Week 6: Audio Processing and Compression

DIGITAL ASSET DEVELOPMENT

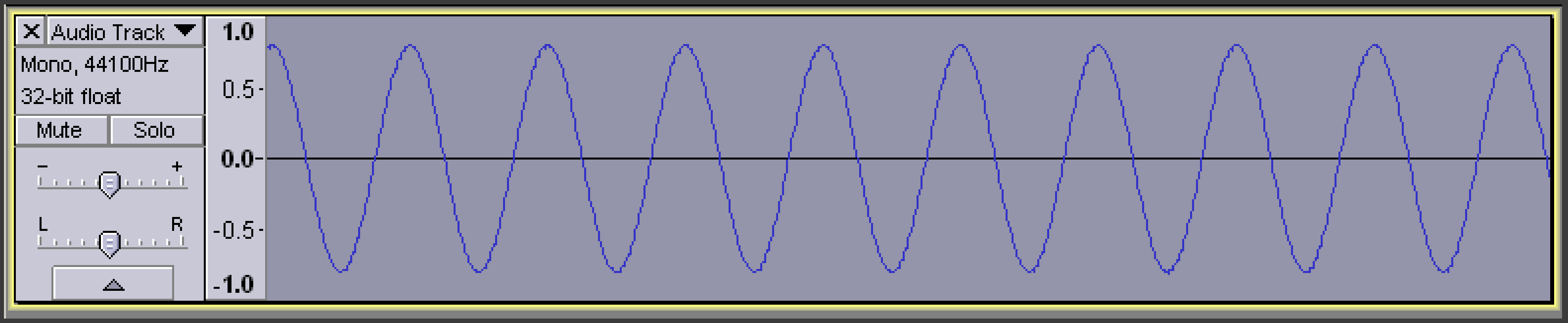
Contents

- Frequency-based audio analysis
- Audio filtering
- Audio compression

Frequencies and Tones

- As we saw last week, the tone or pitch of a sound relates to its frequency
  - Higher pitch implies higher frequency
- Most audio consists of a combination of sounds with different tones
  - Known as component frequencies
- The simplest sound is one that consists of a single tone
  - The corresponding sound wave has the appearance of a sine wave

Sine Wave (Sinusoidal Curve)



- Any sound is made up of a combination of pure tones
  - In other words, it is made by combining sine waves of different frequencies
  - These waves are simply added together

Decomposing Audio Data

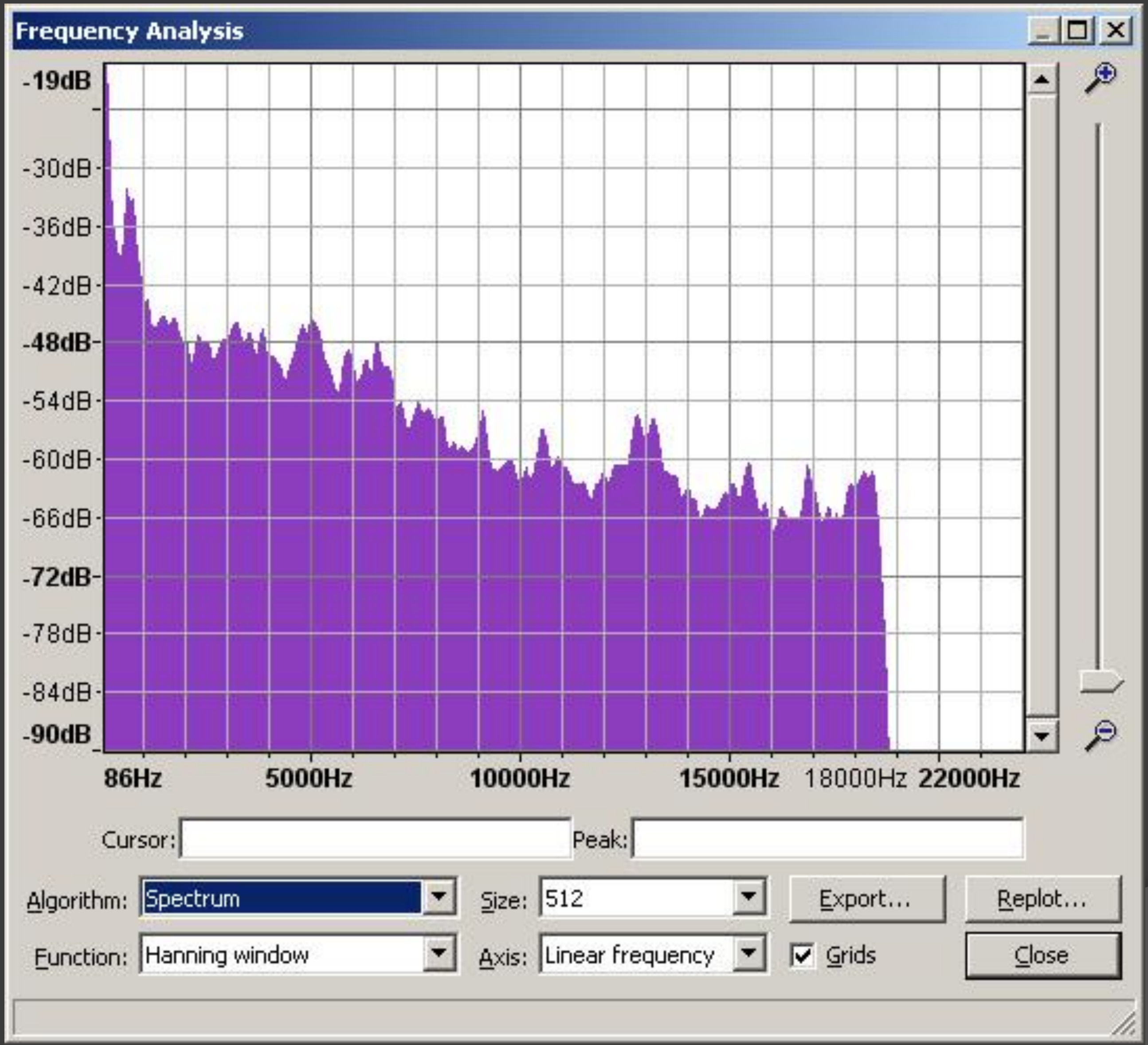
- To process complex sounds, we need to access their component frequencies
  - This entails splitting a sound waveform into a collection of sine waves (decomposition)
  - We can then process specific frequencies individually within the overall sound
- The Fourier Transform is a mathematical operation that does exactly this
  - Similar (though not identical) to the process used to compress a JPEG image

Fourier Transform

- The transform remaps the audio data into frequency space
- Measures the contribution to the overall sound at specific frequencies
  - Parts of the sound wave that only change slowly contain low frequency energy
  - Parts that change rapidly represent high frequencies
- Note that this process is reversible

Frequency Spectrum

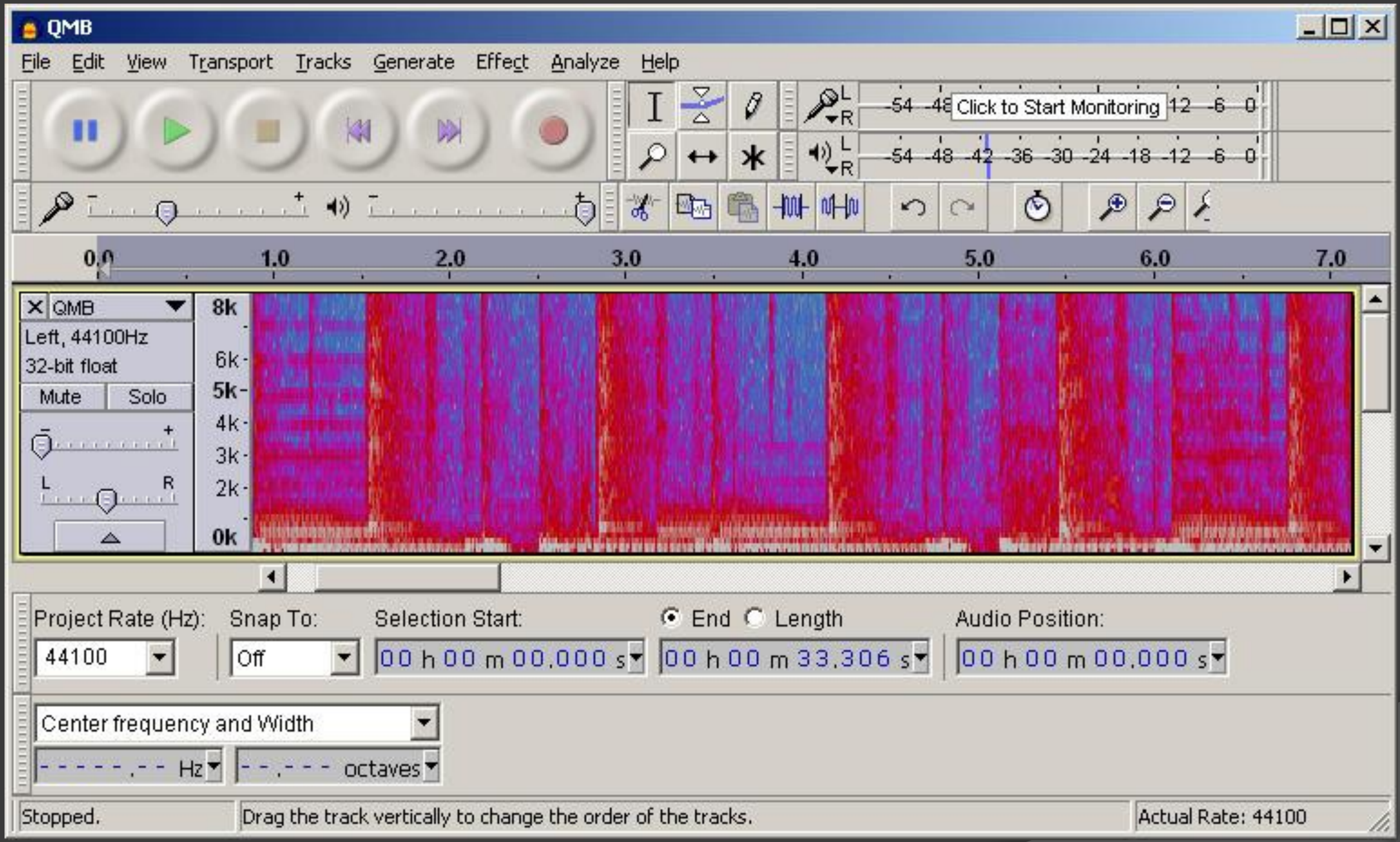
- We can view the Fourier Transform of a sound as its frequency spectrum
- Plot Spectrum function in Audacity
  - Plot frequencies along the x-axis (low to high)
  - Height for a given frequency corresponds to its contribution to the audio signal
- Peaks in the graph represent dominant frequencies in the sound


Frequency Analysis
-19dB
-30dB
-36dB
-42dB
-48dB
-54dB
-60dB
-66dB
-72dB
-78dB
-84dB
-90dB
86Hz
5000Hz
10000Hz
15000Hz
18000Hz
22000Hz
Cursor:
Peak:
Algorithm:
Spectrum
Size:
512
Export...
Replot...
Function:
Hanning window
Axis:
Linear frequency
Grids
Close

Spectrogram

- Sometimes it is more useful to see how the dominant frequencies vary over time
- We can do this using an audio clip's spectrogram
- This creates a 2D plot:
  - Time runs from left to right
  - Frequency is plotted vertically
  - Contribution (energy) at a given frequency and time is colour coded

Spectrogram Example



Using the Spectrogram

- The spectrogram is a very useful tool for processing audio files
- With experience, specific problems with audio files can be identified from the plot
  - Background noise at a specific frequency will appear as a horizontal bar
  - This is a common audio issue, often caused by electrical interference
  - Clicks and other glitches on an audio track may appear as a vertical spike

Example (from lab exercise)

horizontal banding caused by (unwanted?) harmonics of a fundamental frequency

Audio Cleanup

- Using the spectrogram it is possible to isolate and rectify some audio problems
  - Select the affected set of frequencies
  - Apply some form of filter to reduce or eliminate the noise
  - Low pass, high pass and notch filters
- Other approaches are possible, including “redrawing” the waveform
  - Useful for fine-tuning audio files

Frequency-based Filters

- A low pass filter suppresses high frequencies from a sound
  - Lets low frequencies “pass”
- A high pass filter does the opposite
- These are the audio equivalents of blur and edge detection filters for images
- A notch or band pass filter suppresses frequencies within a set range
  - Useful for targeting specific types of noise

Compressing Audio Data

- As mentioned last week, it is convenient to be able to compress audio files
- Ideally, we should do this without hurting overall audio quality
- Mostly use frequency-based approach
  - Identify component frequencies
  - Remove or suppress those frequencies which humans will not hear (much!)
  - Store quantised frequency data

Telephony

- An early use of digital audio encoding was in telephone systems
- Base standards for digital telecoms are μ-law (N America / Japan) and A-law
  - Variations on pulse code modulation (PCM)
  - Use statistical properties of human voice patterns to improve data quality
  - Reduce effective 12 bit sample resolution to 8 bit requirement for transmission
- Still the basis for telephone audio

Psychoacoustic Models

- For more sophisticated compression, we need to study human hearing factors
  - Known as psychoacoustic modelling
- Main characteristics:
  - Frequency limits (typically 20 Hz – 20 kHz)
  - Our sensitivity peaks around 1-5 kHz
  - Simultaneous masking: sounds “merge” if played together (frequency dependent)
  - Temporal masking: louder sounds mask others immediately before or after them

Audio Formats

- Uncompressed audio formats include WAV and AIFF (Windows and Mac)
  - Used for raw recordings
  - Allow a choice of bit depths
- A wide range of compression formats exist, though many are related
  - MP3 is the most well known
- MP3 supports a number of codecs
  - Compressor/decompressor algorithm

The MP3 Format

- MPEG-1 (or 2) Audio Layer 3
  - Part of the MPEG video standard
  - De facto standard for digital music
- Follows frequency-based compression process described in earlier slide
- Uses psychoacoustic modelling to optimise perceived quality
- Choice of compression ratios allowed

MP3 Compression

- Most common bit rate is 128 kbits/sec
  - Approx 10:1 compression ratio
- Higher bit rates should offer improved sound quality
- This is for constant bit rate (CBR) files
- MP3 also allows for variable bit rate (VBR) compression
  - Can be useful where the complexity of the audio signal varies greatly during a piece

Choosing an Audio Format

- For audio processing, it is best to use uncompressed formats
- For distribution, key issues are sound quality and interoperability
- MP3 is a reasonable default, though AAC and Vorbis reputedly sound better
  - Quality / file size also depends on codec
  - Some MP3 compression algorithms give better results than others