

AVSL Guide to Digital Audio

Introduction

Digital Audio is a technology that allows audio signals to be handled and manipulated by microprocessors as a stream of data.

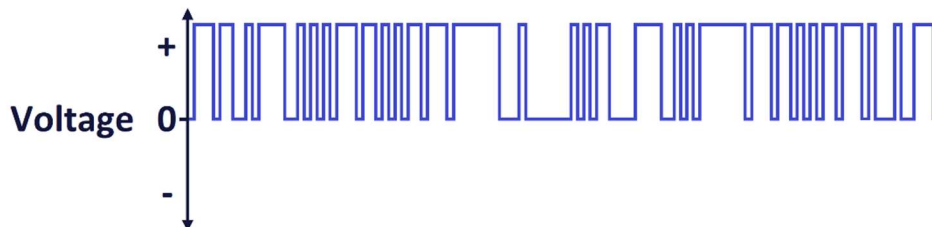
Audio (also described as Analog Audio) is a term to describe sound that has been converted into an electrical signal. In this format, it can be transmitted along wires or over radio waves to a receiving device and then converted back into sound.

**Analog
Audio
Signal**



Digital is a term to describe information that is expressed as numbers (digits). When information is in digital format, it can be manipulated mathematically and transferred between devices without signal degradation.

**Digital
Signal**



Making an Audio Signal

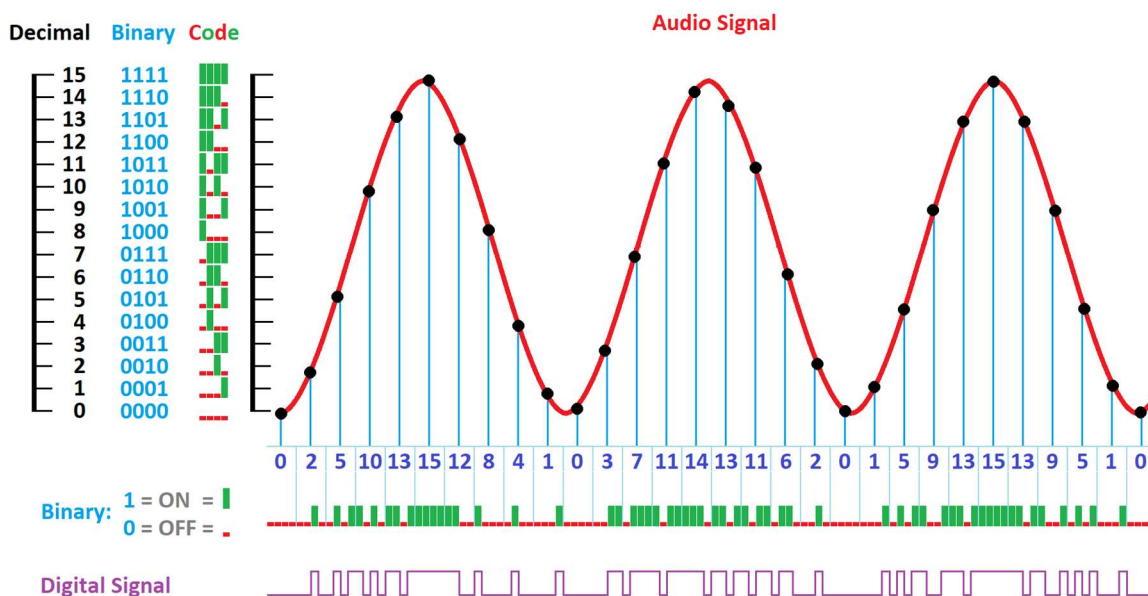
The first step to converting sound into digital information is to convert it into Analog Audio. This is usually done using a microphone or pickup, which turns the air vibrations of natural sounds into small electrical signals. The electrical audio signal is then amplified (made bigger) to make it useful for further equipment to alter the character of the sound or record it and ultimately, play it back through a loudspeaker as a sound wave again.



Converting Analog to Digital

For Digital Audio, there is an intermediate stage where the Analog Audio signal is converted into numerical data, where it can be processed, stored, transmitted and played back before being converted back into Analog Audio ready for being converted back into a sound wave. This process is handled by an electronic device called an Analog-to-Digital Converter (**ADC**)

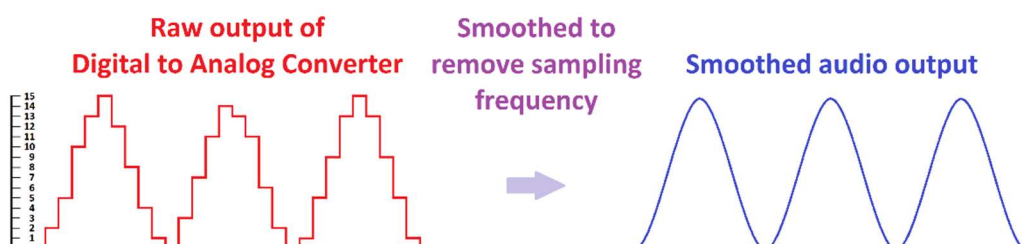
The diagram below shows this process for a sine wave in a 4-bit digital environment. This simplified image uses a sine wave, which is the most basic audio waveform, and 4-bit binary resolution at a sample rate of 10 times the audio frequency just for illustration.



Conversion back to Analog Audio

In reality, an audio signal is usually much more complex than the sine wave shown above, and there would usually be at least 16-bit resolution at 44.1kHz (44,100 times per second). This sample rate is important because it needs to be faster than the human ear can detect. Human hearing can theoretically hear frequencies up to 20kHz and the sample rate ideally should be more than double the audible maximum, so that when it is converted back to Analog Audio, there is no lost detail.

The Digital Audio is transmitted as numbers with staggered jumps between sample values and these jumps happen at the sampling frequency. These are smoothed out by the Digital-to-Analog Converter (**DAC**) as part of its conversion process. It must be converted to analog voltages at some point before it can be converted into sound again.



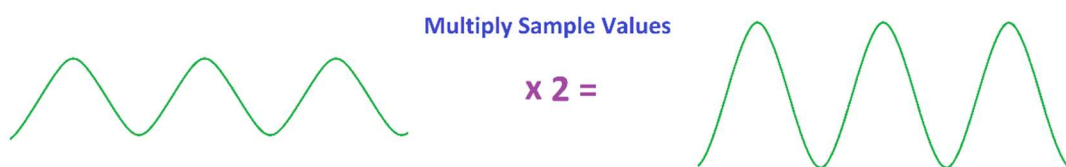


Storage and Transfer

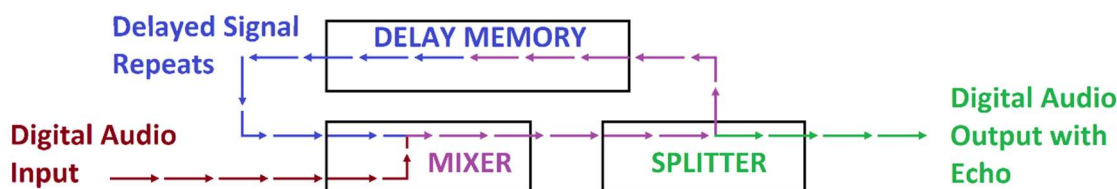
Once the sound information is converted to digital format, it can be stored like any other digital data on a hard disk, Compact Disc (CD), USB or SD flash drive, or onto Digital Audio Tape (DAT). It can also be transmitted along cables or through radio waves in the same way as other digital information. Much of the digital audio produced today resides on servers, which are accessible via the internet for on-demand streaming services to personal smart phones, computers, or standalone audio streaming devices.

Digital Signal Processors (DSPs)

When audio is in digital format, the data can be manipulated by mathematical operations instead of electrical processes. For instance, to double the level of the audio signal, simply double the value of each sample. When it is converted back to audio and then to sound, it will be louder.



Another example is the effect of Digital Delay, which takes a stream of Digital Audio and stores then plays it back slightly later. If this is continually layered over the original signal, it can produce an "echo" effect.



New, faster DSP engines can now change the tonal character or apply spatial or dynamic effects to a signal using sophisticated algorithms, even drastically re-modelling the sound if needed.

Digital Audio and Networks

Digital Audio can be streamed in real-time at the same rate that it is sampled. When playing audio from a CD, the sample rate is 44,100 samples per second, and it is played back continuously at this rate.

Since this is in the form of data, it can also be transferred from one device to another in the same way as all other digital information. To stream audio over a network (or indeed, over the internet), it must share the same feed as all other data on the network. This data is transferred between devices in short bursts called "packets", and audio streaming is no different. The audio stream is transmitted in very small fragments at much higher speed than it would need for playback, time-sharing the bit-stream with any other data being transmitted over the network. When it reaches its intended target, the fragments are re-constructed into a continuous digital audio stream and played back at normal speed.

All of this happens without the listener even realising that it was ever split into tiny pieces and transferred in a shared lightning-fast stream of ones and zeros.

Compressed Digital Audio

The advent of Digital Audio has been a very important development for many industries, including film and television. In traditional film reels, the audio would be printed onto a thin strip at the edge of the celluloid as light and dark patches or shapes, imitating the peaks and troughs of a vinyl record groove. When Digital Audio became available, it could be printed onto the film reel in a similar way with small dark and clear dots representing the ones and zeros, a bit like a tiny continuous QR code.

Since film-makers wanted eventually to digitize both video and audio together, this would require a huge amount of data storage for every scene, so methods were being sought to try to reduce the size of data per project and make it easier to store and transfer the video and audio. This task led to the launch of an alliance called the "*Motion Pictures Experts Group*" (**MPEG**), which oversaw the development of data compression formats for both video and audio.

The techniques used for video and audio data compression aim to minimise the repetitive and predictable elements of the signal to reduce the amount of overall data needed. To describe this using a metaphor...

It takes longer to say... ***"Do this... then do it again... then do it again... then do it again"***...

... than it does to say.... ***"Do this four times"***

In reality, the compacting of this information is much more sophisticated and uses phenomena called "*psychovisual and psychoacoustic redundancy*", which means that our eyes and ears do not scrutinise every pixel of every frame and every frequency of every sound as it is played back. The compression algorithm predicts what we need to see and hear and provides only the necessary data.

The resulting video formats became known progressively as **MPEG-1**, **MPEG-2** and **MPEG-4** (number 3 was scrapped). When stored as a data file, these are given the file extensions **".mpg"** or **".mp4"**.

MPEG encoding also provides additional layers within the data stream for audio, and the most prevalent of these is "MPEG-2 layer 3", which has the file extension **".mp3"**. The high quality and small storage requirements of mp3 has led it to be used extensively outside of the motion picture industry as a valid audio format in its own right. The amount of compression used in mp3 audio can vary (the bit rate can be selected from 96kbps to 320kbps) and lower bitrates are more "lossy" (i.e. losing more detail and quality), but even at the highest bit-rate, it takes up around a fifth of the data that audio stored on a CD would need. This has brought us to the point where we can carry an entire music collection on a very small microSD card to play back through a smart phone or mp3 player.

In addition to **MP3**, further digital audio compression types have been developed, such as **AAC** (proprietary to Apple), **FLAC**, **OGG** and **WMA**. Some formats are lossless (no lost audio detail) but will have larger file sizes. Files with the extension **".wav"** are uncompressed audio and have the largest file size.

Summary

Digital Audio is here to stay and is now used extensively from domestic playback devices to multi-channel networked systems for live events. As processing technology evolves and data storage becomes cheaper and more convenient, new technologies will continue to emerge to take advantage of them for audio and sound processing.

E&OE