Audio

Affects Quality of Audio

- Ears do not respond to sound in a linear fashion
- Decibel (dB) a logarithmic measurement of sound
- 16-Bit has a signal-to-noise ratio of 98 dB | virtually
- inaudible
- 8-bit has a signal-to-noise ratio of 50 dB
- 6 dB increment is twice as loud

Audio Sample Rate and Bit Size Examples

File Type	Audio File (all mono)
44Hz 16 bit	
44KHz 8-bit	
22 KHz 16-bit	
22KHz 8-Bit	
11KHz 8-bit	

Digital Audio Formats

It is important to distinguish between the audio coding format, the container containing the raw audio data, and an audio codec. A codec performs the encoding and decoding of the raw audio data while this encoded data is (usually) stored in a container file. Although most audio file formats support only one type of audio coding data (created with an audio coder), a multimedia container format (as Matroska or AVI) may support multiple types of audio and video data.

Audio Formats Types

- Uncompressed audio formats, such as WAV, AIFF, AU or raw header-less PCM;
  - ☒ WAVe files (record/play sound waves):---
  - ☒ CD - 44,100 times/sec. sampling rate 16bit sound
  - ☒ Stereo- 3.5” HD diskette holds 8 sec.

- Formats with lossless compression, such as FLAC, Monkey's Audio (filename extension .ape), WavPack (filename extension .wv), TTA, ATRAC Advanced

Lossless, ALAC(filename extension .m4a), MPEG-4 SLS, MPEG-4 ALS, MPEG-4 DST, Windows Media Audio Lossless (WMA Lossless), and Shorten (SHN).

- Formats with lossy compression, such as Opus, MP3, Vorbis, Musepack, AAC, ATRAC and Windows Media Audio Lossy (WMA lossy).

Uncompressed audio format

One major uncompressed audio format, LPCM, is the same variety of PCM as used in Compact Disc Digital Audio and is the format most commonly accepted by low level audio APIs and D/A converter hardware. Although LPCM can be stored on a computer as a raw audio format, it is usually stored in a .wav file on Windows or in a .aiff file on macOS. The AIFF format is based on the Interchange File Format (IFF), and the WAV format is based on the similar Resource Interchange File Format (RIFF). WAV and AIFF are designed to store a wide variety of audio formats, lossless and lossy; they just add a small, metadata-containing header before the audio data to declare the format of the audio data, such as

LPCM with a particular sample rate, bit depth, endianness and number of channels. Since WAV and AIFF are widely supported and can store LPCM, they are suitable file formats for storing and archiving an original recording.

BWF (Broadcast Wave Format) is a standard audio format created by the European Broadcasting Union as a successor to WAV. Among other enhancements, BWF allows more robust metadata to be stored in the file. (EBU Technical document 3285, July 1997). This is the primary recording format used in many professional audio workstations in the television and film industry. BWF files include a standardized timestamp reference which allows for easy synchronization with a separate picture element. Stand-alone, file based, multi-track recorders from AETA,[1] Sound Devices, Zaxcom , HHB Communications Ltd, Fostex, Nagra, Aaton, and TASCAM all use BWF as their preferred format.

Lossless compressed audio format

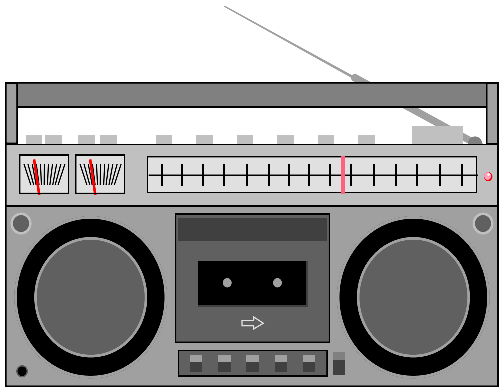
A lossless compressed format stores data in less space without losing any information. The original, uncompressed data can be recreated from the compressed version.

Uncompressed audio formats encode both sound and silence with the same number of bits per unit of time. Encoding an uncompressed minute of absolute silence produces a file of the same size as encoding an uncompressed minute of music. In a lossless compressed format, however, the music would occupy a smaller file than an uncompressed format and the silence would take up almost no space at all.

Lossless compression formats include the common FLAC, WavPack, Monkey's Audio, ALAC (Apple Lossless). They provide a compression ratio of about 2:1 (i.e. their files take up half the space of PCM). Development in lossless compression formats aims to reduce processing time while maintaining a good compression ratio.

Lossy compressed audio format

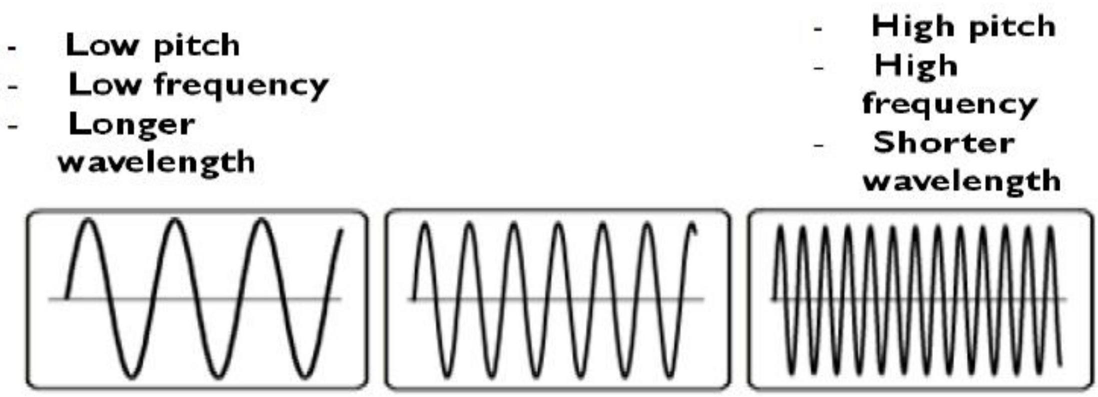
Lossy compression enables even greater reductions in file size by removing some of the audio information and simplifying the data. This of course results in a reduction in audio quality, but a variety of techniques are used, mainly by exploiting psychoacoustics, to remove the parts of the sound that have the least effect on perceived quality, and to minimize the amount of audible noise added during the process. The popular MP3 format is probably the best-known example, but the AAC format found on the iTunes Music Store is also common. Most formats offer a range of degrees of compression, generally measured in bit rate. The lower the rate, the smaller the file and the more significant the quality loss.



Audio Fundamentals:

1. Tone(pitch)

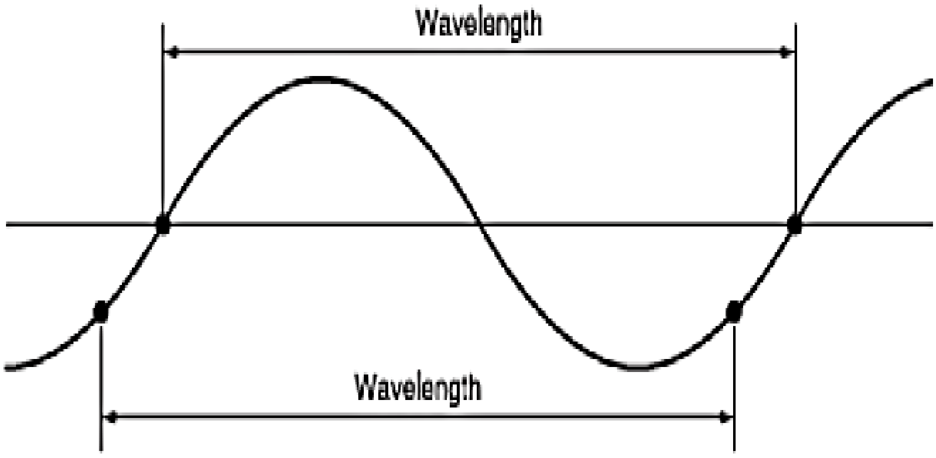
The quality that allows us to classify a sound as relatively high or low.It represents the cyclic, repetitive nature of the vibrations that make up sound. For simple sounds, Tone relates to the frequency of the slowest
vibration in the sound (called the fundamental harmonic).
Sometimes
individuals identify different pitches for the same sound, based on theirpersonal experience of particular sound patterns.



Sound output from the man's throat for that output of the Throat women varies. The fundamental frequency of male voice lies in the range of 85 to 180 Hz while that of the female in the range of 150 to 300 Hz, and for the voice of women is characterized by smooth, and sharpness, while the characterized of male voice rough- and heavy-handed.

2. Wavelength

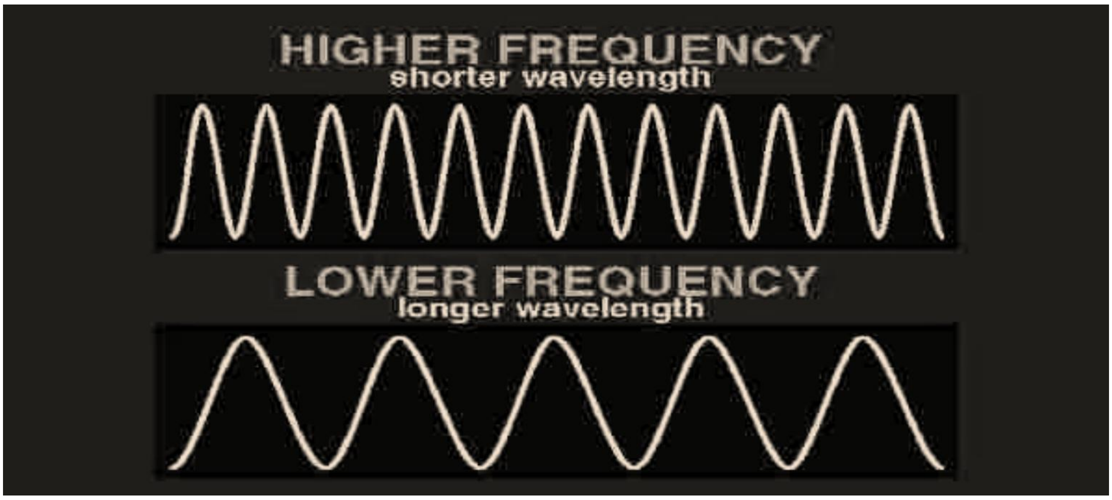
Wave length : is the distance between one part of a wave and the same part of the next wave.

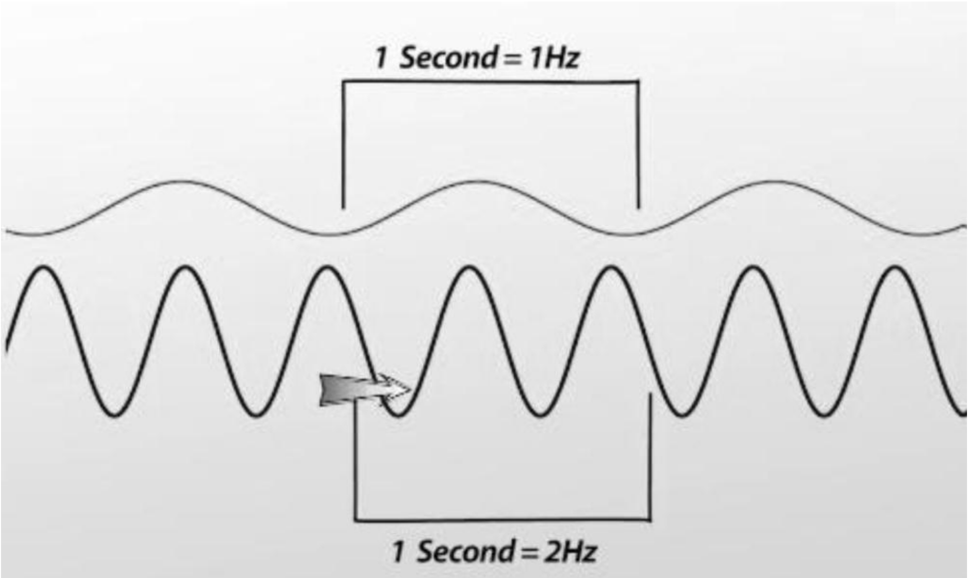


Wavelength is defined as the distance between two points at the same place on adjacent cycl

3. Frequency

Is the number of occurrences of a repeating event per unit of time (number of cycle per second of the wave). The unit of frequency is the Hertz (Hz). The hertz is equivalent to cycles per second, i.e., "1/second". A note that is vibrating at 256 Hz will be caused by sound waves that vibrate at 256 times a second. The generally accepted standard range of audible frequencies is 20 Hz to 20 kHz. Frequencies below 20 Hz are generally felt rather than heard. High frequencies are the first to be affected by hearing loss due to age and/or prolonged exposure to very loud noises.





Sound can be classified depending on the frequency into the types:

1. Infrasound, which is less than 20 Hz and not audible to the human ear, where the frequency is too low, and the human ear cannot sense it. The most important source is a vibrating and sliding movement of the layers of the earth's crust and the resulting earthquakes and volcanoes and therefore it's very important to monitor earthquakes and volcanic activity tracking. Some animals can sense earthquakes before they occur.

2. Range of hearing, which extends from about 20 Hz to 20,000 Hz, which sounds audible to humans,
3. Ultrasound, greater than 20,000 Hz, which is not audible to humans and fall outside the scope of the sense of the human ear. This type of wave is still under consideration and attention given intensive task for applications that affect many areas in industry, medicine and others.

Signal-to-Quantization-Noise Ratio (SQNR)

For digital signals, we must take into account the fact that only quantized values are stored. For a digital audio signal, the precision of each sample is determined by the number of bits per sample, typically 8 or 16.

_ Aside from any noise that may have been present in the original analog signal, there is also an additional error that results from quantization.

(a) If voltages are actually in 0 to 1 but we have only 8 bits in which to store values, then effectively we force all continuous values of voltage into only 256 different values.

(b) This introduces a round off error. It is not really "noise". Nevertheless it is called quantization noise (or quantization error).

The quality of the quantization is characterized by the *signal-to-quantization-noise ratio(SQNR).* Quantization noise is defined as: the difference between the actual value of the analog signal, for the particular sampling time, and the nearest quantization interval value.