# Intelligent Signal Processing (CM3065)

**Course Notes**

Felipe Balbi

April 13, 2021

# Contents

# Week 1

Key Concepts

- Use an audio editor and write simple programs to work with digital audio signals
- Explain the relevance of bit depth and sampling rates for digital audio signals and select appropriate parameters for different types of signals
- Describe the characteristics of sound waves and how they are perceived by humans

## 1.004 Technical information and tools used in this module

- Audacity
- P5.js
- Python
- ffmpeg

## 1.101 What is sound?

From a physics perspective, sound is a form of mechanical energy *produced* by vibrating matter. It needs a medium to propagate. In contrast to electromagnetic energy, sound cannot travel in a vacuum.

From a physiological or psychological perspective, sound is the reception of these waves and its processing by the brain. We capture sound waves with our inner ear and the brain converts it into what we call sound.

## 1.102 Characteristics of sound waves

In figure 1 below, we see several sine waves. Remember that the sine wave equation is given by $f(x) = A\sin(2\pi ft + \phi)$ where $A$ is the amplitude, $f$ is the frequency, $t$ is the time, and $\phi$ is the phase.

The speed of sound waves depends on the temperature, elasticity, and density of the medium the sound is travelling through. As an example, at $0\,^{\circ}\mathrm{C}$ the speed of sound travelling through the air is around $331\,\mathrm{m\,s^{-1}}$. When the air is around $20\,^{\circ}\mathrm{C}$ the speed of sound is around $343\,\mathrm{m\,s^{-1}}$.

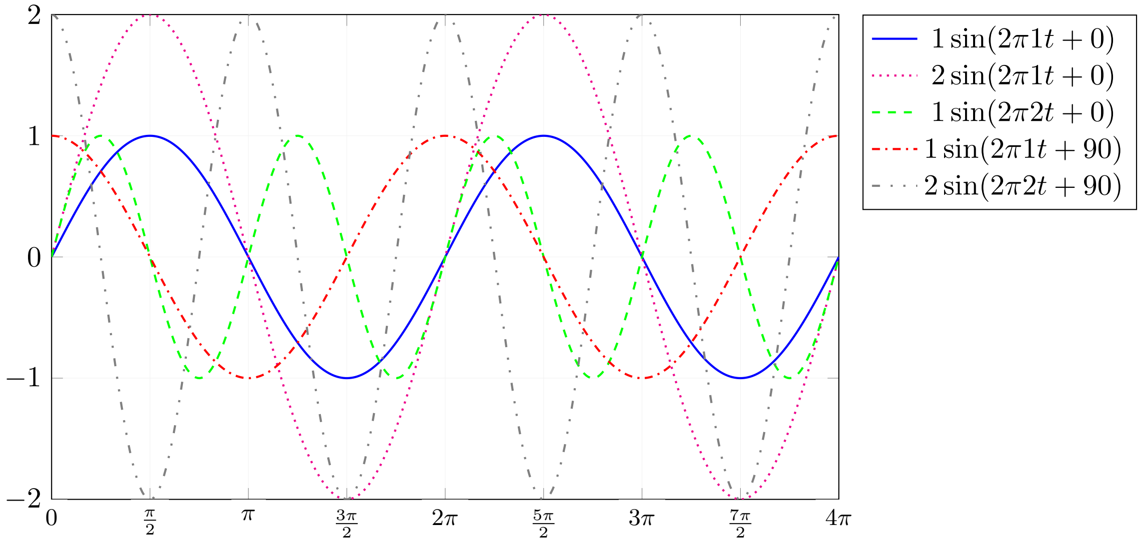


Figure 1: Sine waves

Waves can be represented in the domain of space. In this scenario we can characterize the wave by its wavelength and amplitude. The wavelength ($\lambda$) is given by the distance between two subsequent crests or troughs on the wave. The amplitude is given by the distance from the x axis to a peak or trough of the wave. All of these can be seen in figure 2 below. The pressure is measured in pascals (Pa) although we use dB for measuring sound amplitude.

Waves can also be represented in the domain of time. In this scenario we can characterize the wave by its amplitude, frequency, and period as shown in figure 3 below. Note that frequency, measured in hertz (Hz), is given by the inverse of the period, i.e. $f = \frac{1}{T}$ where $T$ is the period in seconds.

There is a direct relation between sound speed, wavelength, and frequency given by:

$$v = f \cdot \lambda$$

For example, if the wavelength is $\lambda = 2\,\mathrm{m}$, $v = 4\,\mathrm{m\,s^{-1}}$, then:

$$\begin{aligned}
4\,\mathrm{m\,s^{-1}} &= f \cdot 2\,\mathrm{m} \\
f &= \frac{4\,\mathrm{m\,s^{-1}}}{2\,\mathrm{m}} \\
f &= 2\,\mathrm{Hz}
\end{aligned}$$

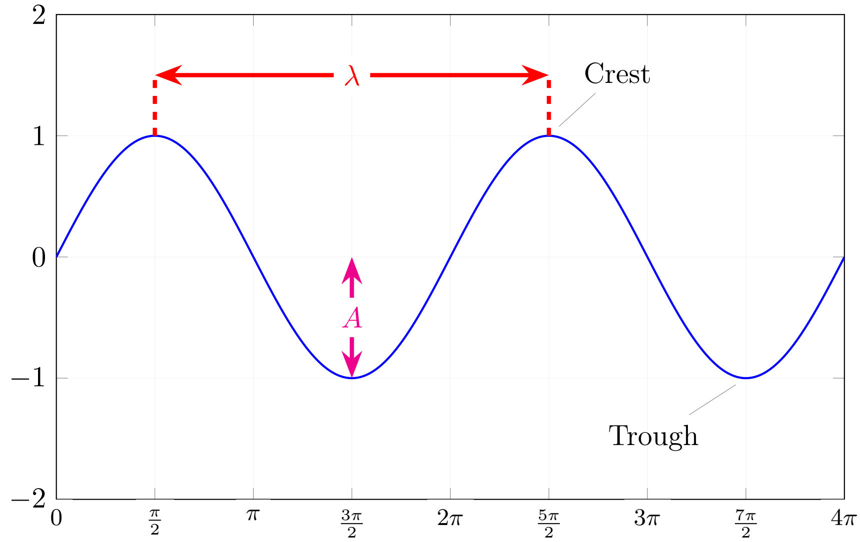


Figure 2: Sine wave in space domain

## 1.104 Human sound perception

Humans perceive sound through *pitch*, *loudness*, and *timbre*. In table 1 we summarize some of the properties of sound, both physical and psychological.

Table 1: Physical vs Psychological Properties of Sound

| **Physical** | **Psychological** |
|---|---|
| Frequency | Pitch |
| Amplitude | Loudness |
| Waveform | Timbre |
| Wavelength | |
| Period | |
| Duration | |

We can say that there is a relation between the matching pairs in table 1. That is, *Frequency* is related to *Pitch*, *Amplitude* is related to *Loudness*, and *Waveform* is related to *Timbre*. The relation, however, is not linear. That is, a certain change in frequency does not correspond to a similar change in pitch.

What is said above is a simplification of reality. A change in frequency also has an impact in loudness and timbre.

*Pure Tone* is a sound wave composed of a single sine wave. In figure 4 we have a representation of a pure tone while in figure 5 we have a representation of a composition of several sine waves.

The relationship between Pitch and Frequency is non-linear. In practice this means that as the pitch increases, a greater change in frequency is required. Another impor-

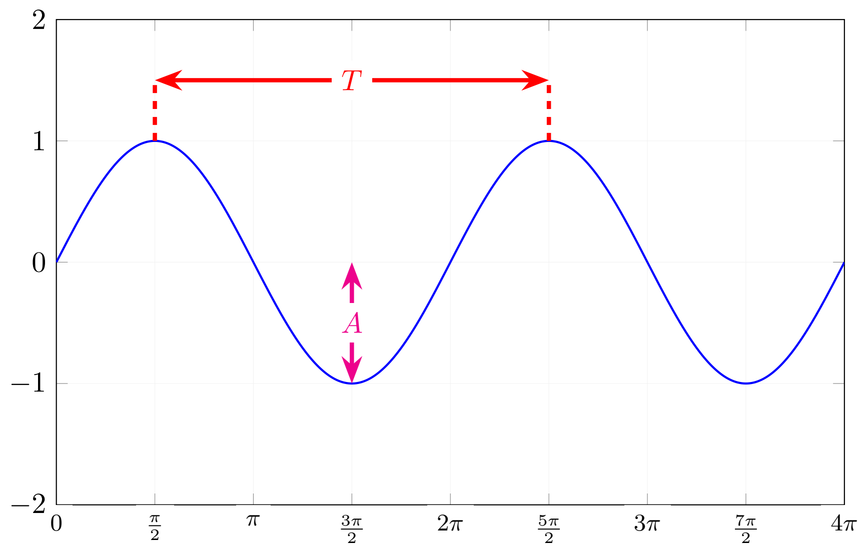


Figure 3: Sine wave in time domain

tant details is that humans can hear sound within the range from 20 Hz to 20 000 Hz. Frequencies below 20 Hz are known as infra-sound and frequencies above 20 000 Hz are known as ultrasound.

Loudness is a sensation to the perception of amplitude in sound waves. The pressure is measured in pascals (Pa). A quiet bedroom as night would measure sound pressure on the order of 630 µPa.

It turns out that measuring sound pressure in pascals is not very convenient, because of that we generally use a logarithmic scale called *dB SPL*. The conversion for pascal to dB is given by:

$$SPL = 20 \log_{10} \left( \frac{p}{p_{ref}} \right) dB$$

Where $p_{ref} = 20\,\mu\text{Pa}$, for sound pressure in air.

Timbre, or sound quality, helps us differentiate between two sounds with the same frequency and amplitude. Timbre is related to the waveform and the sound spectrum. In summary, we can hear the different component sound waves of a sound and interpret it as timbre.

## 1.106 Audio fundamentals

- Hosken, D. W. An introduction to music technology. (New York: Routledge, 2011). – Chapter 1: What is sound (pp.7–9) – Chapter 2: Sound properties and the waveform view (pp.17–26)

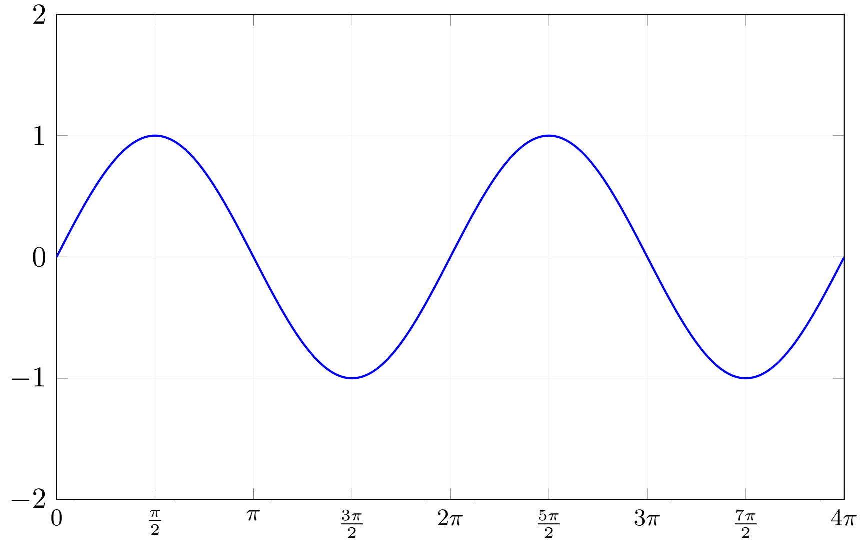


Figure 4: Pure Tone

## 1.201 The digital world: sample rate and bit depth

A sound level meter uses a microphone to convert air pressure changes into an electrical signal measured in volts. This is the first step to digitize a sound wave.

The signal output of a microphone is very small, therefore we need to amplifier to amplify the signal before passing it along to an Analog to Digital Converter, or ADC. The ADC measures the incoming voltage periodically and converts each sample to a numerical value within its dynamic range.

The sampling rate is the rate at which the ADC is *sampling* the input voltage. The sampling rate is given in Hertz (Hz). Each measurement of the input signal's amplitude is called a *sample*, the faster we sample, the better the quality, but the resulting amount of data per second will be larger. Table 2 below gives a set of common sampling rates and where they're used.

Table 2: Common Sampling Rates

| **Application** | **Sample Rate** |
|---|---|
| CD | 44.1 kHz |
| DVD | 48 kHz |
| Professional Audio (?) | 88.2 kHz |
| Professional Audio (?) | 96 kHz |

The Nyquist-Shannon Sampling Theorem states that the sampling rate must be at twice the highest frequency to be sampled. In other words, we can say that the Nyquist Frequency is $\frac{samplerate}{2}$ and any frequency above the Nyquist Frequency will not be recorded properly by the ADC, introducing artificial frequencies through aliasing.

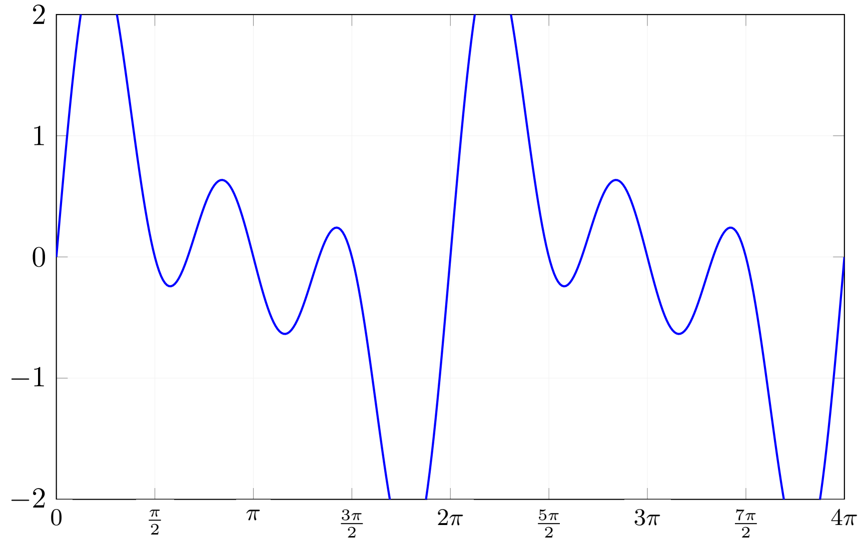


Figure 5: Composite Tone

To counter the problem of aliasing, most audio systems including an anti-aliasing frequencies which are basically a low-pass filter that cuts off frequencies above the Nyquist Frequency of the system.

Bit Depth is the number of bits used to record the samples. It's also referred to as Dynamic Range, sample width, or quantization level. The more bits we use, the more accurately we can measure the analogue waveform, but it results in more memory usage per sample. Common bit widths for digital audio are 8, 16, 24, and 32 bits.

Clipping occurs when the sampled amplitude is outside the ADC's dynamic range, then the value will either be clipped to 0 or maximum allowed value. Clipping generates perceivable distortions in the audio and should be avoided.

## 1.203 Digital audio representation: the time domain

There are two ways to represent audio:

**Time Domain** plotted as amplitude vs time

**Frequency Domain** plotted as amplitude vs frequency

Amplitude is usually represented in *dbFS* or db Full Scale. The relation is given by:

$$dBFS = 20\log_{10}(|value|)$$

## 1.205 Digital audio

- Hosken, D. W. An introduction to music technology. (New York: Routledge, 2011). – Chapter 4: Analog and digital (p.51) – Chapter 4: The audio recording path (pp.51–53) – Chapter 5: Digital audio data (pp.72–85) – Chapter 6: Recording (pp.91–94)

- Smith, S. Digital signal processing: a practical guide for engineers and scientists. (Burlington, MA: Elsevier, 2013).