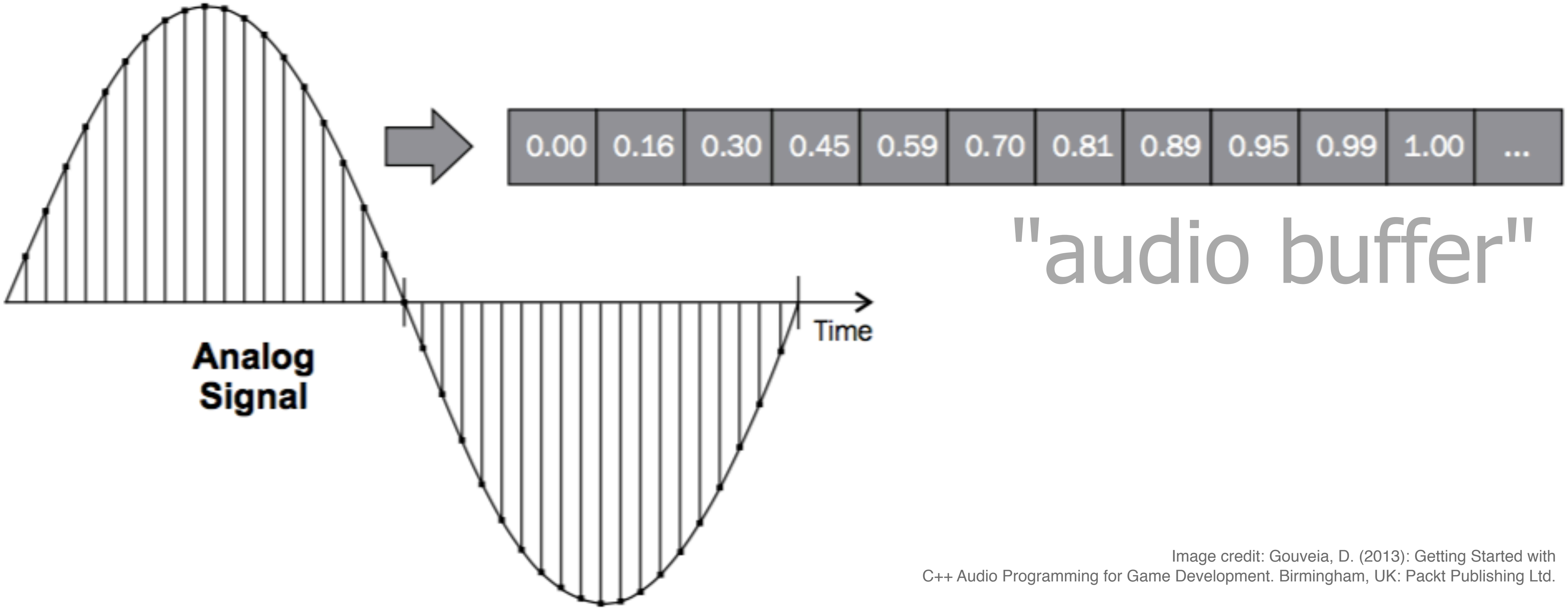
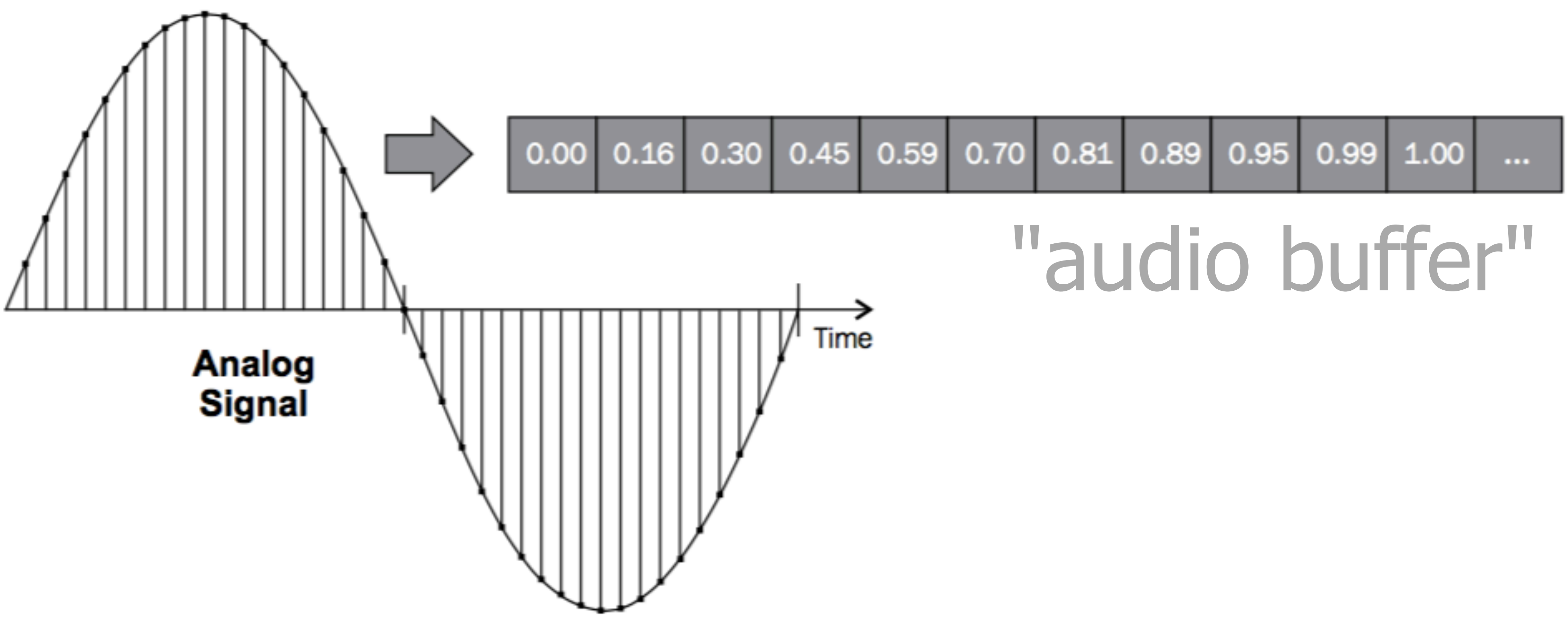



Image credit: Gouveia, D. (2013): Getting Started with C++ Audio Programming for Game Development. Birmingham, UK: Packt Publishing Ltd.

# Digital Audio

Digitisation consists of two steps

1. sampling
2. quantization



# Sampling

amplitude of analog signal is measured at fixed (equidistant) time intervals determining its *sampling rate* or *sampling frequency*

*sampling rate:* how many samples are "taken" per second

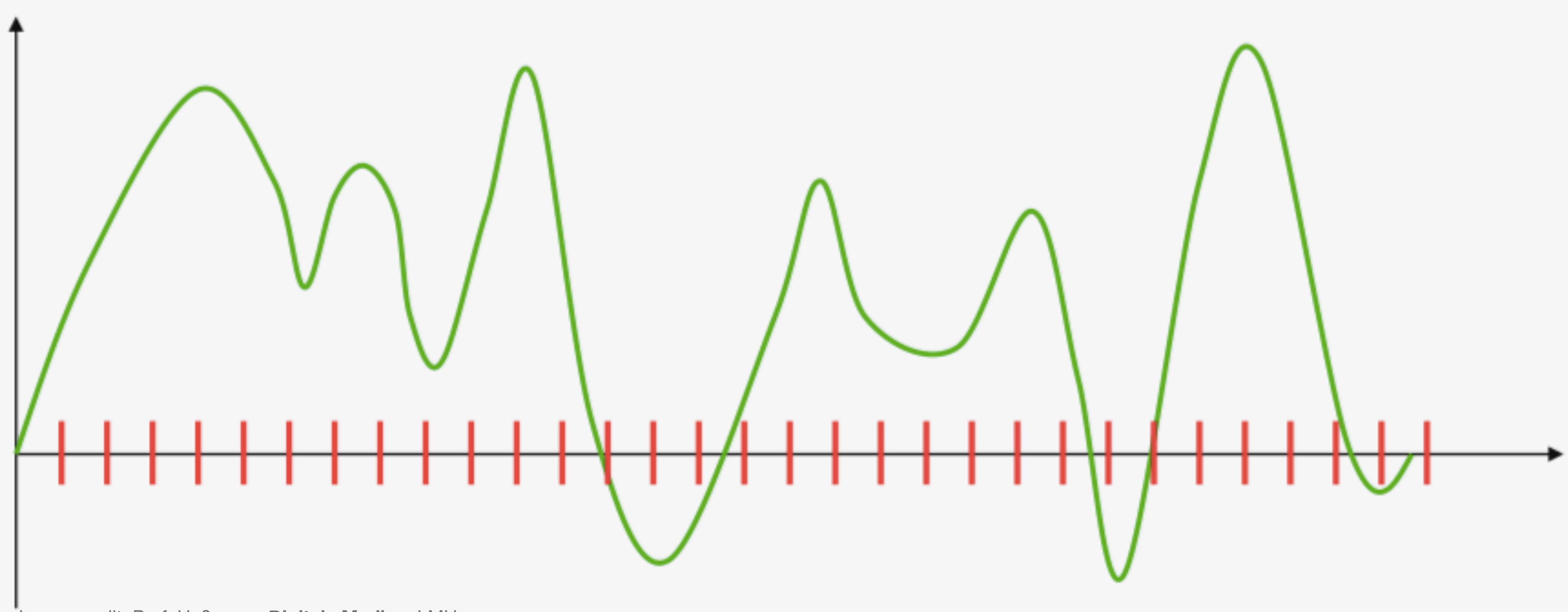


Image credit: Prof. Hußmann. **Digitale Medien**. LMU.
http://www.medien.ifi.lmu.de/fileadmin/mimuc/dm_ss04/dm2b.pdf

# Sampling

Standard sampling rates are

- 44.1kHz (Audio CD)
- 48kHz (film production)

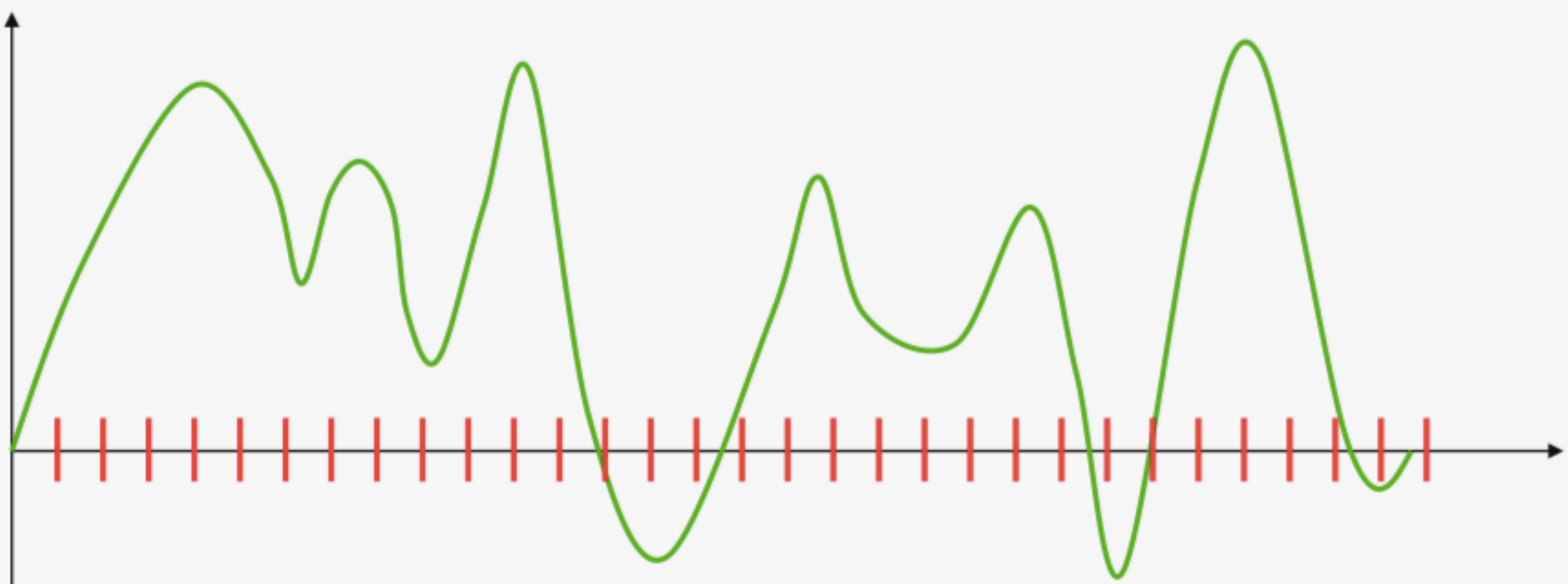


Image credit: Prof. Hußmann. **Digitale Medien**. LMU.
http://www.medien.ifi.lmu.de/fileadmin/mimuc/dm_ss04/dm2b.pdf

# Quantisation

- sampled amplitude values get mapped onto discrete values defined by the *bit depth (resolution)*
- values usually stored as int, float24 or float32

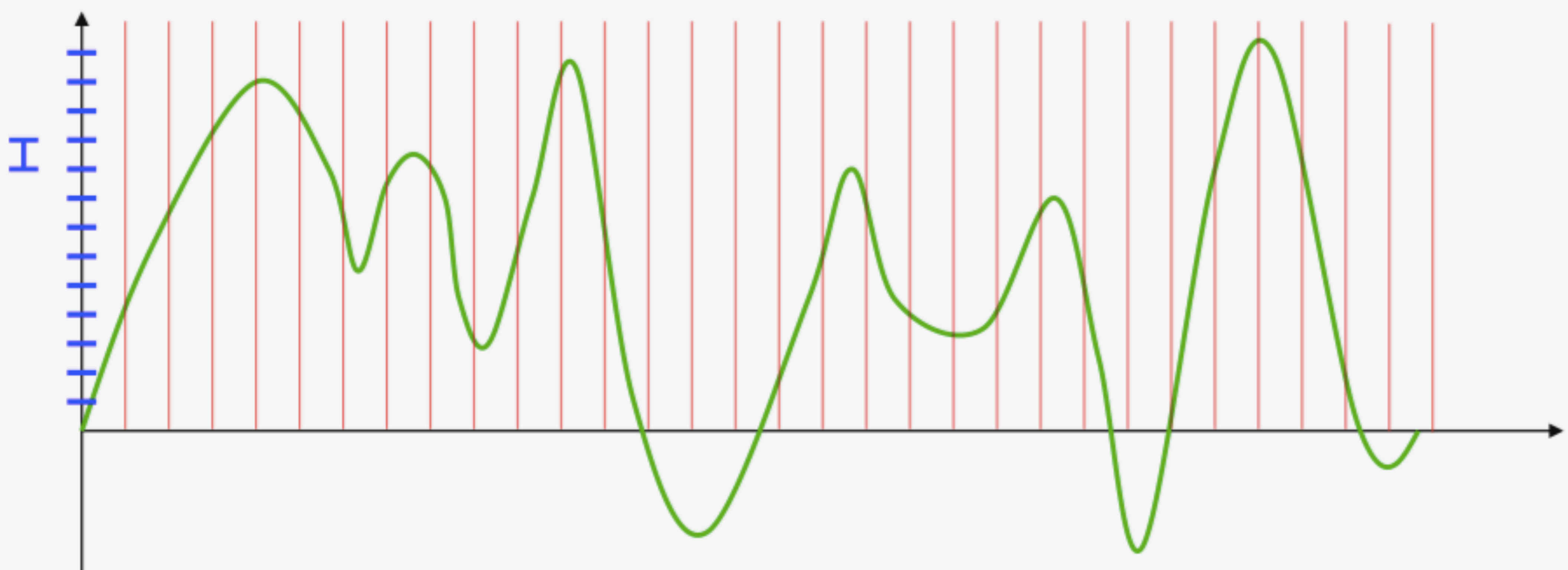


Image credit: Prof. Hußmann. **Digitale Medien**. LMU.
http://www.medien.ifi.lmu.de/fileadmin/mimuc/dm_ss04/dm2b.pdf

# Digital representation

- signal *quality* controlled by *sampling rate* and *bit depth*
- highest frequency that can be represented is sRate/2

  (Nyquist sampling theorem)

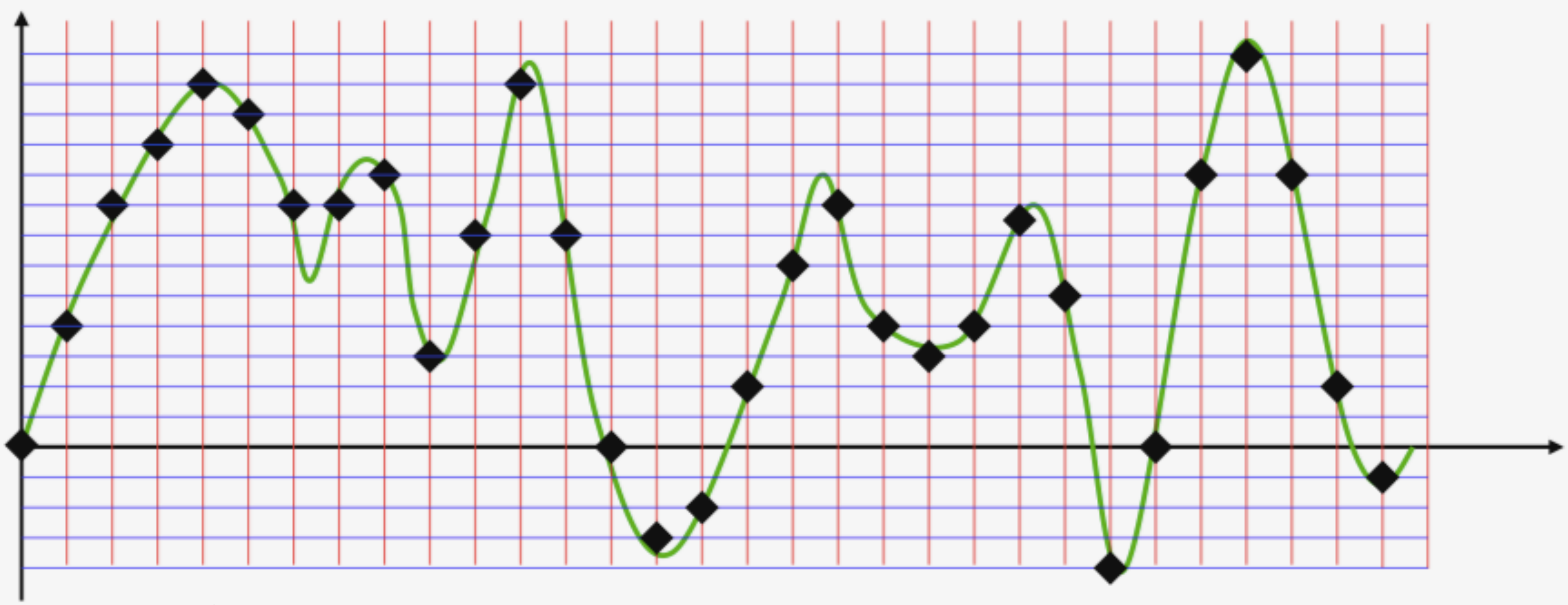


Image credit: Prof. Hußmann. **Digitale Medien**. LMU.
http://www.medien.ifi.lmu.de/fileadmin/mimuc/dm_ss04/dm2b.pdf

# Quantization

- Sampling & quantization always introduce a certain **digitization error**
- The array of numbers that is stored (or played back) is always an approximation of the original analog audio signal only

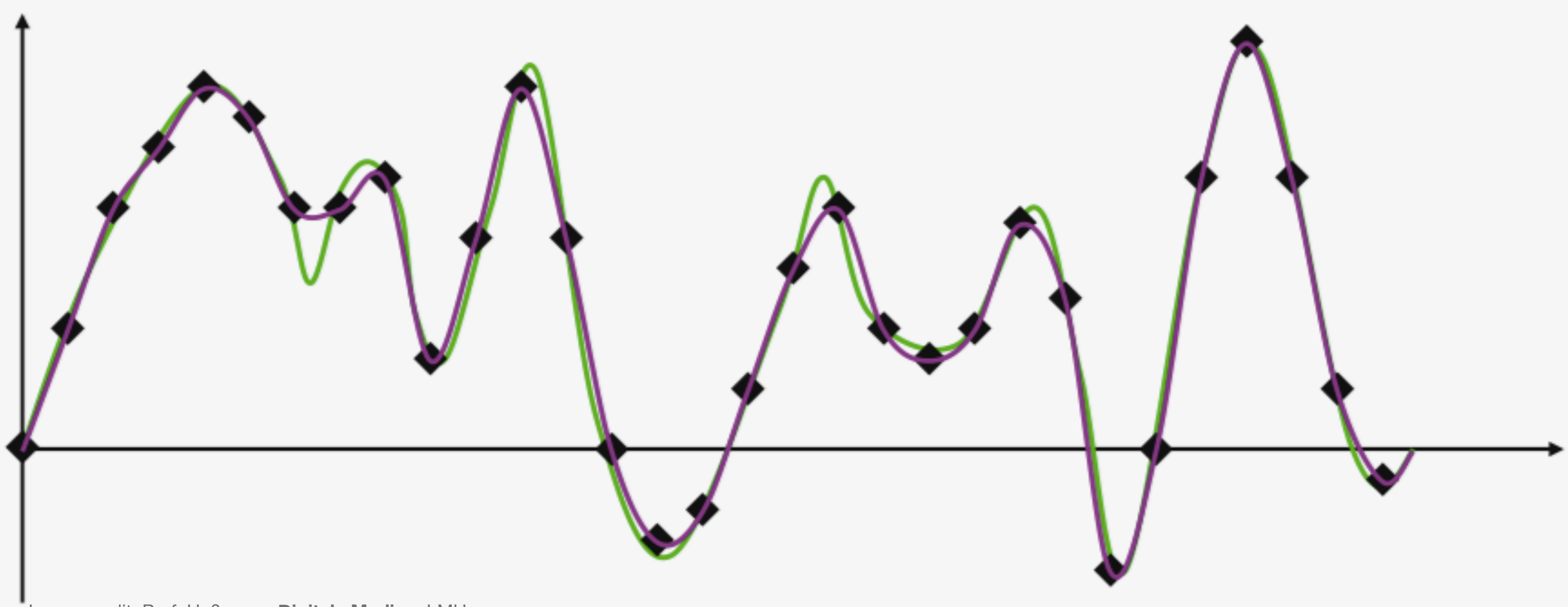
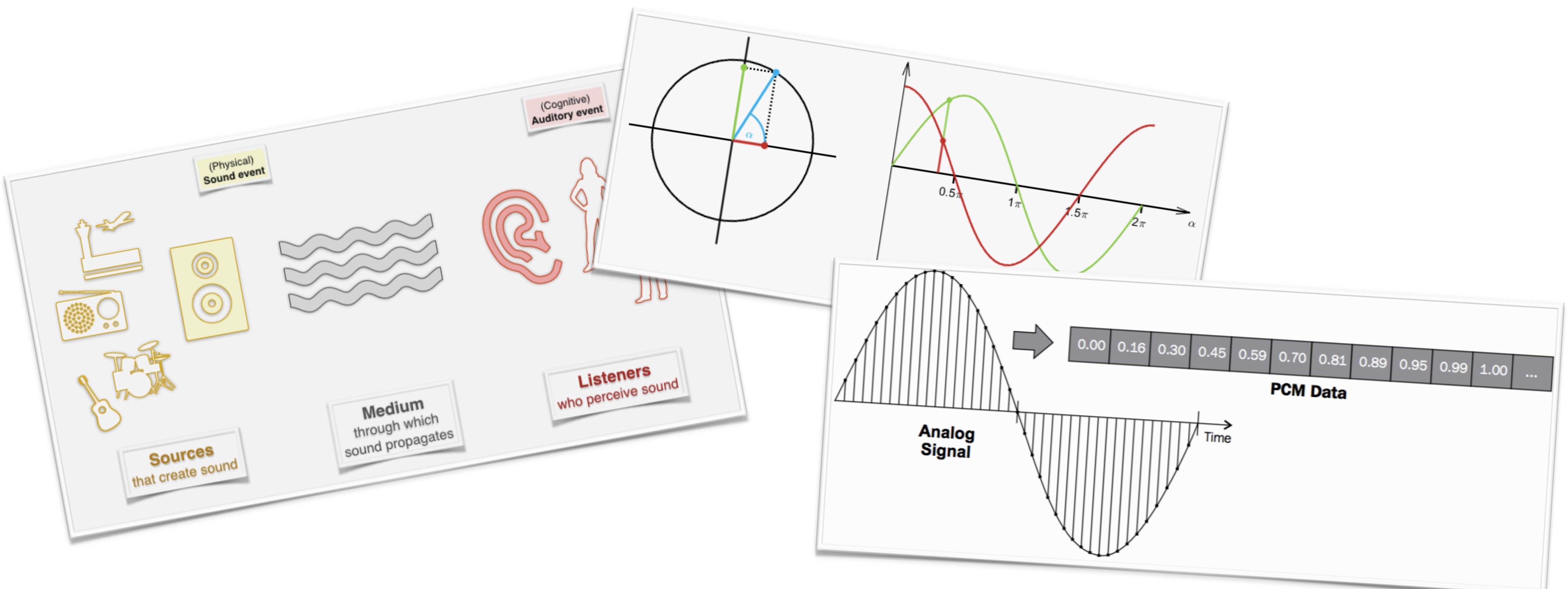


Image credit: Prof. Hußmann. **Digitale Medien**. LMU.
http://www.medien.ifi.lmu.de/fileadmin/mimuc/dm_ss04/dm2b.pdf

# Digital Audio

- What is the digital representation of sound?
  - Digital representation of sound is expressed by arrays of numbers
  - Proper signal reconstruction requires thousands of samples/sec
  - Digital sound signals are approximations of analog sound signals



# Audio Programming

# Generating A Digital Sound Signal

- Create a digital sound of **440 Hz** that lasts for **2 seconds** and plays back at a frequency of **44100 Hz** (sampling rate)

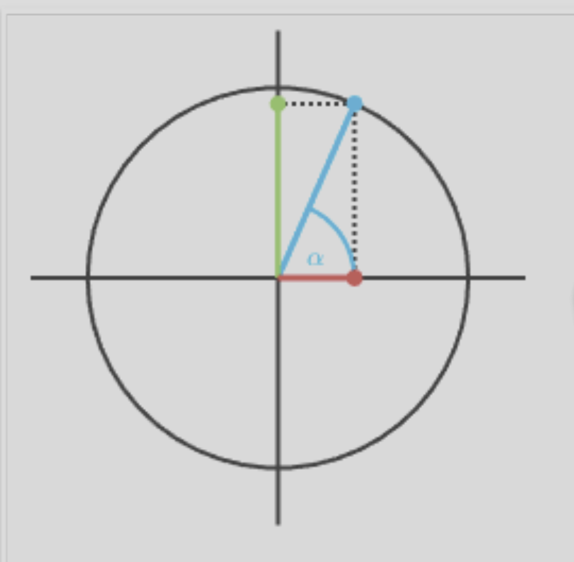


Image credit: https://upload.wikimedia.org/wikipedia/commons/f/f3/Sinus_


$$y = A \sin(2\pi f t + \phi)$$

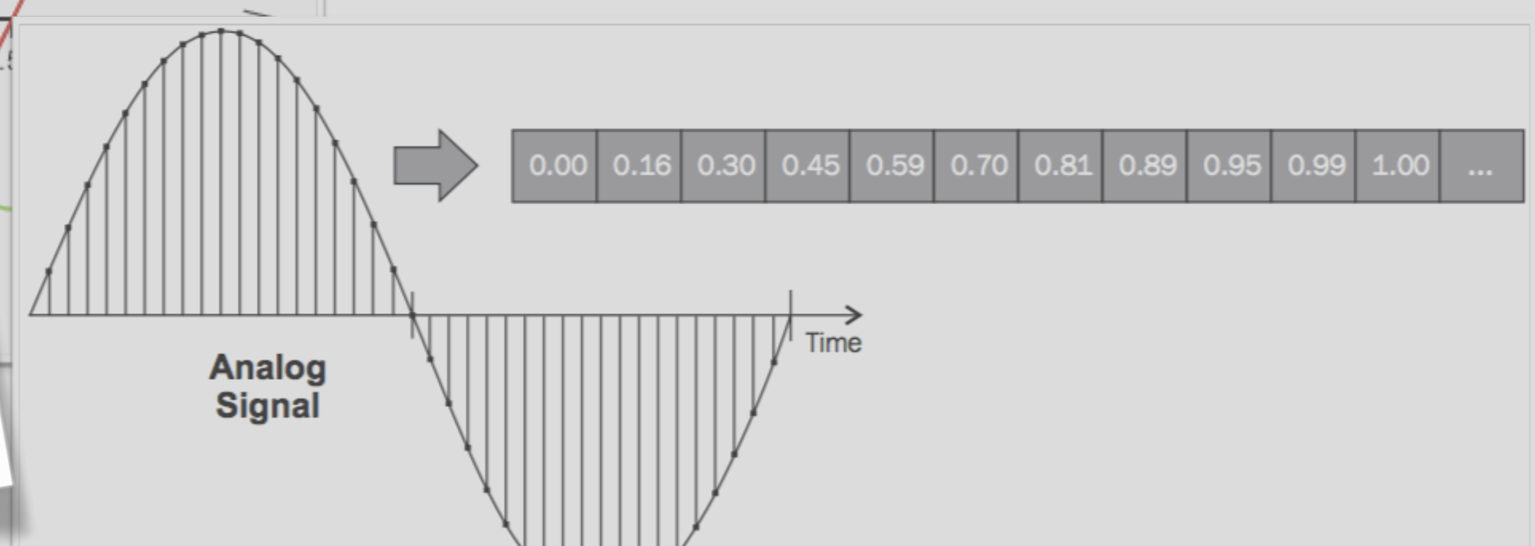



Image credit: Gouveia, D. (2013): Getting Started with C++ Audio Programming for Game Development. Birmingham, UK: Packt Publishing Ltd.

# Generating A Digital Sound Signal

- Create a digital sound of **440 Hz** that lasts for **2 seconds** and plays back at a frequency of **44100 Hz** (sampling rate)

$$y = A \sin(2\pi f t + \varnothing)$$

- Todo
  - Create an audio buffer of size …
  - Fill the audio buffer with sine wave values at frequency 440 Hz

# Generating A Digital Sound Signal

- Create a digital sound of **440 Hz** that lasts for **2 seconds** and plays back at a frequency of **44100 Hz** (sampling rate)

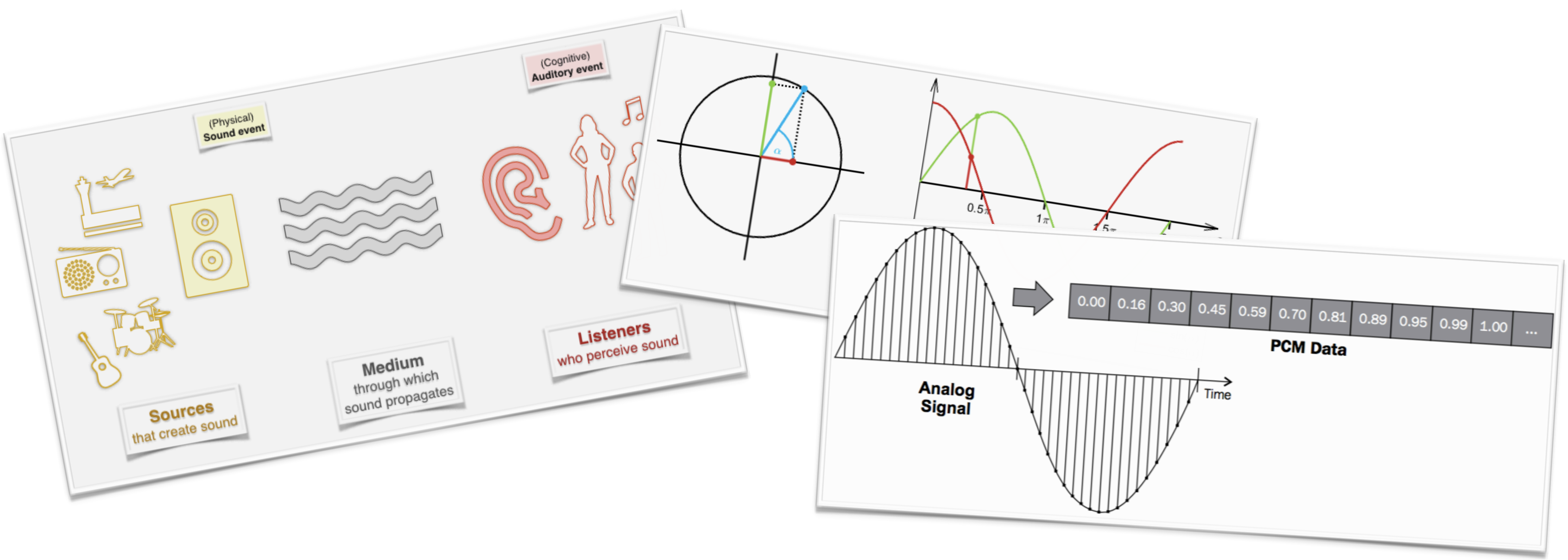
```
audioBuffer[44100 * 2];

for (t= 0; t< 88200; t++) {
  A = 1;
  y = A * sin (2π * 440 * (t / 44100));
  audioBuffer[ t ] = y;
}
```

$$y = A \sin(2\pi f t + \varnothing)$$

# Generating A Digital Sound Signal

- How do we work with „digital sound“?
  - The audio buffer is the central type of data that we can work with
  - The audio buffer contains amplitude values of a digitized or digitally generated sound wave
  - The sampling rate is crucial as it represents our time unit



# Take Home Message

1. What is sound, actually?
2. What are its fundamental physical aspects?
3. How do we describe sound formally?
4. What is the digital representation of sound?
5. How do we work with „digital sound“?

Further Reading

# Literature


- Bernadini, N. & de Poli, G. (2007): The Sound and Music Computing Field: Present & Future. Journal of new Music Research, vol. 36.
- Burg, J., Romney, J. & Schwartz, E. (2014): Digital Sound and Music. Concepts, Applications, and Science. http://digitalsoundandmusic.com
- Burk, P., Polansky, L., Repetto, D., Roberts, M. & Rockmore, D. (2011): Music and Computers. A Theoretical and Historical Approach. http://cmc.music.columbia.edu/MusicAndComputers/
- Gouveia, D. (2013): Getting Started with C++ Audio Programming for Game Development. Birmingham, UK: Packt Publishing Ltd.
- Howard, D.M. & Angus, J.A.S. (2009): Acoustics and Psychoacoustics. 4th Edition. Oxford, UK: Focal Press.
- Papula, Lothar (2014): Mathematik für Ingenieure & Naturwissenschaftler Band 1. 14. überarb. Aufl., Wiesbaden: Springer Vieweg.
- Parviainen, T. (2016): Learn Web Audio from the Ground Up, Part 1: Signals and Sine Waves. http://teropa.info/blog/2016/08/04/sine-waves.html
- Smith, S. W. (2011): The Scientist and Engineer's Guide to Digital Signal Processing. http://www.dspguide.com
- Sethares, W.A. (2005): Tuning, Timbre, Spectrum, Scale. 2nd Edition. London, UK: Springer-Verlag.