# Audio

## Introduction

**Sound** is a type of energy generated by vibration. When an object vibrates, the surrounding air molecules move. These molecules collide with nearby molecules and cause them to vibrate as well. This causes them to encounter closer air molecules. This "chain reaction motion," known as a sound wave, continues until the molecule runs out of energy.

**Frequency**. If your ears are within the range of such vibrations, you will hear the sound. However, to hear the vibration, the vibration requires a certain speed. For example, you may not hear the slow vibrations that occur when you wave your hand in the air. The slowest vibration audible to the human ear is 20 vibrations per second. It will be a very low tone. The fastest vibration we can hear is 20,000 vibrations per second, which is very high-pitched. The frequency per second is called the frequency of the object and is measured in hertz (Hz).

**Pitch** is frequency related. Frequency is a scientific measure of pitch. The frequency is objective, but the pitch is completely subjective. The sound wave itself has no pitch. It is possible to measure their vibrations and give them frequencies, but the human brain needs to map them to their inherent pitch quality. The pitch of the sound is mainly determined by the mass (weight) of the vibrating object. In general, the heavier the mass, the slower the vibration and the lower the pitch. However, you can change the pitch by changing the tension or stiffness of the object. For example, a heavy E-string on an instrument may sound higher than a light E-string by tightening the tuning pegs to tension the strings.

When molecules vibrate in a medium, they can move back and forth or up and down. Sound energy causes molecules to move back and forth in the same direction that sound is propagating. This is called a longitudinal wave. (Shear waves occur when a molecule oscillates up and down perpendicular to the direction of wave propagation).

Speech (like hearing) is accompanied by vibration. When we speak, we move the air across the vocal cords, causing them to vibrate. Stretching these vocal cords changes the sound we make. Stretching the vocal cords produces high-pitched sounds and loosening them produces low-pitched sounds. This is called the pitch.

Text

Description automatically generated

<https://www.scienceworld.ca/resource/sound/>

## Digital audio files

Digital audio files are formats for storing digital audio data on a computer system and digital representation of the sound that can heard on audio device. Audio files contain information about recordings such as voice, music, and even white noise. This information includes how the volume and pitch of these recorded sounds change, and the total recording time. The sound can be very quiet and high-pitched, like the bark of a cat, or loud and deep, like the roar of an explosion. As shown below, the frequency and amplitude of sound can also be visually represented by so-called waveforms.

Diagram

Description automatically generated

<https://www.omnicalculator.com/other/audio-file-size>

All audio files people are working with in the internet has to be stored somewhere. Depending on preferable format each of them displays the same information at different storage sizes or quality levels. In addition, some audio file formats contain metadata that provides information about the file or its contents.

## Lossless and lossy audio files

There are two main types of audio file—lossless file formats, and lossy file formats.

The difference between the two has to do with data compression.

Some methods of data compression make the file smaller, but still retain 100% of the information in the raw audio stream. These are known as **lossless** compression formats.

Other types of compression work by eliminating data in the audio that does not significantly affect the sound. These are called **lossy** compression formats because this method discards some information.

## **Uncompressed audio formats** digital audio workstation

There are other audio file formats that do not use data compression. These are called **uncompressed** audio formats. These file types act as containers for raw audio data without affecting size or quality. Uncompressed audio files are most commonly used for recording and mixing music in digital audio workstation. Nevertheless, there are also uncompressed audio files with different quality levels. These are based on the precision and precision with which analog audio signals are converted to digital. The higher the sample rate and bit depth used, the more information will be captured by the conversion process. Bit depth represents the accuracy of the AD / DA converter for measuring the amplitude or loudness level of the signal.

**Sample rate** means how many measurements are made per second. A higher sampling rate means that more individual measurements have been made. measured in hertz ([***Hz***](https://manual.audacityteam.org/man/glossary.html#hz)), or [***cycles***](https://manual.audacityteam.org/man/glossary.html#cycle) per second.

Some of the most common quality levels are listed below:

Graphical user interface, application

Description automatically generated

The most common formats of audio files

There are only few commonly world-wide used audio files.

Audio File Formats: How to Choose the Right File Type

Chart, treemap chart

Description automatically generated

<https://pt.slideshare.net/lesleyw/audio-codec-presentation/5>

MP3

mp3 are the most common file type for general listening. MPEG-1 [Audio](https://www.collinsdictionary.com/dictionary/english/audio) Layer-3: [software](https://www.collinsdictionary.com/dictionary/english/software) that [enables](https://www.collinsdictionary.com/dictionary/english/enable) [files](https://www.collinsdictionary.com/dictionary/english/file) to be [compressed](https://www.collinsdictionary.com/dictionary/english/compress) quickly to 10% or less of their [original](https://www.collinsdictionary.com/dictionary/english/original) [size](https://www.collinsdictionary.com/dictionary/english/size) for [storage](https://www.collinsdictionary.com/dictionary/english/storage) on [disk](https://www.collinsdictionary.com/dictionary/english/disk) or [hard](https://www.collinsdictionary.com/dictionary/english/hard) [drive](https://www.collinsdictionary.com/dictionary/english/drive) or esp (electronic skip protection) for [transfer](https://www.collinsdictionary.com/dictionary/english/transfer) over the [internet](https://www.collinsdictionary.com/dictionary/english/internet)(https://www.collinsdictionary.com/dictionary/english/mp3).

AAC

An audio coding standard for lossy digital audio compression. Developed as a successor to the MP3 format, AAC typically delivers better sound quality than MP3 encoders at the same bit rate.

https://en.wikipedia.org/wiki/Advanced\_Audio\_Coding

AAC is a standard format for an Apple devices(iPod).

Ogg Vorbis

Newer open-source codec. It is of good quality, free to use and does not require a license fee. It's also an irreversible codec, but much more efficient than MP3s, especially at low data rates.

This format is alternative to lossy compressed formats like MP3. It’s notable for being the file type used for audio material on Wikipedia.

### FLAC

FLAC is an audio coding format for lossless compression of digital audio. This file format provides a bit-perfect copy of a CD, but at half the size. FLAC files make listening to lossless audio possible on devices with limited storage.

https://www.cnet.com/tech/home-entertainment/what-is-flac-the-high-def-mp3-explained/

### WAV

Waveform audio files (also known as WAV files) are one of the most popular digital audio formats and the “gold” standard for studio recording. WAV files are not compressed,the data is stored unchanged in its original format and does not require decoding.

AIFF

AIFF ([Audio Interchange File Format](https://en.wikipedia.org/wiki/Audio_Interchange_File_Format)) works almost the same way: Provides studio quality audio recording and playback. AIFF offers sample rate and bit depth options such as WAV files and registers audio waveforms in the PCM as accurate samples (slices) to provide the best possible audio recording quality and sound reproduction. Like WAV, AIFF stores data in uncompressed, lossless format. In other words, the quality is not compromised, only pure sound enjoyment is obtained.

## How to choose an audio file format

Simple guidelines for choosing an audio file format:

* Use uncompressed audio with high sample rate and bit depth (24-bit / 48kHz WAV or AIFF) for music production
* High bitrate compressed format (320 kbps MP3, AAC etc) - for general listening
* Lossless compressed format (FLAC) - for critical listening,

Tip: In the music production process, if the audio quality does not already exist, it cannot be restored. To get the highest quality files from LANDR mastering, you need to start with high resolution files. It is recommended that you set your DAW session to at least 24-bit / 44.1kHz, or higher if possible.

https://blog.landr.com/audio-file-formats/

## Codec

(**coder**-decoder or compression-decompression) A standard for compressing and decompressing digital media, especially audio and video. Codecs are used to save files to disk and transfer media over a computer network. Rapid compression and decompression of this data reduces bandwidth requirements and increases the amount of interactive and multimedia content accessed and transmitted over the network.

Different types of video are better encoded in different formats, just as some audio codecs better encode human voice and others better encode musical instrument music. In general, the most efficient codecs also require considerable processing power. Multimedia distribution always requires a balance between computing power and bandwidth.

https://www.britannica.com/technology/codec

Embedding audio in HTML documents

Inserting audio into a web page has never been easier, as web browsers did not have a consistent standard for defining embedded media files such as audio. Using HTML5 audio elements. The newly introduced HTML5 <audio > Element provides a standard way to embed audio in web pages. However, although the audio element is relatively new, it works with most modern web browsers

https://www.tutorialrepublic.com/html-tutorial/html5-audio.php

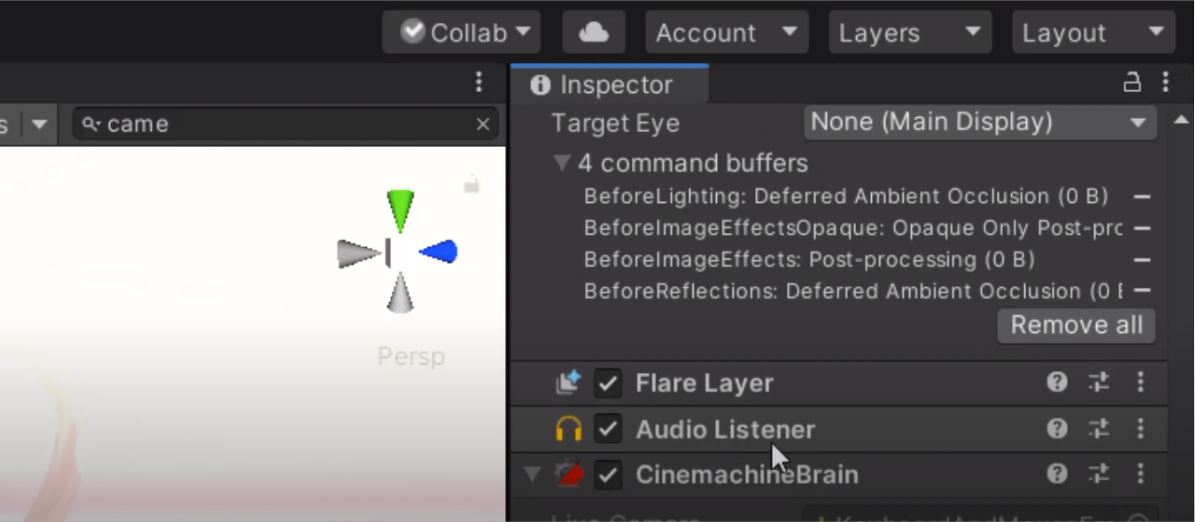
Graphical user interface, text, application, email

Description automatically generated

<https://www.w3schools.com/html/html5_audio.asp>

## Audio in Unity

**Audio Listener**: Think of an audio listener as an element of a game that listens to music. Basically, this is part of the world that captures the sounds created in the game. For example, as a character passes through a portal, the portal may emit a gentle “hum” or electrical noise. Something in the game needs to hear these sounds. In most cases, game developers add audio listeners to their cameras. This is because in most games the camera tracks the character, so the emitted sound is picked up and registered by a camera that shares the player's point of view. Therefore, all game sounds appear to surround the character from the correct position in 3D space.



<https://www.sovereignmoon.studio/how-to-add-music-to-unity-3d-game-kit/>

**Creating Audio Sources.** The audio source is not going to work without the associated audio clip. Clips are the sound files that are actually perform. A source is like a controller that starts and stops playing that clip and changes other audio properties. TO be able to create a new audio source: 1. Import the audio file into Unity project. These are now audio clips. 2. From the menu bar, choose Game Object > Create Empty. 3. With the new Game Object selected, select Components > Audio > Audio Sources. 4. By using Inspector panel on the right side, choose the Audio Clip property on the Audio Source Component and assign a clip, either by dragging one from the Project Window or by clicking the small circle icon to the right of the Inspector property, then selecting a clip from the list. Note: If you only want to create an audio source for just one audio clip in your Assets folder, just drag that clip into the scene view. Graphical user interface, text

Description automatically generated

A Game Object with an audio source component is automatically created. When you drag a clip onto an existing Game Object, the clip will be attached with a new audio source (if it doesn't already exist). If the object already has an audio source, the newly dragged clip will replace the clip currently in use by the source.

<https://docs.unity3d.com/2019.3/Documentation/Manual/class-AudioSource.html>

To determine the most suitable format, the table below was analyzed with various bit rates and sample, the evaluation was carried out by an independent participant.

The analysis was done in the application GoldWave.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Format type | Sample Rate | Bit Rate | File Size | User rating |
| Original MP3 | 44.100Hz | 256Kbps | 502kb | Very good |
| MP3 (ACM) | 8000Hz | 8Kbps | 16kb | Bad, like from the next room |
| MP3(LAME) | 8000Hz | 8Kbps | 16Kb | Terrible |
| MP3(LAME) | 44100Hz | 32Kbps | 64Kb | Bad |
| MP3(LAME) | 44100Hz | 64Kbps | 126Kb | Bad |
| MP3(LAME) | 44100Hz | 128Kbps | 252Kb | Very good |
| MP3(LAME) | 44100Hz | 256Kbps | 252Kb | Very good |
| MP3 | 48.000Hz | 320Kbps | 628kb | Very good, the same as original |
| Wav -16 bit | 44100Hz | 1411Kbps | 2775Kb | Average |
| FLAC-max compression | 44100Hz | 24bit | 3249Kb | Good |
| FLAC-max compression | 44100Hz | 8bit | 670Kb | Good |
| Ogg Vorbis | 44100Hz | 500Kbps | 905Kb | Very good |
| Ogg Vorbis | 44100Hz | 320Kbps | 639Kb | Very good |
| Ogg Vorbis | 44100Hz | 128Kbps | 245Kb | Very good |
| Ogg Vorbis | 44100Hz | 64Kbps | 127Kb | good |
| Ogg Vorbis | 44100Hz | 45Kbps | 48Kb | bad |
| Ogg Vorbis | 8000Hz | 12Kbps | 16Kb | Very bad |

### Summary

The highest marks were received by MP3 and Ogg formats, so it makes sense to choose between them. As bit rate was reducing – quality was reducing as well, but it got noticed especially on MP3 format. Significant different in a quality of a sound appears when bit rate is lowered from 128 Kbps down to 64 Kbps on MP3 format, but there is no change in Ogg format quality. The best combination of format/bit rate / size are

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| MP3(LAME) | 44100Hz | 128Kbps | 252Kb | Very good |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Ogg Vorbis | 44100Hz | 128Kbps | 245Kb | Very good |

Flac format represented a high quality but has a very big size of a file – 3249 kb for a 15 second audio.