- **Digital signal processing: Understanding digital sound files**
- **Recordings for phonetic analysis**

To read along with these slides:

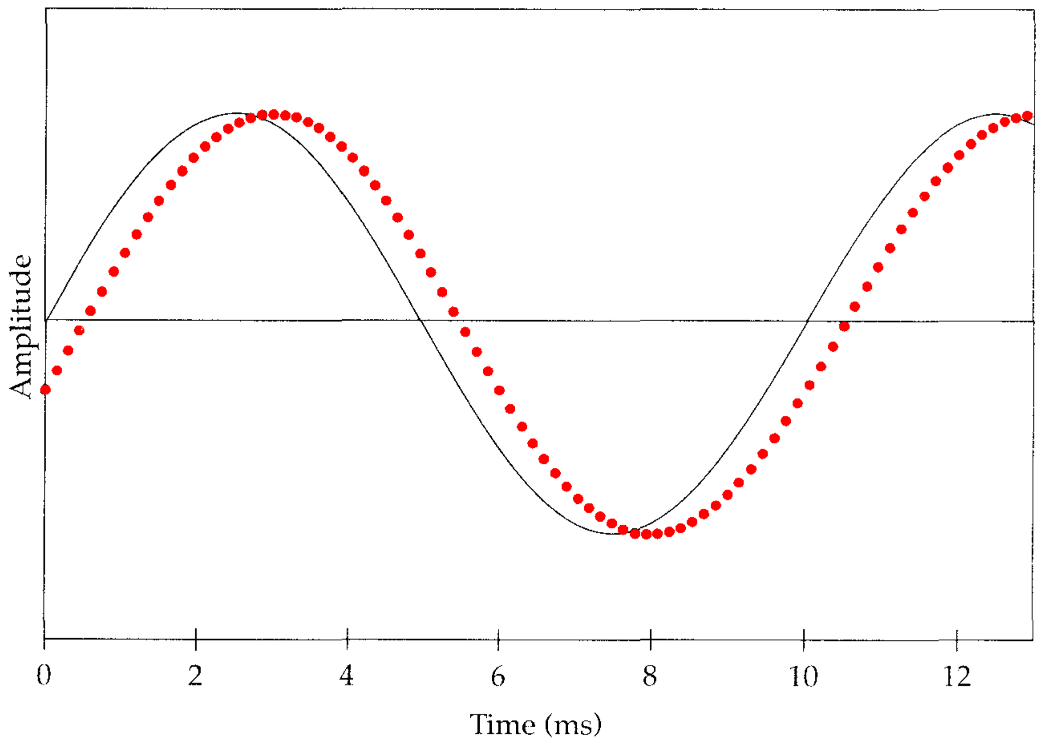
- *AAP* Ch 3, sec 3.1-3.2 (digital signal processing)

Overview

- DSP: Theory
- Making good recordings: Practice

1. Continuous and discrete signals

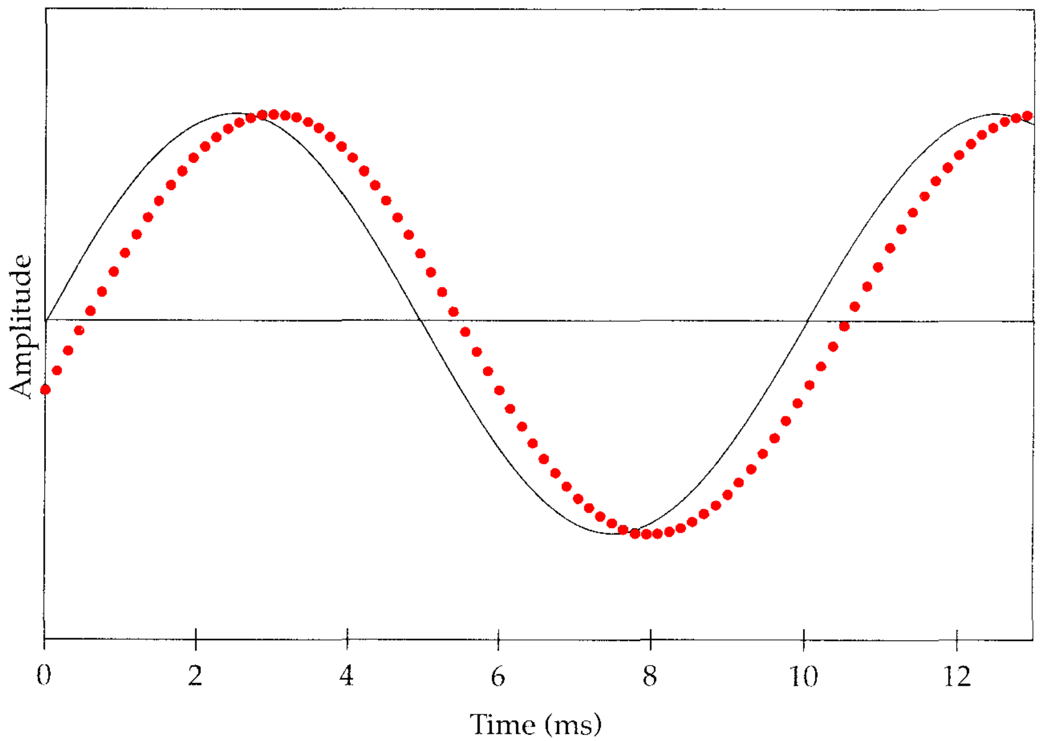
- This graphic (based on *AAP*, Fig. 3.1) represents a **continuous** sine wave and a **discrete** sine wave



- Which is which?

1. Continuous and discrete signals

- This graphic (based on *AAP*, Fig. 3.1) represents a **continuous** sine wave and a **discrete** sine wave



- Which is which?

continuous

discrete

1. Continuous and discrete signals

- Continuous or discrete?
  - Air-pressure fluctuations over time: __________
  - Sound files stored on a computer: __________
  - Analog signals: __________
  - Digital signals: __________

1. Continuous and discrete signals

- Continuous or discrete?
  - Air-pressure fluctuations over time: continuous
  - Sound files stored on a computer: discrete
  - Analog signals: continuous
  - Digital signals: discrete

1. Continuous and discrete signals

- Compare **analog** and **digital** clocks

1. Continuous and discrete signals

- Compare **analog** and **digital** clocks
  - **Analog** clock: The hand sweeps around continuously and smoothly, and it points to every spot on the circle of the clock face
  - **Digital** clock: It divides time into intervals, and jumps from one interval to the next, with no information between intervals
    - An alarm clock shows you 5:07 and then 5:08, but no information between those two time points
    - A digital stopwatch shows you 5:07:35.02 and 5:07:35.03, but no information between those two time points (though the interval is much smaller!)

2. Understanding digital sound files

- A sound wave (disturbance in a medium), measured in terms of air-pressure fluctuations, is **continuous**
  - Time: There is *some* numerical air-pressure value at **any** given instant
  - Amplitude: The air-pressure value can be any magnitude whatsoever

2. Understanding digital sound files

- A sound wave (disturbance in a medium), measured in terms of air-pressure fluctuations, is **continuous**
- Recording this sound wave and storing it on a computer converts it into a **digital** representation
  - Time: Pressure measured at **discrete intervals**
    - Imagine a tiny observer, measuring the air pressure hundreds or thousands of times every second
    - There are still timepoints where *no* measurement is taken = *no* air-pressure value is represented
  - Amplitude: Pressure measured in **discrete units**
    - The units can be very small—but not infinitely small

3. Sampling rate

- **Sampling rate**
  - What does this refer to, for a digital sound file?

3. Sampling rate

- **Sampling rate**
  - What does this refer to, for a digital sound file?

    AAP, p 51: the “number of times per second that we measure the continuous wave in producing the discrete representation of the signal”

  - How can we find it in Praat?

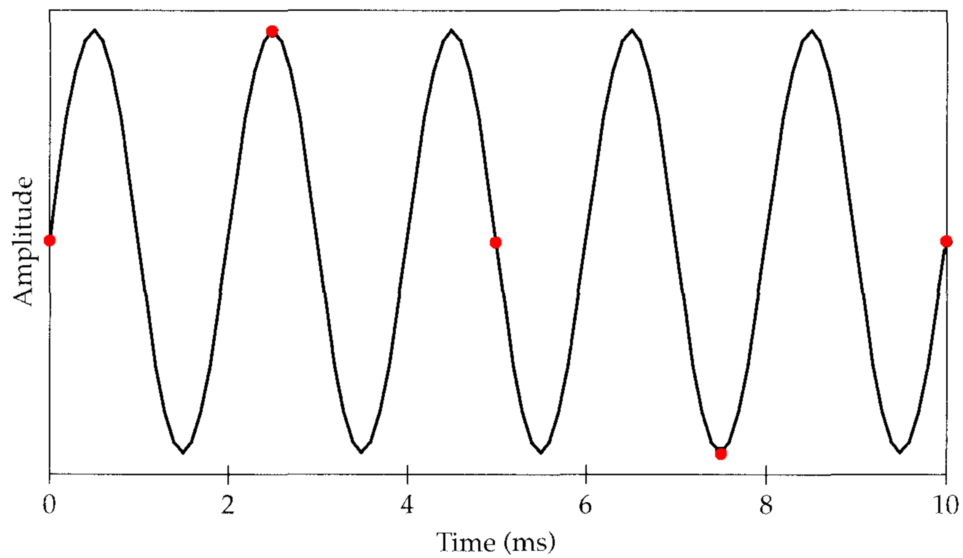
3. Sampling rate

- What is **aliasing**?
- What is the **Nyquist frequency**, and how does it help us avoid aliasing in a digital recording?

3. Sampling rate

- What is **aliasing**?

→ Suppose we sample this sine wave every 2.5ms, as shown by the red dots



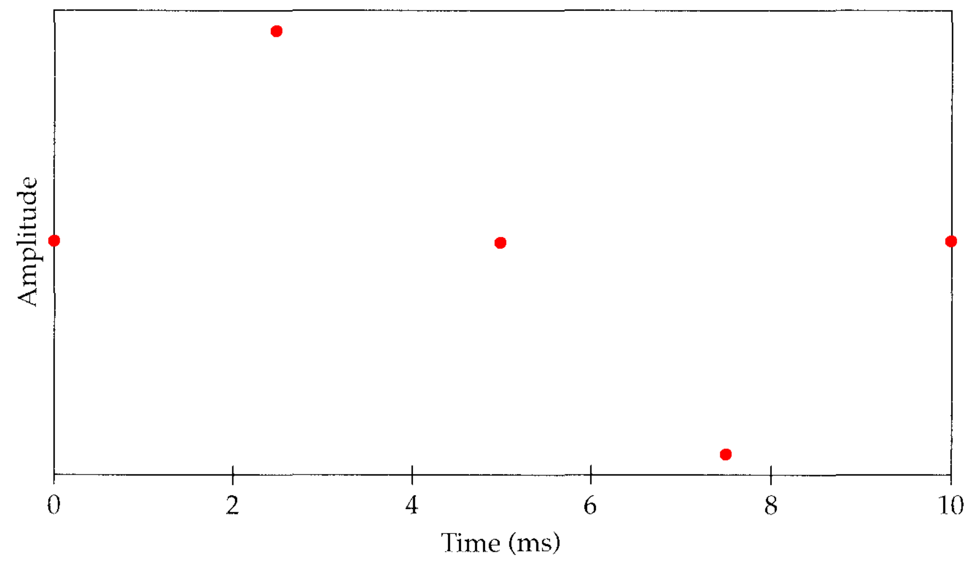
(adapted from *AAP* Fig 3.4, p 54)

- sampling rate = 1/2.5ms = 400Hz
- (black) sine wave *f* = 5 cycles/.01s = 500Hz

3. Sampling rate

- What is **aliasing**?

→ The computer's representation of this wave, being **digital**, looks like this



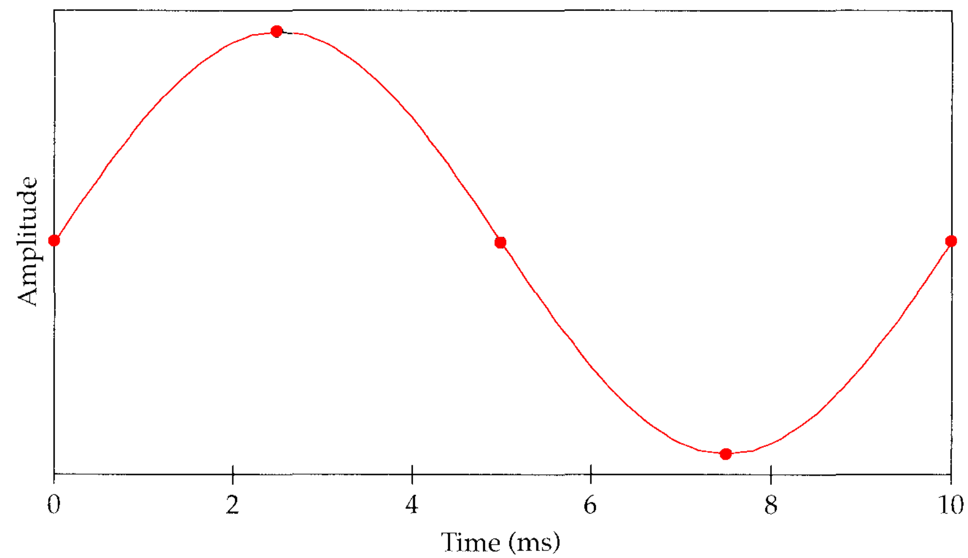
(adapted from *AAP* Fig 3.4, p 54)

- sampling rate = 1/2.5ms = 400Hz

3. Sampling rate

- What is **aliasing**?

→ This information will be interpreted as corresponding to the **red** sine wave now shown



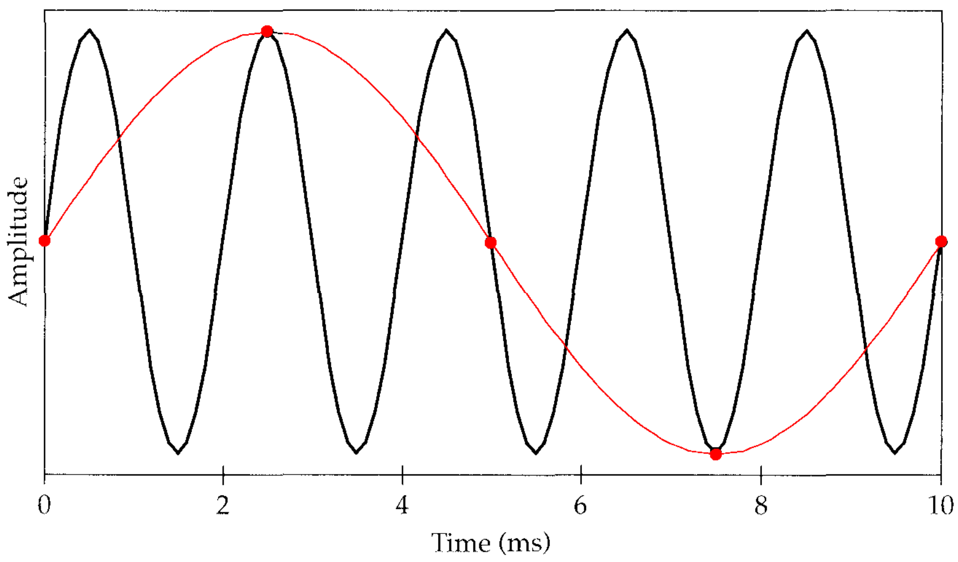
(adapted from *AAP* Fig 3.4, p 54)

- sampling rate = 1/2.5ms = 400Hz
- red sine wave *f* = 1 cycle/.01s = 100Hz

3. Sampling rate

- What is **aliasing**?

→ Analysis is therefore *missing* the (real) **black** component (500Hz) and *including* the (spurious) **red** component (100Hz)



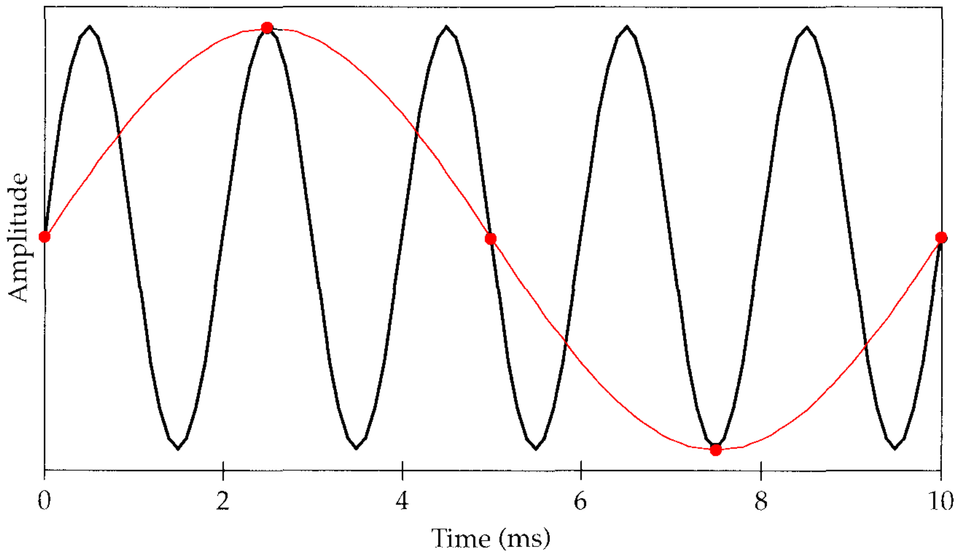
(adapted from *AAP* Fig 3.4, p 54)

- sampling rate = 1/2.5ms = 400Hz
- black sine wave f = 5 cycles/.01s = 500Hz
- red sine wave f = 1 cycle/.01s = 100Hz

3. Sampling rate

- What is **aliasing**? | *AAP*, p 54: "This misrepresentation of a continuous signal in a discrete waveform is called aliasing..."

→ Analysis is therefore *missing* the (real) **black** component (500Hz) and *including* the (spurious) **red** component (100Hz)



(adapted from *AAP* Fig 3.4, p 54)

3. Sampling rate

- What is the **Nyquist frequency**, and how does it help us avoid aliasing in a digital recording?
  - The Nyquist frequency is ________
  - Aliasing occurs when ________

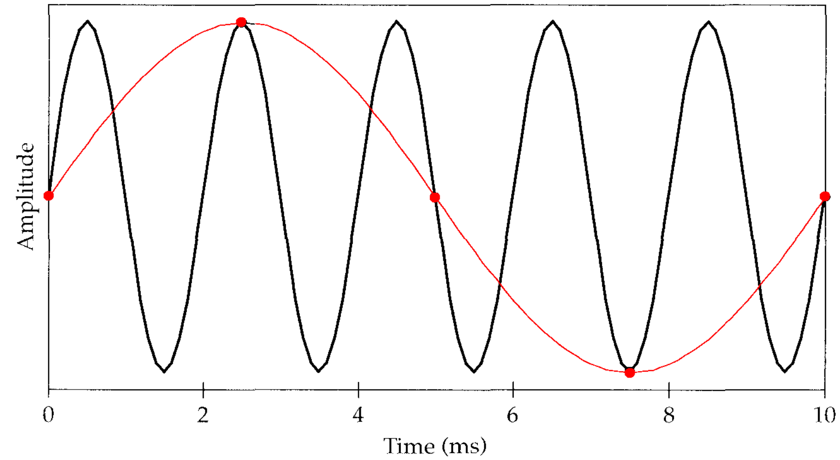
3. Sampling rate

- What is the **Nyquist frequency**, and how does it help us avoid aliasing in a digital recording?
  - The Nyquist frequency is half the sampling rate
  - Aliasing occurs when there are components at frequencies *higher* than the Nyquist frequency

- sampling rate = 400Hz
- **Nyquist frequency = 200Hz**

- black sine wave *f* = 500Hz

→ 500Hz > 200Hz, so: aliasing!



(adapted from *AAP* Fig 3.4, p 54)

3. Sampling rate

- We need to…
  - **Filter out** frequencies above the Nyquist frequency to avoid aliasing
  - Make sure our **sampling rate** is **high enough** to *include* components of interest

3. Sampling rate

- We need to…
  - **Filter out** frequencies above the Nyquist frequency to avoid aliasing

    → A microphone has a "frequency response band" (often 20Hz–20,000Hz) — it filters out components outside this range

  - Make sure our **sampling rate** is **high enough** to *include* components of interest

    → This is a setting we can control when recording in Praat: 44.1 kHz is recommended — why? (see above point about microphones!)

4. Quantization

- **Quantization**
  - What does this refer to, for a digital sound file?

4. Quantization

- **Quantization**
  - What does this refer to, for a digital sound file? See discussion in *AAP*, p 55, summarized here: Quantization refers to how finely grained the amplitude measurements will be when the sound wave is sampled
    - Essentially: How many decimal places will the amplitude value be rounded off to?

4. Quantization

- What are the **trade-offs** in determining a **setting** for quantization?

4. Quantization

- What are the **trade-offs** in determining a **setting** for quantization?
  - More accurate amplitude values require more computer memory for storage
    - Fortunately, this is much less of an issue than it used to be!

→ Computer sound cards these days use *at least* 16-bit quantization, which is very good
(*AAP*, p 57: signal-to-noise ratio for 16-bit quantization is 65,536:1!)
- We don't need to set a value for quantization in Praat

5. Audio file formats

- In acoustic analysis, we want to use **digital** sound files that are as **close** as possible to the original **analog** sound waves they represent

  Ideally, we would use…

  - A **high sampling rate** (44.1kHz is CD-quality)
  - **Higher-bit quantization** (16-bit is CD-quality)

- We also want to **avoid** using sound file formats that **compress** the data and **lose information** ("lossy compression")

5. Audio file formats

- Good **file formats** to use for **phonetic analysis** are:
  - .WAV files
  - .AIFF, .AIFC files (this is a Mac format, but they open in Praat on Windows machines too)
  - Many of the other formats Praat offers in the "Save as..." menu (in the Objects window) would be fine

- Note that this list **does not include MP3**

5. Audio file formats

- MP3s use **lossy compression** — information about components is lost when this file format is created
  - Therefore, it is preferable not to use MP3 files for phonetic analysis
  - If using MP3s is unavoidable, try to use a **compression setting of 192 kb/s**, which is the least compression possible

- For more information, see "What does MP3 do to your recordings?" by Sydney Wood

6. Making recordings in Praat

- See Praat handout #5 and *AAP*, p 58

  What are some things to watch out for?

  - Placement of microphone — why?
  - Sound input level — why? What happens if it is too low? Too high?