

Sampling Theory

Counting in Binary

Binary describes a numbering scheme in which there are only two possible values for each digit: 0 and 1. The term also refers to any digital encoding/decoding system in which there are exactly two possible states. In digital data memory, storage, processing, and communications, the 0 and 1 values are sometimes called "low" and "high," respectively.

A bit (short for binary digit) is the smallest unit of data on a computer; each bit has a single value of either 1 or 0.

Binary numbers look strange when they are written out directly. This is because the digits' weight increases by powers of 2, rather than by powers of 10. The highest represented value is located to the left, and the lowest to the right.

A Bit = A **B**inary Digit

A Byte = 8 Bits

Binary System	4 Bit System			
	8	4	2	1
Decimal Number	0	0	0	0
	1	0	0	1
	2	0	0	1
	3	0	0	1
	4	0	1	0
	5	0	1	0
	6	0	1	0
	7	0	1	1
	8	1	0	0
	9	1	0	0
	10	1	0	1
	11	1	0	1
	12	1	1	0
	13	1	1	0
	14	1	1	1
	15	1	1	1

Sampling Rates

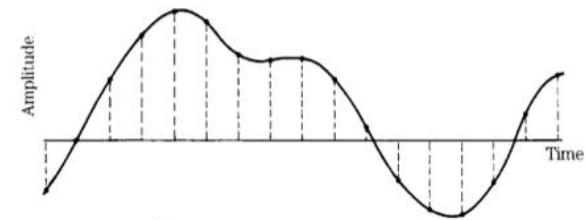
In order for a computer to record and represent audio in a digital format, it has to take a series of samples of a binary value input at regular time intervals. You can think of this as a series of 'snapshots' of the input values, or like stills on an old film reel.

Common sampling rates found within music technology are:

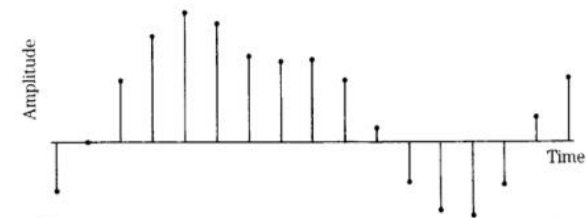
44.1 kHz 48 kHz 88.2 kHz 96 kHz 192 kHz

In practice, this means the computer is taking between 44,100 and 192,000 samples (depending on sample rate) every second.

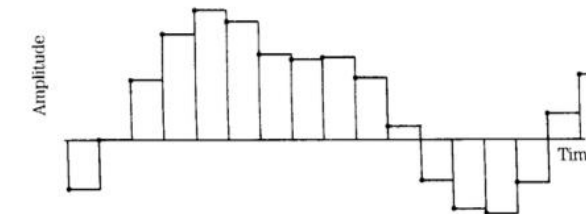
The higher the sampling rate is, the more accurate the digital representation will be.



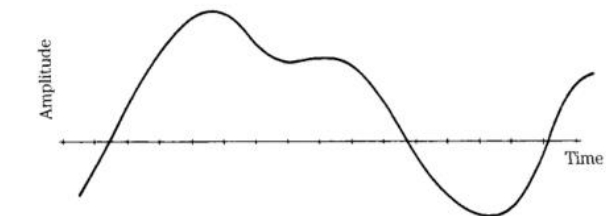
A The input analog signal is sampled.



B The numerical values of these samples are stored or transmitted (effect of quantization not shown).



C Samples are held to form a staircase representation of the signal.



D An output lowpass filter interpolates the staircase to reconstruct the input waveform.

2-1 With discrete time sampling, a bandlimited signal can be sampled and reconstructed without loss because of sampling.

Why these specific sampling rates?

All audio signals can be broken down into a series of discrete sine waves of various frequencies and amplitudes. This rather extraordinary fact allows us to, for example, create simple harmonic waveforms (such as square, sawtooth and triangle) by adding up partials that are integer multiples of a fundamental frequency, and adjusting their relative amplitudes to produce the desired timbre.

More complex waveforms, such as for example a high quality, resonant cello bowed slowly, with varying degrees of pressure on the string, and with varying degrees of performer vibrato, result in complex, time-varying interactions between the resonant frequencies of the sound. To capture this performance, we must be certain to sample the highest audible resonant sine wave frequency at a minimum of twice the frequency of the wave — this allows us to sample both the peak and the trough of the wave and thereby capture both its amplitude and frequency.

Since the highest audible frequency is, to choose a convenient number, 20,000 Hz (cycles per second), it follows then that we'd need to sample a musical signal at twice that frequency, or 40,000 samples per second.

In the digital audio field, this is known, famously, as the **Nyquist Frequency**, after Harry Nyquist who developed these concepts way back in the 1920s. We can define this as follows (this taken directly from Pellman):

The highest frequency that can be adequately represented at a given sample rate— that is, the frequency whose value is half that of the sampling rate — is often referred to as the Nyquist frequency. So long as the Nyquist frequency is greater than or equal to the highest audio frequency to be represented, the signal will be reasonably accurate and free of misleading information.

In step one, above, in the digital audio conversion chain, we use a low pass filter to filter all frequencies above 20,000 Hz, and then we sample that signal at the Nyquist frequency, or 40,000 Hz (40 kHz) to obtain an accurate representation of the highest possible signal. The highest possible sine wave frequency then, the 20,000 hz signal, is *critically sampled* at twice the frequency. It is sampled at the Nyquist frequency.

NOTE: Since no low-pass filter can filter as a "brick wall," cutting off completely every frequency above 20,000 hz, a sloping filter is used to more gradually filter the high frequencies above 20,000. As CD audio rates were being devised by Sony and Phillips, a rate of 44.1 kHz was chosen because they were able to more easily modify existing video equipment to create audio sampling technologies. This is the reason for seemingly random rate of 44.1 (as opposed to say 44).

Bit-Depth

Going hand in hand with sampling rate, is bit depth. Sampling rate tells you how many time to take a sample per second, and bit depth represents the number of available values to assign the digitised input to. For example, there are two commonly used bit depths within music technology.

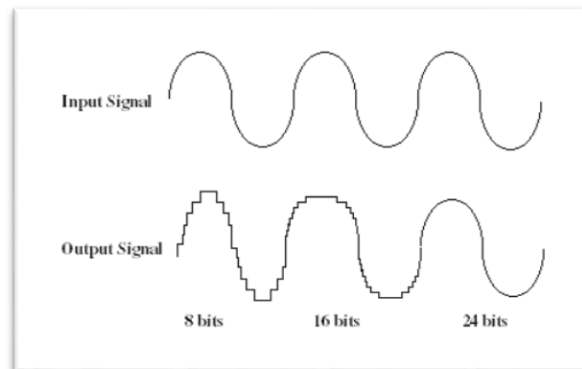
16 bit

24 bit

A system using a 16 bit bit-depth will allow each sample to be mapped over a range of 65,535. Whereas a system using a 24 bit bit-depth will allow a value between 0 and 16,777,215.

The diagram opposite shows how a change in bit depth changes the resolution and accuracy when sampling an analogue signal.

Notice how the size of the wave does not change when a higher bit-depth allowing more values is used. Instead, the accuracy (resolution) of the digital representation is increased.



Number of Bits	Binary Values	Highest Possible Value
1	1	1
2	2	3
3	4	7
4	8	15
5	16	31
6	32	63
7	64	127
8	128	255
9	256	511
10	512	1023
11	1024	2047
12	2048	4095
13	4096	8191
14	8192	16383
15	16384	32767
16	32768	65535
17	65536	131071
18	131072	262143
19	262144	524287
20	524288	1048575
21	1048576	2097151
22	2097152	4194303
23	4194304	8388607
24	8388608	16777215

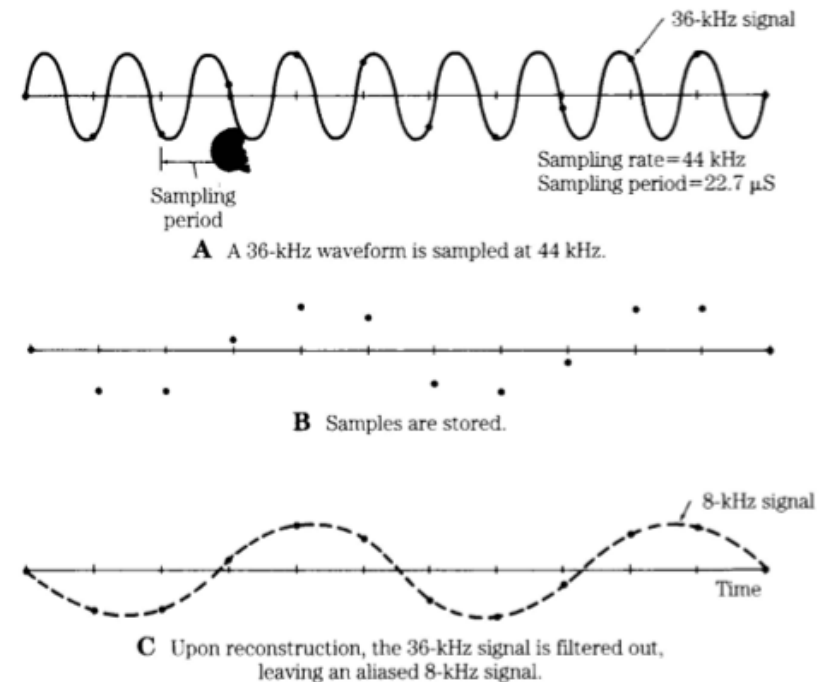
Aliasing

You might ask: What happens when the sampling rate is lower than the highest frequency of the input signal? In short, a misrepresentation of the source input is created. This is called aliasing.

Aliasing is well known in the film world. Film is in fact an excellent analogy for digital audio. Film takes discrete samples (snapshots) of the visual world at a steady frame rate. When that continuous sequence of still images is played back at the same rate, it accurately "plays back" the visual image.

However, when capturing a moving vehicle (classically, it was known as the wagon wheel effect), the movement of the spokes of the wheel were seldom at a resonant frequency with the frame rate of the camera, resulting in a playback that would sometimes appear that the wheel was moving backwards while the vehicle was moving forwards.

Unless we filter an incoming signal at around 20 kHz (20,000 cycles per second), we end up aliasing a frequency that we can't hear, and it appears down in the frequency domain as an audible signal that was not a part of the original waveform. Often, this signal is non-harmonic to the original waveform.



Calculating File Size

File Size in Bytes =

(Sample Rate x Bit Depth x Time x Channels) / 8 bits

Eg. A 30 second stereo recording at 16bit, 44.1kHz

= (44100 x 16 x 30 x 2) / 8

= 42,336,000 / 8

= 5,292,000 Bytes

= 5167.97 KB

= 5 MB

Calculate:

1. 3 minute, stereo, 24bit, 48kHz

2. 5 minute, stereo, 16bit, 44.1kHz

3. 1 hour, octophonic, 24bit, 192kHz

Unit	Value
bit	a binary digit
Byte	8 bits
KB	1000 (1024) Bytes
MB	1000 (1024) KBytes
GB	1000 (1024) Mytes
TB	1000 (1024) GBytes