BASICS OF [DIGITAL] AUDIO

Frequency of Sound

Frequency is the **number of times a waveform repeats** (i.e. completes one full cycle/period) **within a specific amount of time** (usually 1 second). Higher frequencies are perceived as higher pitches, and vice versa. Frequency of sound is usually measured in **Hertz (Hz), 1 Hertz = 1 cycle per second**.

NOTE:

Cycle usually indicates one full waveform: starting at 0, going up to the highest point, going back down to 0, going down to the lowest point, and then finally returning to 0.

Period is the length of time 1 cycle takes, but is often marked/measured as the time between 2 peaks/troughs in the waveform.

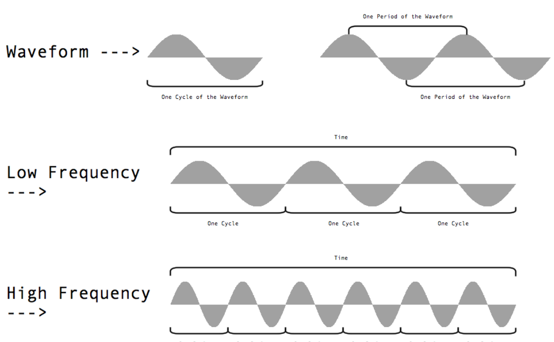

FIGURE 1. Relationship of cycle/period to the frequency of a sound.

http://ears.pierrecouprie.fr/spip.php?article73

http://www.physicsclassroom.com/class/sound/Lesson-2/Pitch-and-Frequency

https://en.wikipedia.org/wiki/Audio_frequency

http://www.indiana.edu/~emusic/acoustics/frequency.htm

http://www.nchearingloss.org/freq.htm?fromncshhh

Sampling Rate

Sampling, in signal processing, is the conversion of a continuos signal (i.e. audio) into a series of discreet steps. The "sampling rate" is the number of samples taken per second, and directly effects the highest frequency you can sample/create (see Nyquist Frequency).

https://en.wikipedia.org/wiki/Sampling_(signal_processing)#Sampling_rate

http://wiki.audacityteam.org/wiki/Sample_Rates

http://www.digitizationguidelines.gov/term.php?term=samplingrateaudio

Nyquist Frequency

The "Nyquist Frequency" is the highest frequency possible given by a specific sampling rate. It is always **LESS than half** the sampling rate! So, for example, if your sampling rate is 48000 samples per second (48kHz), then the Nyquist Frequency is $<\frac{48000}{2} = \mathbf{<24kHz}$.

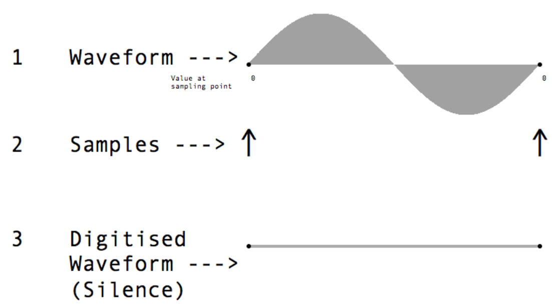

FIGURE 2. Digital Conversion of an analog signal using a sample rate **EQUAL** to the frequency of the waveform.

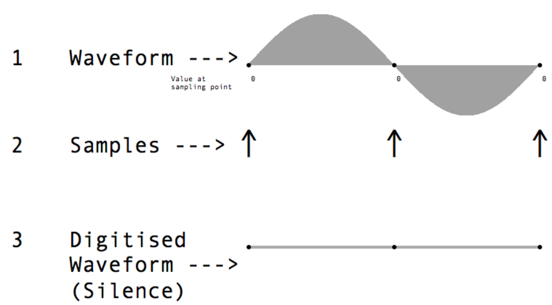

FIGURE 3. Digital Conversion of an analog signal using a sample rate **TWICE** the frequency of the waveform.

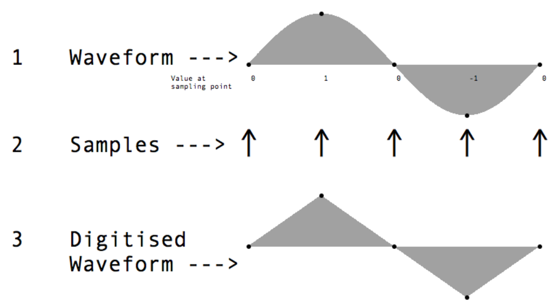

FIGURE 4. Digital Conversion of an analog signal using a sample rate **MORE THAN TWICE** the frequency of the waveform. Note how the original waveform is still distorted, an even higher sampling rate would smooth out the digitised waveform.

Frequencies above the **Nyquist Frequency** result in **aliasing**. This means that that frequencies become folded back to a frequency below the bf Nyquist Frequency.

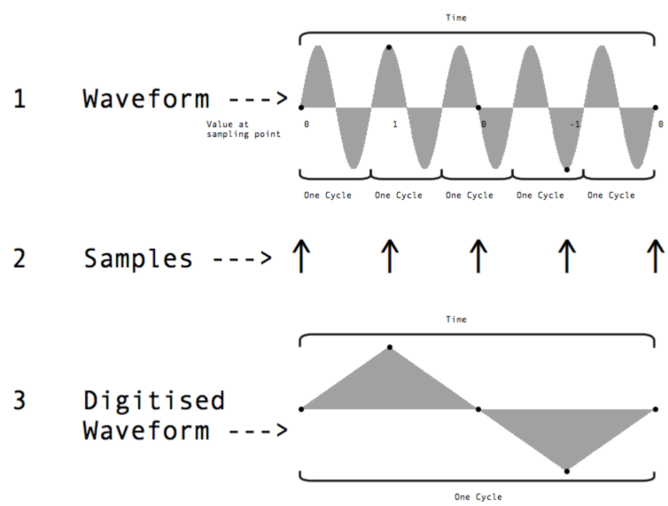

FIGURE 5. Aliasing from using too **SMALL** a sampling rate. Note how the **frequency** of the digitised sound is $\frac{1}{5}$th of the frequency of the original sound!!!

Higher **Sampling Rates** mean higher maximum frequencies, and also more accurate conversion of sounds (see Figs. 2 - 4). **BUT** at a certain point we go beyond the range of frequencies humans can actually hear (roughly 20Hz - 20kHz) so the benefits become, well, questionable...

http://www.indiana.edu/~emusic/etext/digital_audio/chapter5_nyquist.shtml http://mathworld.wolfram.com/NyquistFrequency.html

https://en.wikipedia.org/wiki/Nyquist_frequency

Bit Depth

Bit depth is the number of bits used for each sample. The **higher** the bit depth the more accurate each sample is, and therefore the more closely the sampled (digitised) waveform will represent the original waveform. Each sample taken that does not equal a bit value is rounded to the nearest bit, this is called **quantization** (see Figs. 8 and 9).

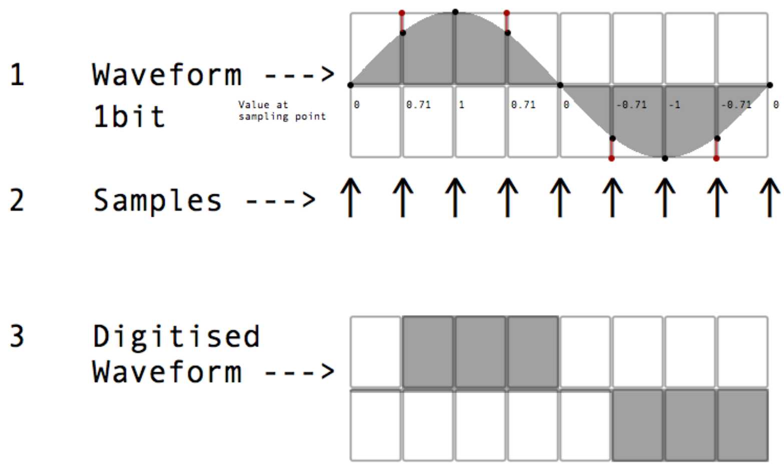

FIGURE 6. Sampling using a bit depth of 1 (red lines and dots represent quantization).

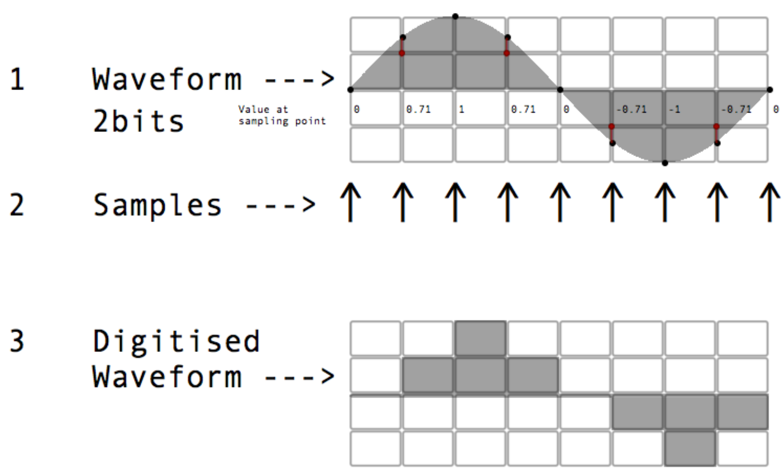

FIGURE 7. Sampling using a bit depth of 2. Note how the digitised waveform already beings to resemble the original more closely.

https://documentation.apple.com/en/finalcutpro/usermanual/index.html#chapter=52%26section=7%26tasks=true http://mp3.about.com/od/glossary/g/Bit-Depth-Definition-What-Does-Bit-Depth-Mean.htm

https://en.wikipedia.org/wiki/Audio_bit_depth

CONCLUSION

- **Sampling Rate** is **how often** samples are taken. It relates to frequency - specifically the highest frequencies which can be sampled.
- **Bit Rate** is **how accurate** the value taken for each sample is. It directly affects the **Dynamic Range**, i.e. the range of different values that can be sampled from quietist sounds to the loudest.

Envelope

"The shape of a sound's **amplitude** in time[1]."

"The profile of the evolution of the intensity and/or spectrum of a sound during its duration. Other uses of the term concern frequency and filter evolution in time, primarily on synthesizers[2]."

With acoustic sounds we often talk about amplitude envelopes in terms of:

- **Attack**: the duration from a zero to a maximum amplitude.
- Sustain: the level of the steady state amplitude.
- **Decay**: the duration from the steady state to its final zero amplitude[3].

Analog and Digital Synthesisers often have **ADSR** controls for amplitude envelopes:

- **Attack** (A): the duration from a zero to a maximum amplitude.
- **Decay** (D): the duration from the initial maximum amplitude to a stable state amplitude.
- **Sustain** (S): the level of the steady state amplitude.
- **Release** (R): the duration from the steady state to its final zero amplitude[4].

[1] http://www.sfu.ca/sonic-studio/handbook/Envelope.html
[2] http://ears.pierrecouprie.fr/spip.php?article70
[3] http://www.sfu.ca/sonic-studio/handbook/Envelope.html
[4] http://ears.pierrecouprie.fr/spip.php?article483

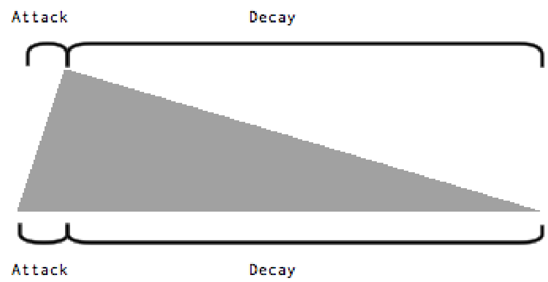

Figure 8. An envelope with a **short/sharp** attack and a **slow/soft/gentle** decay.

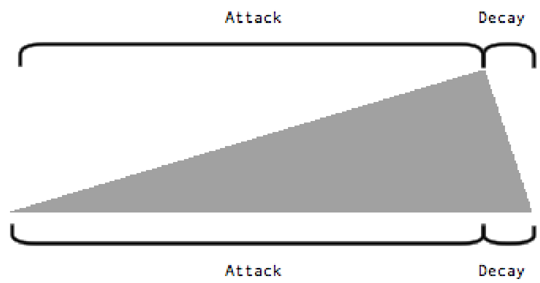

Figure 9. An envelope with a **slow/soft/gentle** attack and a **short/sharp** decay.

Spectrum/Timbre

"The frequency content of a sound or audio signal is it's **spectrum**[5]."

"The behaviour in time of the spectrum of a sound determines it's **timbre**[6]."

Signal-to-noise Ratio

"The ratio of the magnitude of the wanted signal to that of the unwanted noise...

A signal-to-noise ratio is said to be favorable when the signal predominates; that is, it can be clearly distinguished from the noise... When signal and noise are less clearly distinguishable, the signal-to-noise ratio is said to be poor or low[7]."

[5] http://www.sfu.ca/sonic-studio/handbook/Spectrum.html
[6] http://www.sfu.ca/sonic-studio/handbook/Timbre.html
[7] http://www.sfu.ca/sonic-studio/handbook/Signal-To-Noise_Ratio.html

Hi-Fi

"Abbr. for high-fidelity, that is, a system reproducing a full audio frequency spectrum (20 to 20,000 Hz) and a favourable Signal-to-Noise Ratio...

...**the hi-fi environment is one where all sounds may be heard clearly without being crowded or masked by other sounds and noise**[8]."

Lo-Fi

"Abbr. for low-fidelity, that is, a system which reproduces less than a full frequency spectrum, and which has a poor Signal-to-Noise Ratio...

...**the lo-fi environment is one in which signals are overcrowded, resulting in masking and lack of clarity**[9]."

Masking

"The effect one sound has on another by making it harder or impossible to hear[10]."

Other Resources:

Handbook for Acoustic Ecology

EARS (ElectroAcoustic Research Site)

[8] http://www.sfu.ca/sonic-studio/handbook/Hi-Fi.html

[9] http://www.sfu.ca/sonic-studio/handbook/Lo-Fi.html

[10] http://www.sfu.ca/sonic-studio/handbook/Mask.html