

# Digital Representation of sound

## What is a sound?

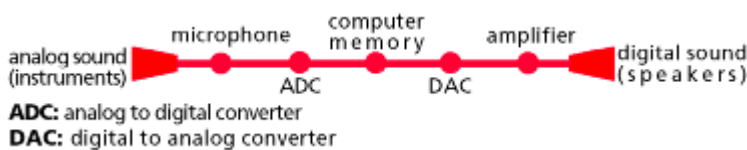


What is a sound from a physical point of view? It is the result of a mechanical disturbance of some object in a physical medium, such as air. This mechanical disturbance generates vibrations that can be represented as electrical signals by means of a device (for example, a microphone), that converts these vibrations into a time-varying voltage.

## Analog vs. Digital Signal

An analog sound signal is the result of measuring the voltage that represents the sound. These kinds of signals are continuous in the sense that they consist of a continuum of constantly changing values. A digital sound is the result of counting all these values many times per second for a certain defined length/time. Each measurement value is called a sample, and this kind of process is called sampling. In order to process sounds on the computer, the analog sound must be converted into a digital format understandable by computer; that is, binary numbers.

## Analog sound into digital sound

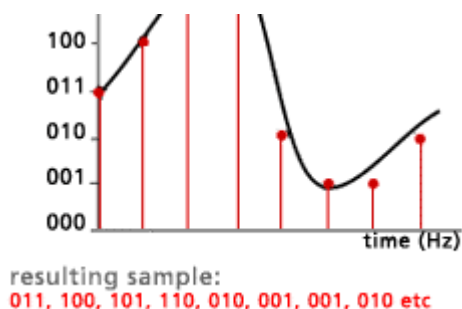


Computers are normally configured to function based upon strings of bits (binary numbers) of fixed size, called words. For example, a computer configured for processing 4-bit words would represent the decimal numbers 0,1,2 and 3 as the binary numbers 0000, 0001,0010 and 0011, respectively. 16-bit conversion suffices for home-quality recording, but nowadays, professional quality recording often uses 32-bit and even 64-bit words. The above figure shows the process of converting a sound (e.g., coming from an instrument such as the violin) into a digital representation: a microphone or a line-level source converts sound into voltage; an analog-to-digital converter then converts voltage into binary numbers; then a computer stores the resulting sound. A digital-to-analog converter converts numbers into voltage and an amplifier relays the voltage to the speakers.

## Converters

Converters are devices used to convert the waveform into a succession of binary numbers, each representing the voltage level at a given instant. This device is called "analog-to-digital" converter (ADC). Another kind of converter, called "digital-to-analog" (DAC), converts binary numbers (i.e., digital sound representation) into analog signals in order to play the sound.

## Sampling theorem

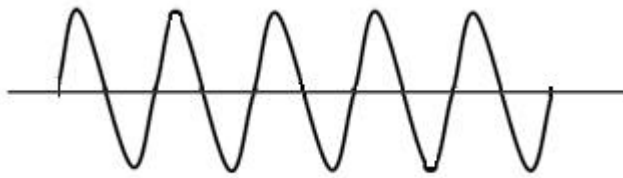


The analog-to-digital conversion process is called sampling. The frequency of a sound is equal to the number of cycles which occur every second ("cycles per second", abbreviated "cps" or "Hz"). In order to convert an analog sound signal into digital representation one needs to sample the signal many times per second. The frequency of this sampling process is called sampling frequency or sampling rate, and it is measured in Hertz (Hz). The sampling theorem states that in order to accurately represent a sound digitally, the sampling rate must be higher than at least twice the value of the highest frequency contained in the signal. The average upper limit

of human hearing is approximately 18 kHz (18000 Hz), which implies a minimum sampling rate of 36 kHz (36000 Hz). The sampling rate frequently used in computer sound design systems is

44.1 kHz (44100 Hz).

## Nyquist frequency and aliasing distortion



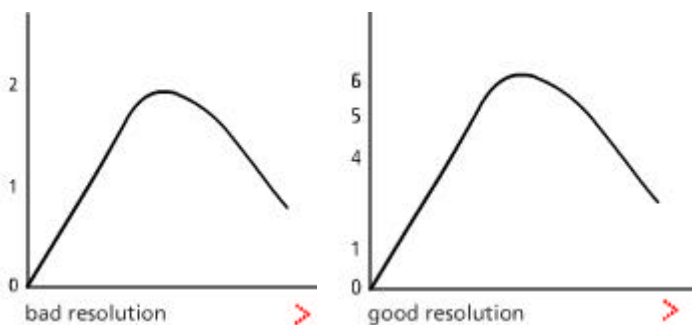
aliasing distortion



*Click on the image*

Nyquist frequency is the name of the highest frequency that can theoretically be represented in a digital audio system. It is calculated as half of the value of the sampling rate. Passing this level causes aliasing distortion, or foldover. Nyquist frequency acts as a mirror, folding over any frequency above it.

## Resolution and quantization noise







*Click on the image*

Another element that influences the quality of sampling is the level of resolution, or quantization of a sampler. The resolution depends upon the size of the word used to represent the amplitude of a sampling sound and is determined by the resolution of the ADC and DAC. A word, for instance, could be 4-bits long or 16-bits long etc. Unsatisfactory lower resolutions are prone to cause a damaging loss of sound quality, referred to as quantization noise.

## Sound examples

The following four sound examples were sampled from a mono sound, 4027 milliseconds long, sampled at 44100 Hz, 11025 Hz, 2756 Hz and 2000 Hz, respectively. The file for the last one uses only 17 kB, but it is considerably distorted when compared to the first example, which uses 367 kB to store the sound.

-  [\[example with good voice\]](#)  
44100 samples per second results in a sound file of 367 kB
-  [\[example with bad voice\]](#)  
11025 samples per second results in a sound file of 93 kB
-  [\[example with distorted voice\]](#)  
2756 samples per second results in a sound file of 23 kB
-  [\[example with very distorted voice\]](#)

2000 samples per second results in a sound file of 17 kB

## Quality and quantity: the right balance

As can be seen above, the ability to adjust the sample rate affects the length of a sample for a given number of samples. This will affect the size of the digital sound file in a storage system. There is a penalty to be paid for the lower rate: a narrower band of frequencies will be recorded. The problem is that acoustic sounds are in fact composed of many sound components called partials. In the majority of cases, most of these partials are of very high frequencies. If the sampling frequency is too low, the computer will not capture these high frequency components and this causes distortion.

## File formats and compression

After digital conversion the sound needs to be stored in order to be re-used or played back. The most basic way to store a sound is to take the stream of samples and write them onto a file. Sound files normally contain other descriptive information related to file properties, text comments and cue pointers; all this information is stored in the sound file header (the initial portion of data of the file). In the header one may find information such as the sampling rate used, the size of the word, whether the sound is mono or stereo, and so on. Frequently used sound file formats are:

- Wave, adopted by Microsoft (.wav)
- VOC, adopted by Creative Lab's Sound Blaster (.voc)
- NeXT/Sun, originated by NeXT and Sun computers (.snd and .au)
- AIFF, originated by Apple computers (.aif)
- AVR, adopted by Atari and Apple computers (.avr)

This kind of file storage is often uneconomical as it might contain a great deal of redundant information like, for example, a silent portion. There are techniques for optimising the representation of samples in order to reduce the size of the file. One of the most popular methods is MPEG3 (or simply MP3) that works by eliminating sound components not normally audible to humans. This compression method can reduce the size of a Wave or Aiff file to up to one-twelfth of its original size. Other compression schemes widely used today are:

- Real Audio by Real Network (also supports live audio over the Internet)
- Atrac3 by Sony
- WMA by Microsoft (also supports live audio over the Internet)

Sound artists and composers normally do not work with compressed sounds. Compression should be applied only after the piece is completed.

2003©Copyright

All the materials on this Web site are subject to copyright protection. Any form of copying without the express permission of the copyright holder, except such copying as may be permitted under the applicable copyright laws, is strictly prohibited. Redistribution of these materials without the permission of the copyright holder is forbidden, including but not limited to, posting, e-mailing, faxing, archiving in a public database, redistribution via a computer network, or in printed form.