UOSH: Archives+ Audio Digitisation Guide

101: Intro to Audio

1. [Intro To Audio Digitisation (Archives+, Manchester)]

For an introduction to Audio Digitisation please watch the 12min video above.

The video also includes a brief explanation of audio theory and file standards (below is a short summary of that information).

1.1 EXPLAINING SOUND WAVES

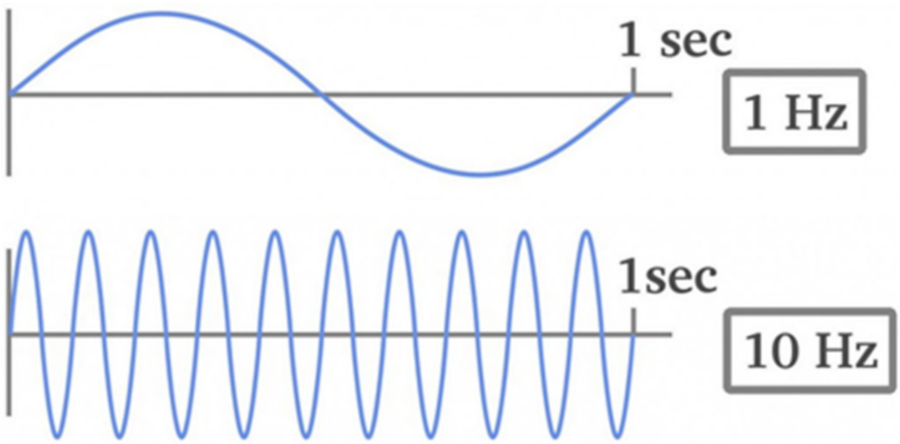
Sound waves are both longitudinal (travel in the same direction as the sound) and transversal (perpendicular). Think of ripples when you throw a stone into a pond:

[Image from http://www.ocdqblog.com/home/the-stone-wars-of-root-cause-analysis.html]

The air particles do not travel, instead they vibrate around a point in space and pass this movement on. The rate of this oscillation is called frequency.

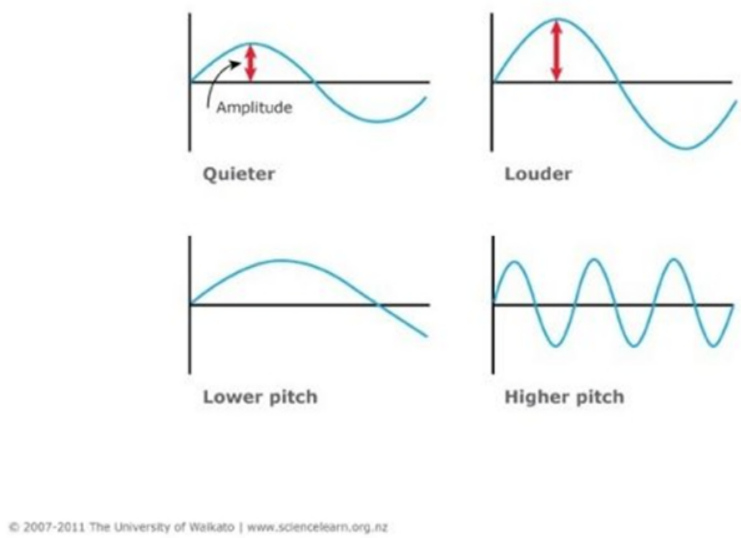
The amount of compression/rarefaction in the air is the amplitude of the sound wave. The distance between these peaks is the wavelength of the sound wave.

One sound wave is one peak and one trough together making it 1 oscillation. Note: 1 oscillation per second is 1 Hertz (Hz).



[image from https://www.scienceabc.com/eyeopeners/why-do-voices-sound-squeaky-when-theyre-sped-up.html]

Our perception of frequency is called pitch. The higher pitch something is the more the oscillations over time such as the 10 Hertz sound wave in the image above.



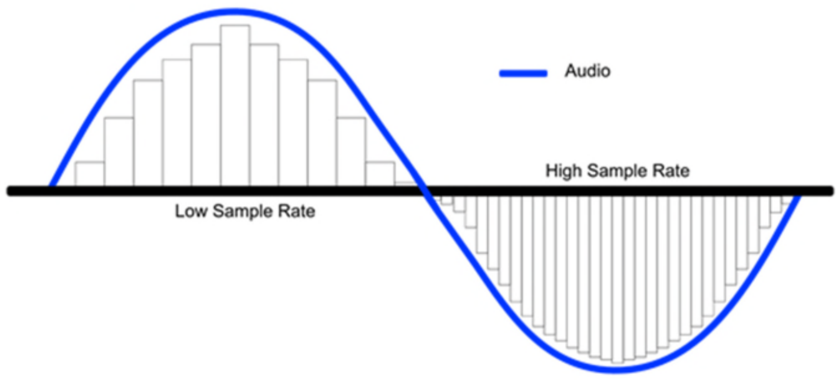
[Image from: https://onca.org.uk/graphs-of-sound-waves/]

1.02 AUDIO QUALITY

The audio quality of digitisation is the accuracy of the original recording depends on: **Sample rate** and **bit depth.**

Sample rate (measurement of amplitude over time in Hertz or Kilohertz)

The higher your sample rate the more points per second that you measure in a piece of audio. Increasing the number of samples means that that we are more accurately capturing the original sound wave:



Sample Rate infographic by Mastering The Mix.

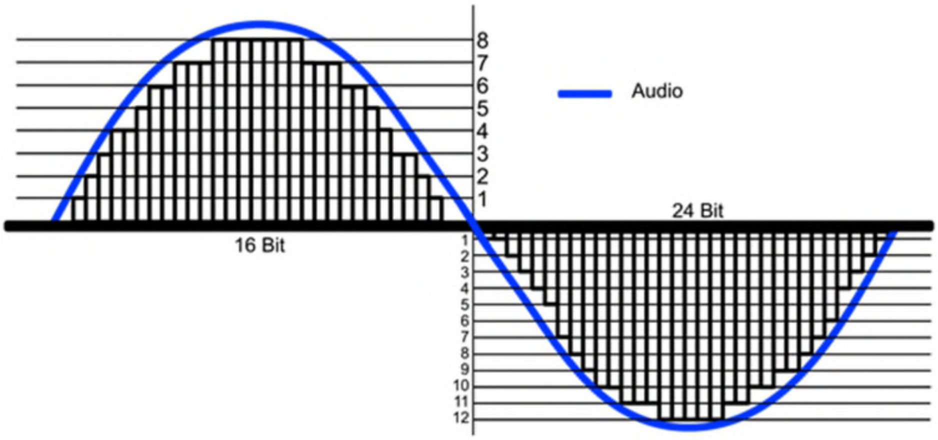
An audio CD is 44.1 kHz (that’s 48,000 Hz or 48,000 samples per second) so due to the Nyquist-Shannon sampling theorem the sample rate must be at least double the highest frequency it is sampling, so a CD captures frequencies up to 22kHz.

Humans’ hearing goes up to frequencies of 20kHz so, at 44.1kHz, Audio CDs include frequencies already outside of our hearing.

For archive digitisation we use a sample rate of 96 kHz. This means that we are capturing higher frequencies contained within recording and future-proofing standards of preservation files.

Bit depth (measured in bits)

The higher the bit depth, the more dynamic range we can capture (quantisation). This means more detail for the original recording is retained by the number of ‘bits’ of information in every sample is increased:



Bit Depth infographic by Mastering The Mix.

An Audio CD is 16 bit (65,000 values)

For archive digitisation we use a bit depth of 24 bits (over 16 million different values). This means that we are capturing more information contained within recording and future-proofing standards of preservation files so that in future digital preservation files can be modified and improved without having to re-run analogue tapes.

Note: Bit-rate (bits per second) should be confused with bit depth!

Bit rate is the number of bits per second that can be transmitted along a digital network. Example:

16 bits per sample x 96 kHz samples per second = 1536 kbps (mono) x 2 (stereo) = 3072 kilo bits per second (kbps)

To work out the size of the full audio file use this formula above and multiply by the length of the audio track:

Thousands of bits per second (kbps) x length of audio in seconds = file size in bites.

Current archive standards

We digitise to 24 bit and 96 kHz - this leaves plenty of headroom to allow for any future restoration work on the files or even any changes of archival standards in the future.

Frequency and Pitch

The limits of human gearing are approx. 35 to 16/17,000 Hz

This changes with age where the upper level changes to

15,000 Hz (aged 40)

12,000 Hz (aged 50)

10,000 Hz or lower beyond 50

We perceive frequency as pitch

LOW BASS (1st & 2nd octaves 20-80 Hz)

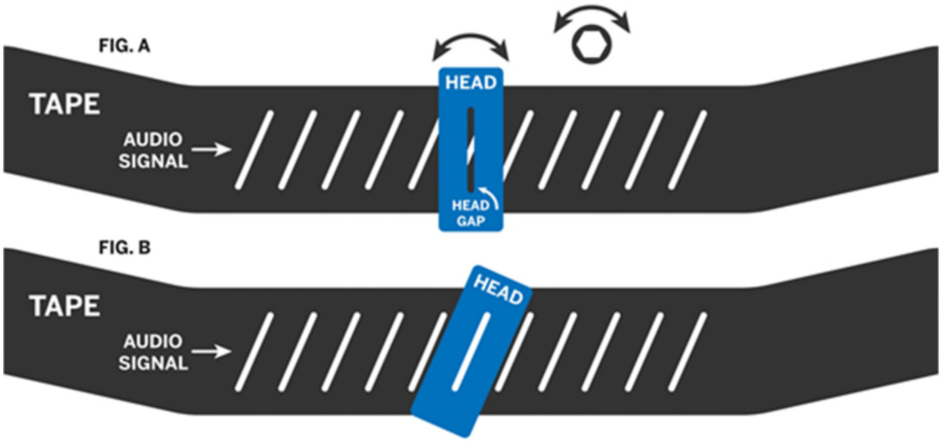
UPPER BASS (3rd & 4th octaves 80-320 Hz)

MIDRANGE (5th, 6th, 7th octaves 320-2,560 Hz)

UPPER MIDRANGE (8th octave 2,560 - 5,120 Hz)

Glossary

Azimuth – The angle that the audio signal has been recorded onto the tape. When playing back we need to manually match this same angle to get all of the signal to play back fully.



Bandwidth - Bandwidth is the difference between the highest and lowest frequencies carried in an audio stream.

Bias Frequency - Tape bias is the term for techniques used to improve the fidelity of tape recorders. AC bias increases the signal quality of most audio recordings significantly by pushing the signal into more linear zones of the tape's magnetic transfer function.

Decibels (dB) - can be used to express a change in value (e.g., +1 dB or −1 dB) or an absolute value.

(The definition of the decibel is based on the measurement of power in telephony of the early 20th century in the Bell System in the United States. One decibel is one tenth (deci-) of one bel, named in honor of Alexander Graham Bell; however, the bel is seldom used.)

dBFS (decibels full scale)

dB(full scale) – the amplitude of a signal compared with the maximum which a device can handle before clipping occurs. Full-scale may be defined as the power level of a full-scale sinusoid or alternatively a full-scale square wave. A signal measured with reference to a full-scale sine-wave will appear 3 dB weaker when referenced to a full-scale square wave, thus: 0 dBFS (fullscale sine wave) = −3 dBFS (fullscale square wave).

Dolby - Dolby Noise Reduction is a proprietary system to reduce unwanted low level tape hiss. If a tape was recorded using a type of Dolby (e.g. A B or C), it must be played back using that same setting.

Frequency - number of occurrences of a repeating event per unit of time (the unit for frequency is Hertz (cycles per second).

Aliasing - temporal aliasing is the phenomenon where the sample rate is too low to capture changes in the sound wave (aliasing occuring in signals over time).

PCM (pulse code modulation) used for storing lossless audio, in our case as a wav file.

Phase Inversion - (Not to be confused with Polarity Reversal) Waveform is delayed making the waves ‘180 degrees out of phase’ where the peak of one wave form meets the low form the other inside of rising and falling at the same time. This results in silence due to delaying.

Polarity Reversal - (Not to be confused with Phase inversion) Waveform is flipped upside down. This results in silence due to inverting.

Noise - unwanted sound.

Signal - this is the audio.

Signal to Noise Ratio (SNR) In any analogue system some of the voltage is what you want to measure (signal) and some is random fluctuations (noise). SNR is usually measured in decibels (dB).

Wow and flutter - Wow is an audible change of pitch (frequency) that occurs once per rotation. E.g., in a tape recorder any changes in frequency can be caused by irregular tape motion during recording or playback. For example, a change in the angular velocity of the capstan, or dragging of the tape within a reel or audio cassette shell.

Flutter is the deviation of frequency caused by irregular mechanical motion.

[The terms "wow and flutter" are often referred to together, flutter being a higher-rate version of wow.]

WAV - Waveform Audio File Format (WAVE, or more commonly known as WAV due to its filename extension; pronounced "wave" or “wav”.

WAV (BMF) - A Broadcast Wave File is a WAVE file with added information (timestamps etc) contained in the header. Can be exported as BWF files with the .WAV extension or .BWF.