Lab 4: Sound Recording

Learning Objectives
By the end of this lab, you will be able to:

- Explain how a microphone and computer convert sound into a digital file
- Define sampling rate, bit depth, and bit rate, and describe how each affects audio quality
- Record and analyze audio waveforms using Audacity
- Compare the file size and quality of different audio formats
- Explain the difference between lossy and lossless compression

When you speak into a microphone, air pressure variations cause a diaphragm to vibrate, producing a continuously varying electrical voltage. A computer cannot store a continuous signal — it must convert it into a series of numbers through a process called **analog-to-digital conversion (ADC)**.

This process has two key parameters:

Sampling rate — How many times per second the computer measures the amplitude of the signal. Measured in Hertz (Hz). CD-quality audio uses 44,100 Hz, meaning 44,100 measurements every second.
Bit depth — How precisely each measurement is stored. A 16-bit recording can represent 2^{16} = 65,536 distinct amplitude levels. More levels means a more accurate representation of the original wave, and a greater **dynamic range** (the difference between the quietest and loudest sounds). Each additional bit adds approximately 6 dB of dynamic range.

The Nyquist Theorem states that to accurately capture a given frequency, you must sample at *least* twice as fast as that frequency. A 44,100 Hz sampling rate can therefore capture frequencies up to 22,050 Hz — just above the upper limit of human hearing (~20,000 Hz).

Key Vocabulary

Term	Definition
Sampling rate	Number of amplitude measurements per second (Hz)
Bit depth	Number of bits used per sample; determines amplitude precision
Bit rate	Data per second = sampling rate × bit depth × channels
Quantization	Rounding each sample to the nearest available level
Quantization noise	The small error introduced by rounding; sounds like hiss
PCM	Pulse-Code Modulation — the standard uncompressed audio encoding
Nyquist theorem	Max capturable frequency = sampling rate ÷ 2
Dynamic range	Ratio of loudest to quietest signal, measured in dB

Setup

1. Open **Audacity** from your Applications folder.

2. Set your input device. Using the Buttons on the toolbar, click **Audio Setup → Recording Device → H2n** to select your microphone**.** Under **Audio Setup → Recording Channels**, set Channels to **1 (Mono)**.

3. Set bit depth: **Audacity → Preferences → Quality**. Set Default Sample Format to **16-bit**. Click OK.

4. Check the Project Rate (Hz) box in the bottom-left of the main window. This controls your sampling rate. Note the current value: _____ Hz

The Audacity Interface — Identify These Before Recording

- **Transport toolbar** — Record (●), Play (▶), Stop (■) buttons at the top
- **Waveform display** — Visual representation of amplitude over time
- **Timeline** — Ruler along the top showing time in seconds
- **Selection tool** — The I-beam cursor; click and drag to select a region
- **Recording level meter** — Shows input level in dB; aim for peaks between −18 dB and −6 dB
- **Project Rate box** — Bottom-left; controls sampling rate for new recordings

Avoiding clipping: If the level meter hits 0 dB (the red zone), the signal is clipping — it has exceeded the maximum amplitude and the waveform peaks are being cut off, causing permanent distortion. Keep levels in the yellow zone.

Part A: The Effect of Sampling Rate

Background

A lower sampling rate means fewer samples are captured each second. By the Nyquist theorem, this directly limits the highest frequency that can be accurately recorded. Below a certain sampling rate, high-frequency sounds become distorted or disappear entirely.

Procedure

1. Set Project Rate to **44,100 Hz**.

2. Record yourself saying the same consistent phrase (e.g., *"The sampling rate affects sound quality"*). Click the track name and label it **"44,100 Hz."**

3. **Export:** File → Export → Export as WAV. Name it sample_44100.wav. Note the file size in Finder (right-click → Get Info).

4. Change Project Rate to **22,050 Hz**. Record the exact same phrase into a new track. Label it. Export as sample_22050.wav.

5. Change Project Rate to **8,000 Hz**. Record the same phrase. Label it. Export as sample_8000.wav.

6. Import all three WAV files back into Audacity (File → Import → Audio or just drag them into the window) and listen to each carefully. *Hint: By clicking the "Solo" button on each track, you can play back just that track without the others*.

Data Table A — Sampling Rate Comparison

Sampling rate	Max capturable frequency (SR ÷ 2)	WAV file size (KB)	Description of audio quality
44,100 Hz			
22,050 Hz			
8,000 Hz			

Analysis

1. At 8,000 Hz, describe specifically what happened to consonant sounds like "s," "f," and "th." Using the Nyquist theorem, explain *why* this happened.

2. 8,000 Hz is the standard sampling rate used for telephone calls. Why can you still understand speech over the phone even at this low sampling rate? What information is being lost that you don't notice in ordinary conversation?

Part B: Bit Depth and Dynamic Range

Background
Bit depth determines how many distinct amplitude levels are available to represent each sample. Fewer levels means more **quantization noise** — a subtle hiss or crunch that becomes especially audible in quiet passages and when audio is amplified.

Procedure
1. Set Project Rate back to **44,100 Hz** and bit depth to **16-bit** (Audacity → Preferences → Quality).

2. Record about 5 seconds of audio that starts as a soft whisper and gradually increases to a normal speaking voice. This dynamic range will help reveal the effect of bit depth.

3. Zoom in closely on a very quiet section of the waveform. Describe its appearance.

4. Export this sample as a .wav file and give it the name "16bit.wav"

5. Export this sample as a .wav file, but change the encoding to "Unsigned 8-bit PCM" and give it the name "8bit.wav". Re-import the 8 bit file into audiacity.

6. Compare the two waveforms, especially in the quiet sections. Listen to the two recordings — how would you describe the differences between the 8-bit and 16-bit recordings?

Data Table B — Bit Depth Comparison

Bit depth	Amplitude levels (2^n)	Theoretical dynamic range (~6 dB × bits)	Audible noise in quiet sections?	Waveform appearance in quiet section
8-bit				
16-bit				

Analysis

1. Calculate the dynamic range of a 16-bit recording using the formula: Dynamic range ≈ 6.02 × bit depth. Show your work. How does this compare to the dynamic range of human hearing (~120 dB)?

Part C: File Formats and Compression

Background: How File Size Is Calculated

For uncompressed audio:

File size (bytes) = Sampling rate × Bit depth × Channels × Duration ÷ 8

Example — 1 minute of stereo audio at 44,100 Hz, 16-bit: 44,100 × 16 × 2 × 60 ÷ 8 = **10,584,000 bytes ≈ 10.1 MB**

Audio File Formats

Format	Extension	Compression	Data lost?	Common use
Waveform Audio	.wav	None (raw PCM)	No	Professional recording, Windows
Audio Interchange	.aiff	None (raw PCM)	No	Professional recording, Mac/Apple
Free Lossless Audio Codec	.flac	Lossless (~50% smaller)	No	Audiophile listening, archiving
MPEG-1 Audio Layer 3	.mp3	Lossy (~90% smaller)	Yes, permanently	Music streaming, general use
Advanced Audio Coding	.aac	Lossy (~90% smaller)	Yes, permanently	Apple Music, YouTube
Ogg Vorbis	.ogg	Lossy (~90% smaller)	Yes, permanently	Open-source apps, video games

Lossy compression (MP3, AAC) works by exploiting **psychoacoustics** — a model of what humans actually hear. It discards:

- Quiet sounds that are masked by louder simultaneous sounds (frequency masking)

- Sounds immediately following a loud sound, when hearing is temporarily less sensitive (temporal masking)
- Very high frequencies (above ~15,000 Hz), which are difficult for most listeners to perceive

Lossless compression (FLAC) finds mathematical redundancies in the data and encodes them more efficiently. No data is discarded; the original file can be perfectly reconstructed.

Procedure

1. Set Project Rate to **44,100 Hz**, bit depth to **16-bit**, mono.

2. Download from the course webpage and import into Audacity the clip "orchestral.aiff"

3. **Before checking**, calculate the predicted file size (in kB). Show your work in your lab notebook.

4. Export two compressed versions of the same clip:
    - File → Export → **Export as MP3** → set Bit Rate Mode to "Constant" → set quality to **64 kbps** → ensure that "Stereo" is set → name it orchestral_64.mp3
    - File → Export → **Export as MP3** → set Bit Rate Mode to "Constant" → set quality to **128 kbps** → ensure that "Stereo" is set → name it orchestral_128.mp3

5. Check the file sizes in Finder (right-click → Get Info) and record them below.

6. Import all three tracks into Audacity. Play the tracks and click the "Solo" button on the left hand side of the track to listen to each of them in turn. Can you hear any difference?

7. **Spectrum comparison:** In the upper-left corner of each track, click the down-arrow next to the track's name and set it "Spectrogram." Compare the frequency spectrum of each clip, with particular attention to the higher frequencies. Can you see anything different?

Data Table File Format Comparison

Format	**Bit rate (kbps)**	**Predicted size (KB)**	**Actual size (KB)**	**Audible difference from WAV?**
.aiff (44,100 Hz, 16-bit)	1,411			—
.mp3 @ 64 kbps	64	—		
.mp3 @ 128 kbps	128	—		

Analysis

1. Describe what you observed in the spectrum analysis above 15,000 Hz for the AIFF vs. the 64 kbps MP3. Why does MP3 compression remove this frequency content first?

2. The 64 kbps MP3 is what fraction of the size of the AIFF file? Show your calculation.

Full Lab Analysis

1. A music producer records at 96,000 Hz, 24-bit, stereo. Using the file size formula, calculate the size of a 4-minute recording at these settings. Then calculate the same recording at CD quality (44,100 Hz, 16-bit). How much larger is the high-resolution file? Do you think the extra quality justifies the extra storage? Defend your answer with evidence from this lab.

2. A friend downloads a 320 kbps MP3, re-encodes it at 64 kbps to save space, then re-encodes it back to 320 kbps "to restore the quality." Using what you learned in this lab, explain clearly why this does not work.

3. A podcast producer has 100 hours of recorded audio and needs to decide between storing it as WAV, FLAC, or MP3 at 128 kbps. Estimate the storage required for each format (assume mono, 44,100 Hz, 16-bit as your baseline). Write a recommendation explaining the trade-offs.

Extra Stuff

1. **Room Acoustics** Record the same spoken phrase in two different spaces: a small enclosed room and a large open space (hallway or gym). Import both into Audacity and compare the waveforms and spectrum. Describe the differences. What does reverberation look like in a waveform?

2. **Noise Floor** Record 10 seconds of complete silence (microphone on, nobody speaking). Zoom in — you'll see small random fluctuations. This is the **noise floor**. Apply +40 dB of amplification (Effect → Amplify). What do you hear? Using your knowledge of bit depth, explain why a lower bit depth has a worse noise floor.

3. **Clipping and Distortion** Record a very loud sound that causes the meter to hit 0 dB, then the same sound at a safe level. Zoom in and compare the waveform peaks. Run Analyze → Plot Spectrum on both. Clipping creates **harmonic distortion** — describe any extra frequency peaks you observe in the clipped recording.