CS 583 – Computational Audio – Fall 2021

Wayne Snyder
Computer Science Department
Boston University

Lecture One

Overview of physical properties of sound waves: frequency, amplitude, phase, intensity, loudness.

Measuring loudness: logarithmic scales and decibels.

Basics of digital signals and sampling: Quantization error, aliasing, and choice of bit depth and sample rate; dynamic range; the Nyquist Theorem.

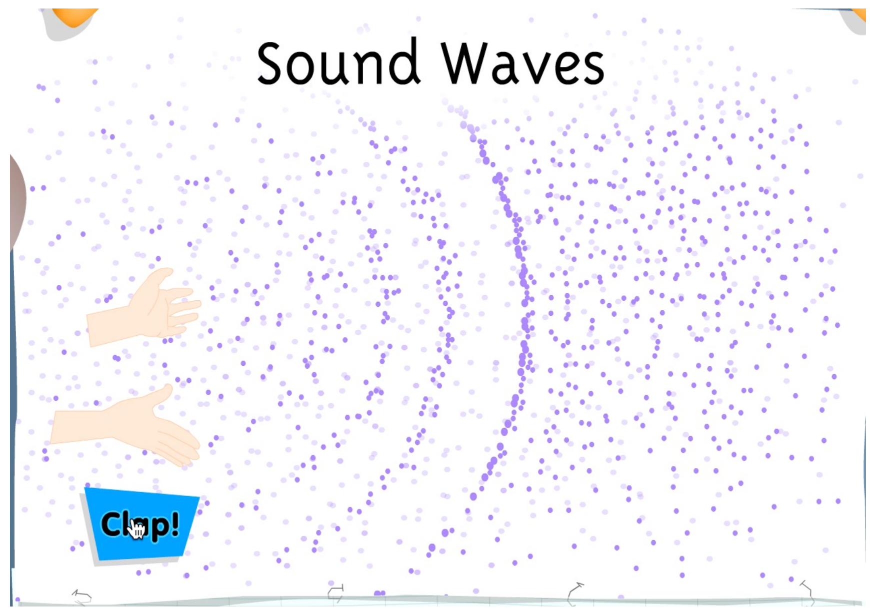
Computer Science

Physical Basis of Sound

Computer Science

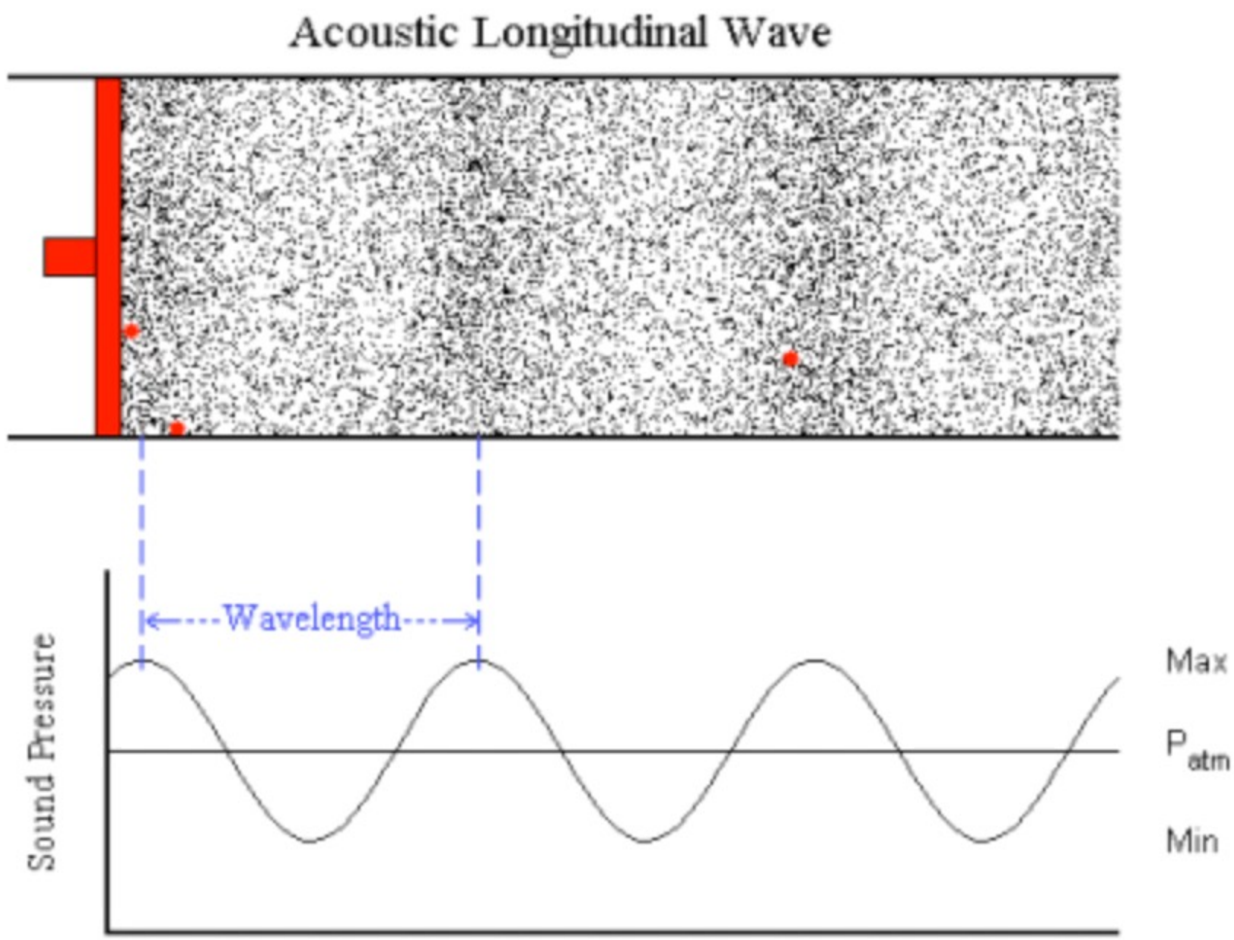
Sound is produced by vibrating objects which produce pressure waves in air, traveling at 343.21 meters/sec (768 mph, or a mile in 4.69 seconds), which are sensed by the ear and interpreted by the brain.

These waves are **longitudinal waves** (the motion is along the direction of travel), as opposed to transverse waves (motion is at right angles to the direction).



Physical Basis of Sound

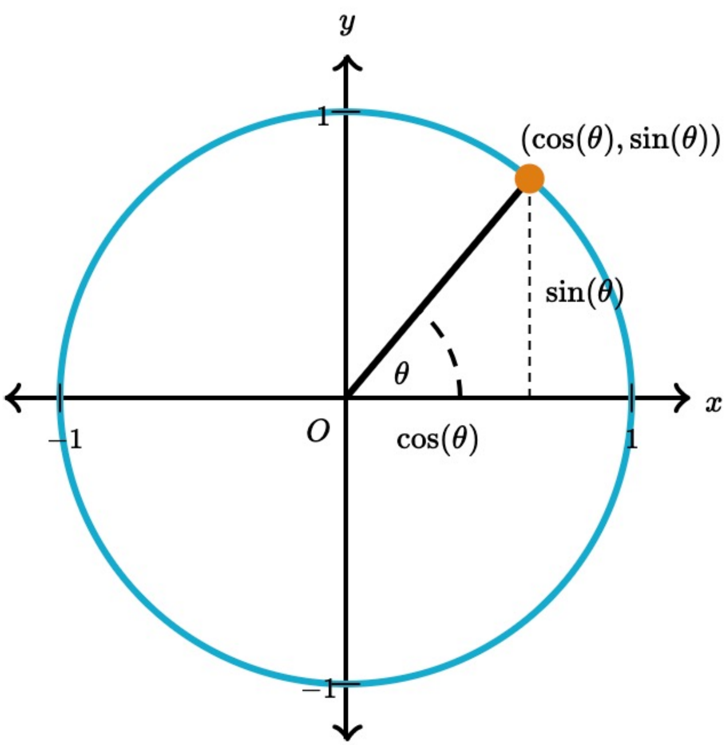
Time Domain Representation: If we record the atmospheric pressure over time, we get a curve with amplitude (in psi) on the y axis and time (in seconds) over the x axis in the shape of the sin (or cos) function:



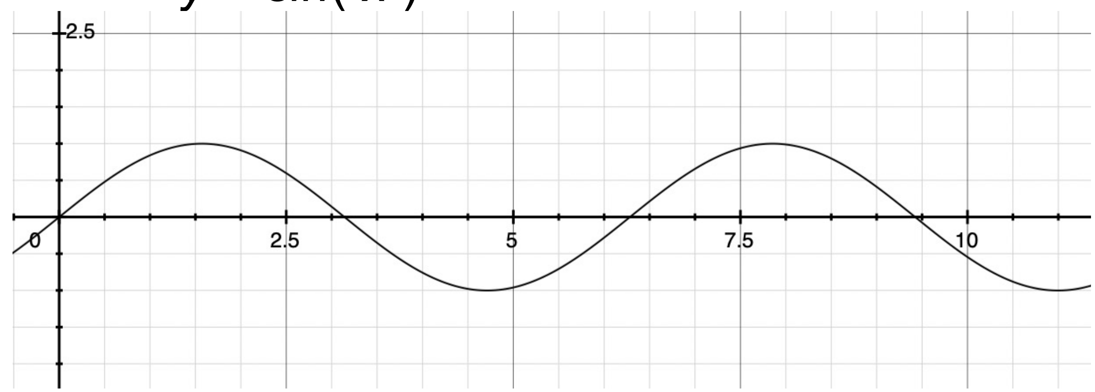
$y = sin(x)$

Physical Basis of Sound

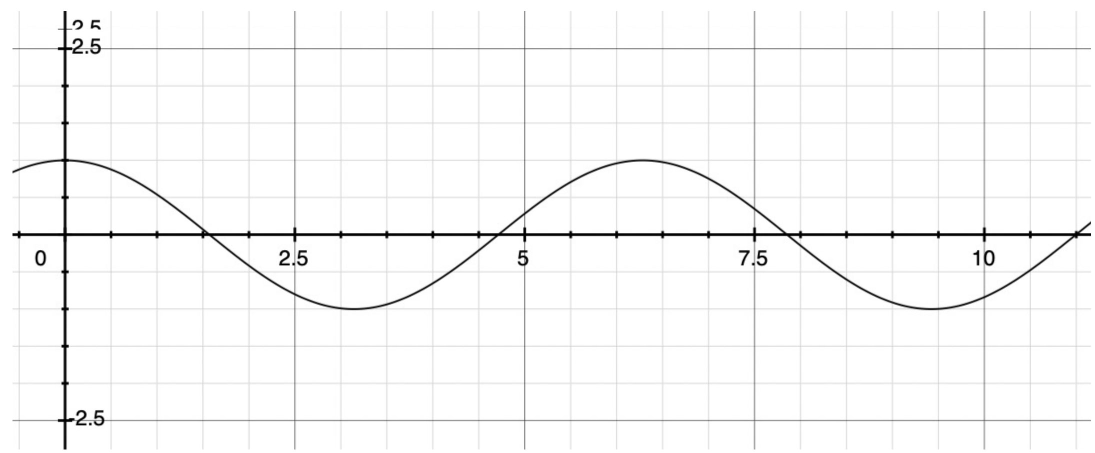
Sine (and Cosine) waves are the fundamental components of musical sounds.



y = sin(x)

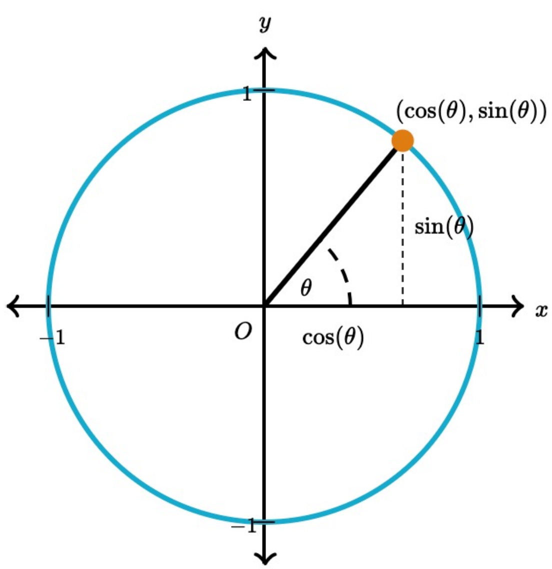


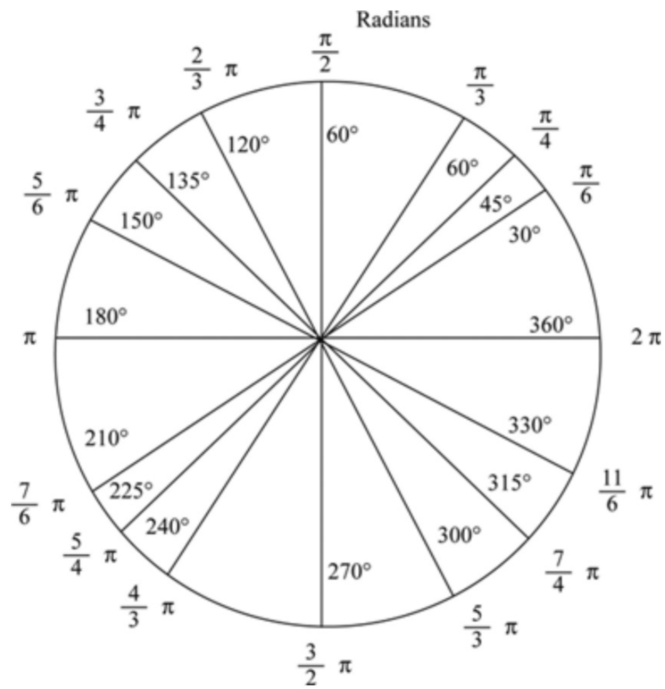
y = cos(x)

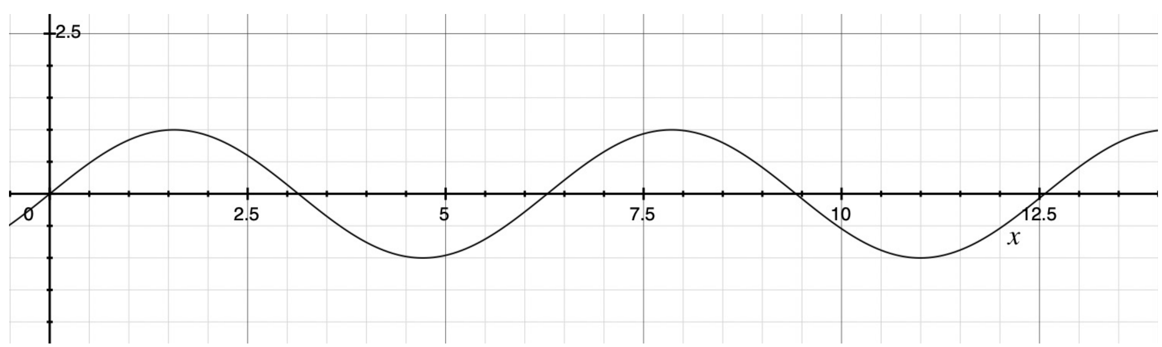


Physical Basis of Sound

Recall that the amplitude of a sin (or cos) wave is 1, and the x axis is expressed in radians.

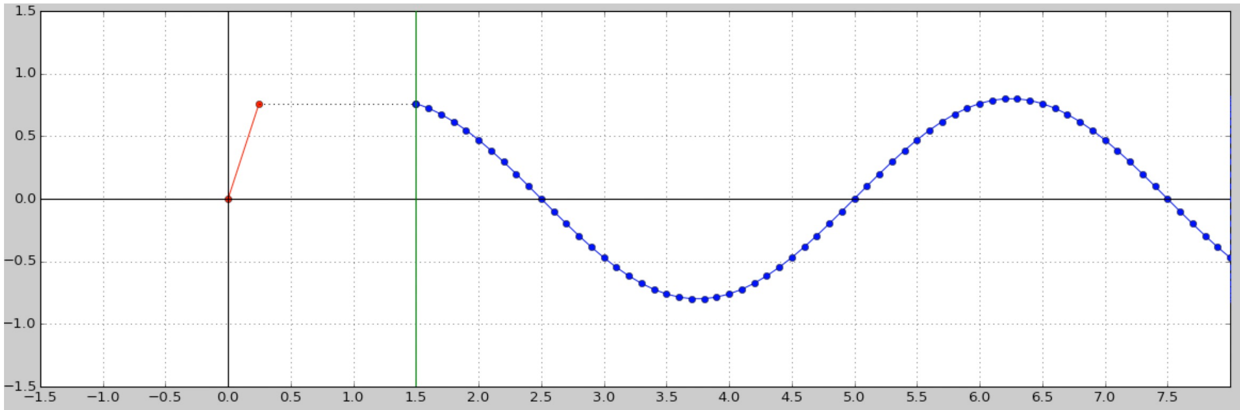




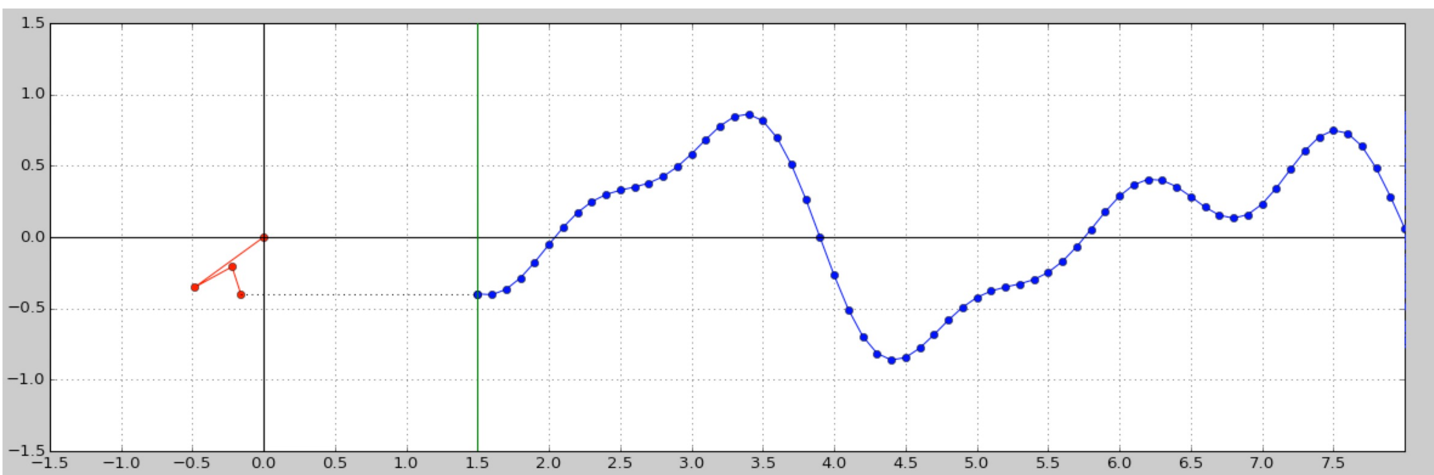


Sine Waves

Phasors: Sine waves can also be produced by rotary motion; this is a convenient way to describe them using phasors (we will return to this later):



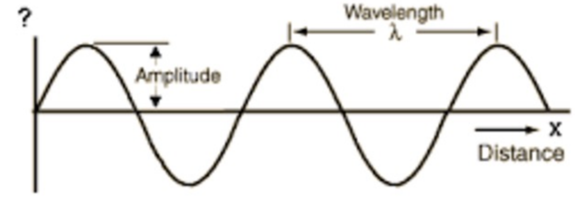
Sinusoids: In fact, most sound waves are more complex, being composed of many individual sine waves added together, such a wave is called a sinusoid:



Properties of (Pure) Sine Waves

Wavelength (λ):

Distance between peaks of a wave (affects pitch -- high or low sounds); measured in meters.

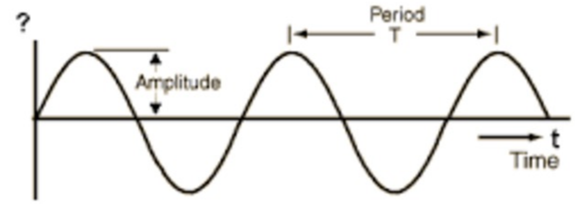


Period (p):

The time between peaks, measured in seconds.

Amplitude (A):

magnitude of the wave above the midpoint (x axis).



Frequency (f):

the number of times a wave occurs in a second. Measured cycles per second or Hertz (Hz) or KiloHertz (kHz).

Some important relationships: (where $v = 343.21$ m/s)

$f = 1 / \text{p}$ **Ex:** 10 Hz = 10 cycles/sec; p = 1/10 sec

$f = v / \lambda$ or $\lambda = v / f$

Ex: Middle C = 261.63 Hz

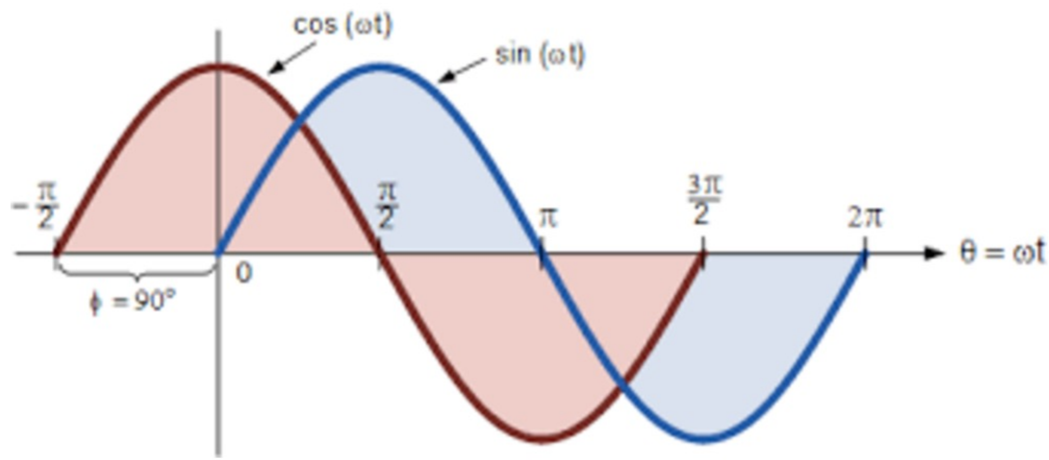
$\lambda = 343.21 \text{ m/s} / (261.63 / \text{s}) = 1.31$

m

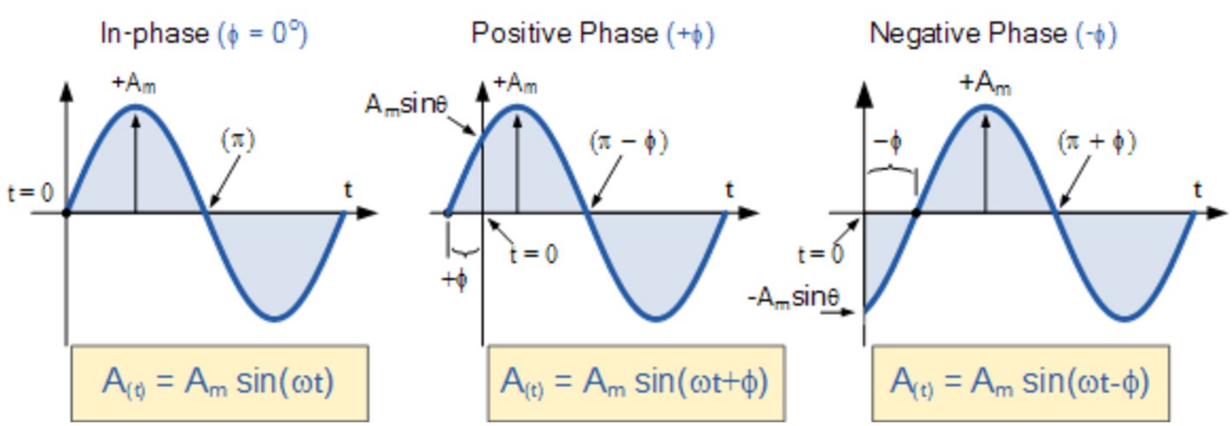
Properties of (Pure) Sine Waves

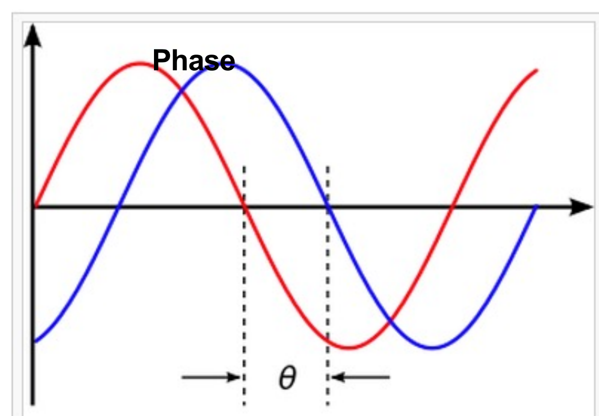
Phase (ϕ):

Where in the oscillation the wave is at a particular point in time (when it starts, or during its oscillation, relative to other waves), measured as offsets around a circle, in radians.



We may also speak of the phase difference between two waves:





Sound Wave Properties: Frequency

Frequency is an absolute measure, and is strongly related to but not absolutely identical to the notion of pitch.

Pitch = **perceived** frequency of a sound

Pitch is subjective, and inaccurate at extremes of frequency or amplitude, and not measured precisely by the ear throughout the range of hearing.

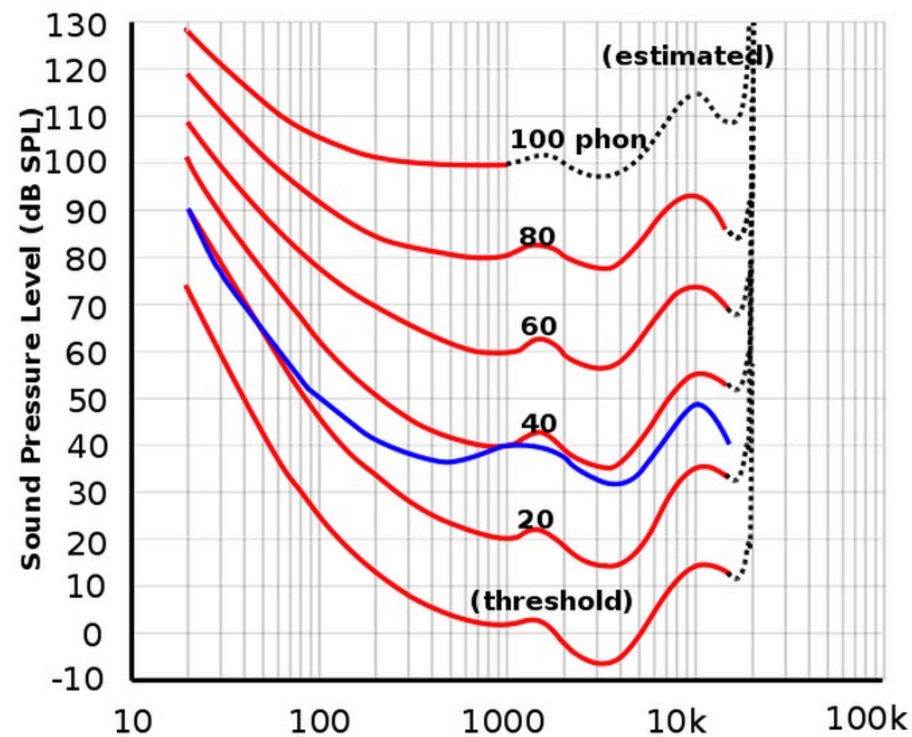
Human: 20Hz – 20 kHz

Dogs: 67 Hz – 44 kHz

Cats: 55 Hz – 79 kHz

Bats: 1Hz – 200 kHz

Sensitivity of human ear at various frequencies; curves represent impressions of equal loudness at various frequencies:



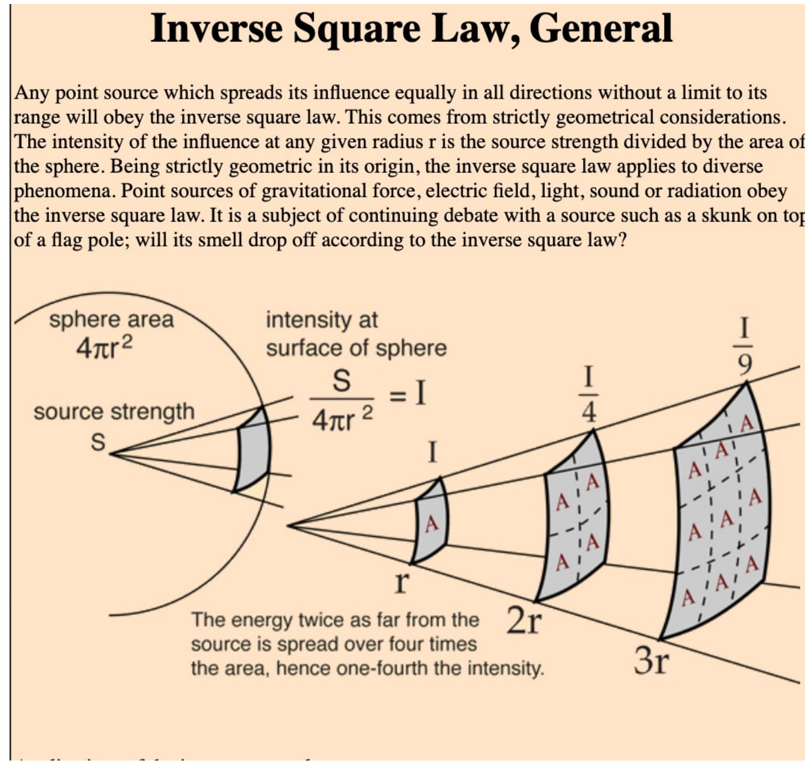
Sound Wave Properties: Intensity and Loudness

Sound Intensity

Notice that sound waves carry energy. We define the intensity I as the rate at which energy E flows through a unit area A perpendicular to the direction of travel of the wave.

Intensity = Power / Area

I = P / A = E / (At)



Inverse Square Law, General

Any point source which spreads its influence equally in all directions without a limit to its range will obey the inverse square law. This comes from strictly geometrical considerations. The intensity of the influence at any given radius r is the source strength divided by the area of the sphere. Being strictly geometric in its origin, the inverse square law applies to diverse phenomena. Point sources of gravitational force, electric field, light, sound or radiation obey the inverse square law. It is a subject of continuing debate with a source such as a skunk on top of a flag pole; will its smell drop off according to the inverse square law?

Thus, Intensity is proportional to the square of the amplitude:

$I = c * A^2$ (c is a constant depending on properties of medium)

Since ear drums tend to be similar in area, often Intensity and Power are used as equivalent terms.

Sound Wave Properties: Intensity and Loudness

Computer Science

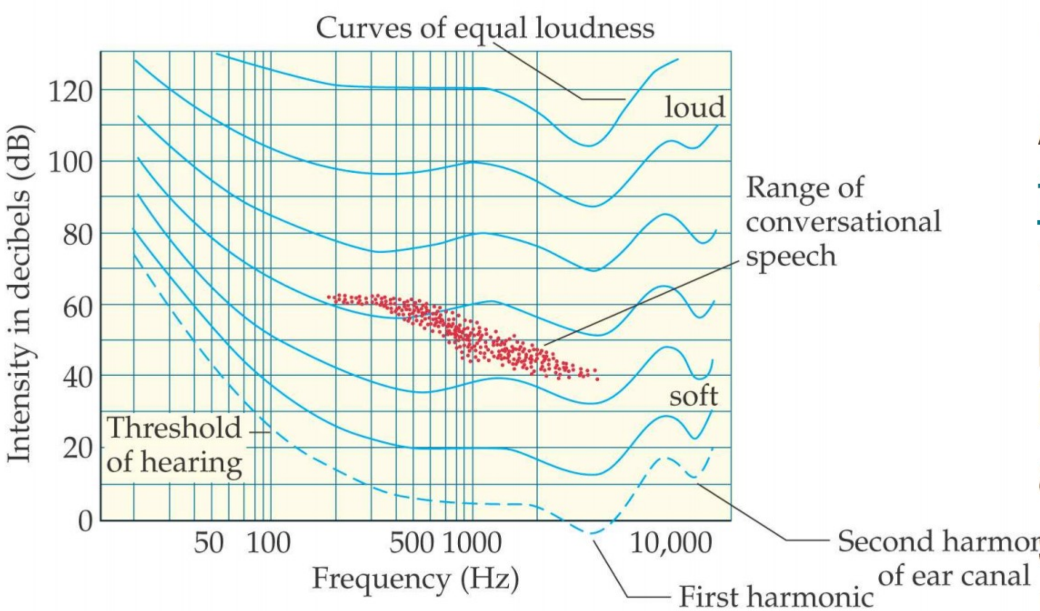
Loudness is not simply sound intensity!

Loudness of a sound is measured by the logarithm of the intensity.

log here is to base 10

The threshold of hearing is at an intensity of 10^{-12} W/m^2.

Sound intensity level is defined by $\beta = (10\text{dB})\log\frac{I}{I_0}$



dB are decibels

A general "rule of thumb" for loudness is that the power must be increased by about a factor of ten to sound twice as loud.

Sound Wave Properties: Intensity and Loudness

Computer Science

A decibel is one tenth of a **bel** (**B**). Devised by engineers of the Bell Telephone Laboratory to quantify the reduction in audio level over a 1 mile (1.6 km) length of standard telephone cable, the bel was originally called the *transmission unit* or *TU*, but was renamed in 1923 or 1924 in honor of the laboratory's founder and telecommunications pioneer Alexander Graham Bell. In many situations, however, the bel proved inconveniently large, so the decibel has become more common.

$$\beta = (10\text{dB})\log\frac{I}{I_0}$$

$$\beta = (10\text{dB})\log\frac{I_0}{I_0} = 0$$

$$\beta = (10\text{dB})\log\frac{10I_0}{I_0} = 10dB$$

$$\beta = (10\text{dB})\log\frac{100I_0}{I_0} = 20dB$$

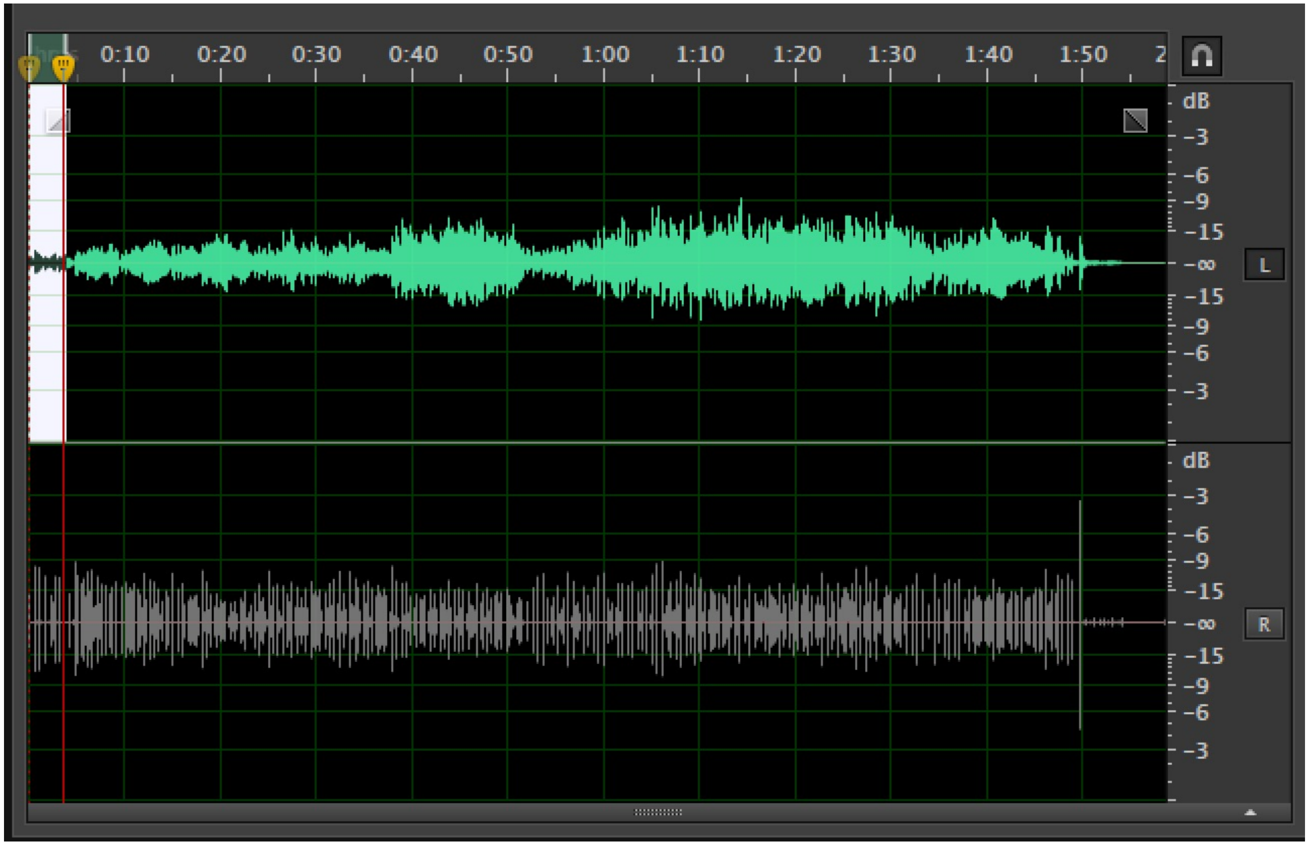
$$\beta = (10\text{dB})\log\frac{1000I_0}{I_0} = 30dB$$

33

Note that dB measures intensity, NOT sound pressure level. When you see dB on the Y axis, you know you are looking a a signal where the amplitude is Intensity (roughly the square of the sound pressure).

Sound Wave Properties: Intensity and Loudness

When dealing with loudness in an abstract setting (say in editing software) usually we are not concerned with the low threshold of hearing, but with the maximum amplitude of the signal; in this case, loudness is measured in negative decibels compared with the maximum loudness:

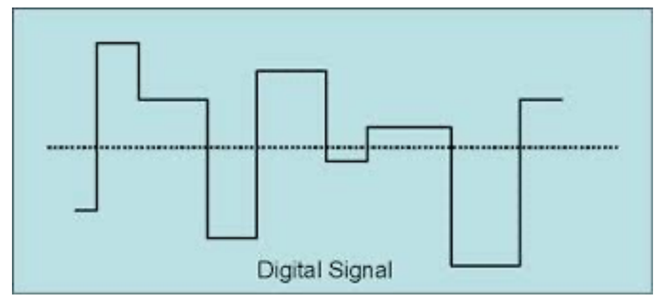


We will explore this in the first homework…..

Digital Audio: Analog vs digital signals

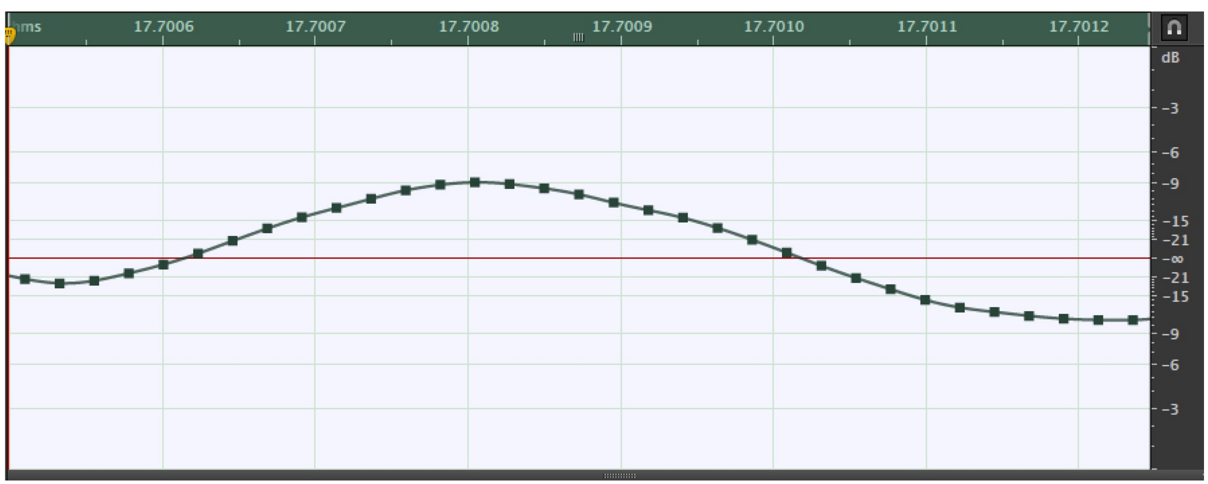
An analog signal continually fluctuates in up and down in the real domain (time is a real number and amplitude is a real number).

A digital signal has a discrete number of amplitudes over a discrete number of time steps, T = 0, 1, 2, ….



Typically, represented as **samples** taken at a regular **sample rate**, e.g.,

16-bit integers **44,100 times a second.**

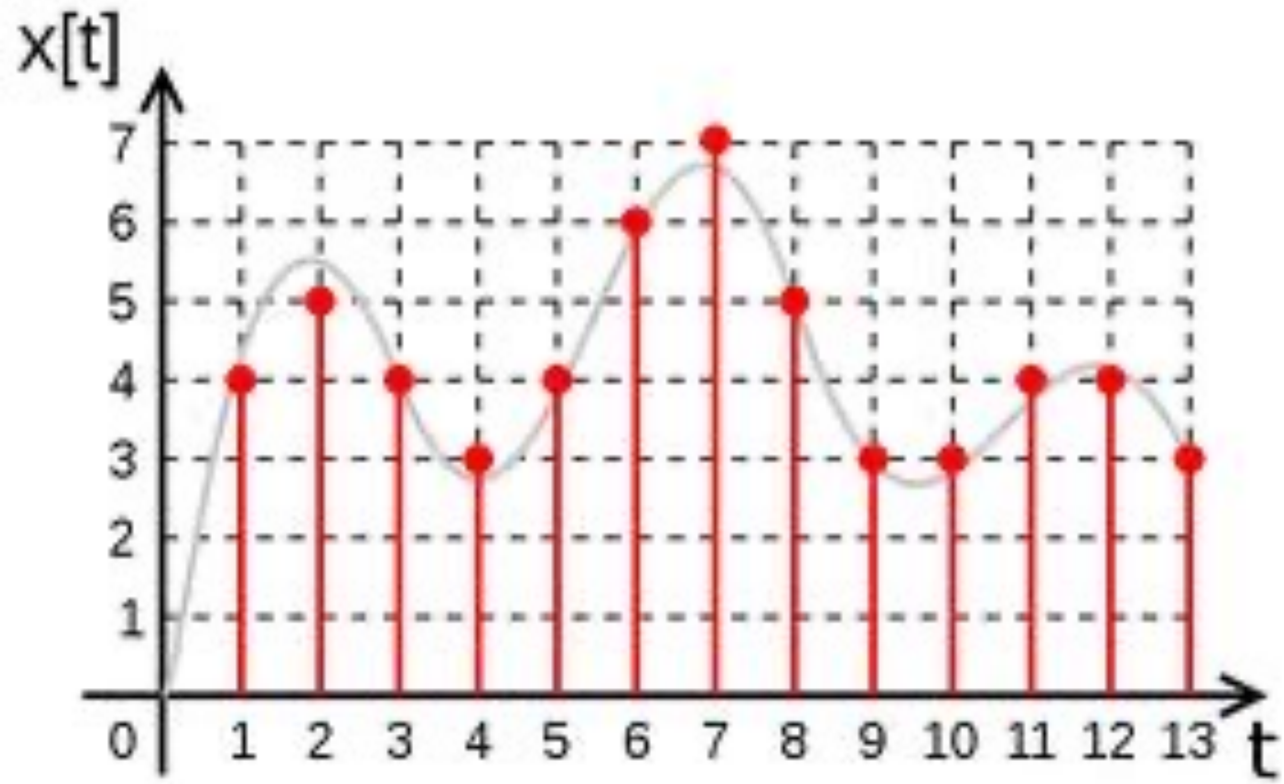


Digital Audio: DAC and ADC

Computer Science

The MOST IMPORTANT concept to understand at the outset is how sound (analog) is digitized to create a digital signal. This is usually called “Analog to Digital Conversion” (ADC)

Turning a signal signal back into an analog signal “Digital to Analog Conversion”” = DAC) for playback is relatively easier!



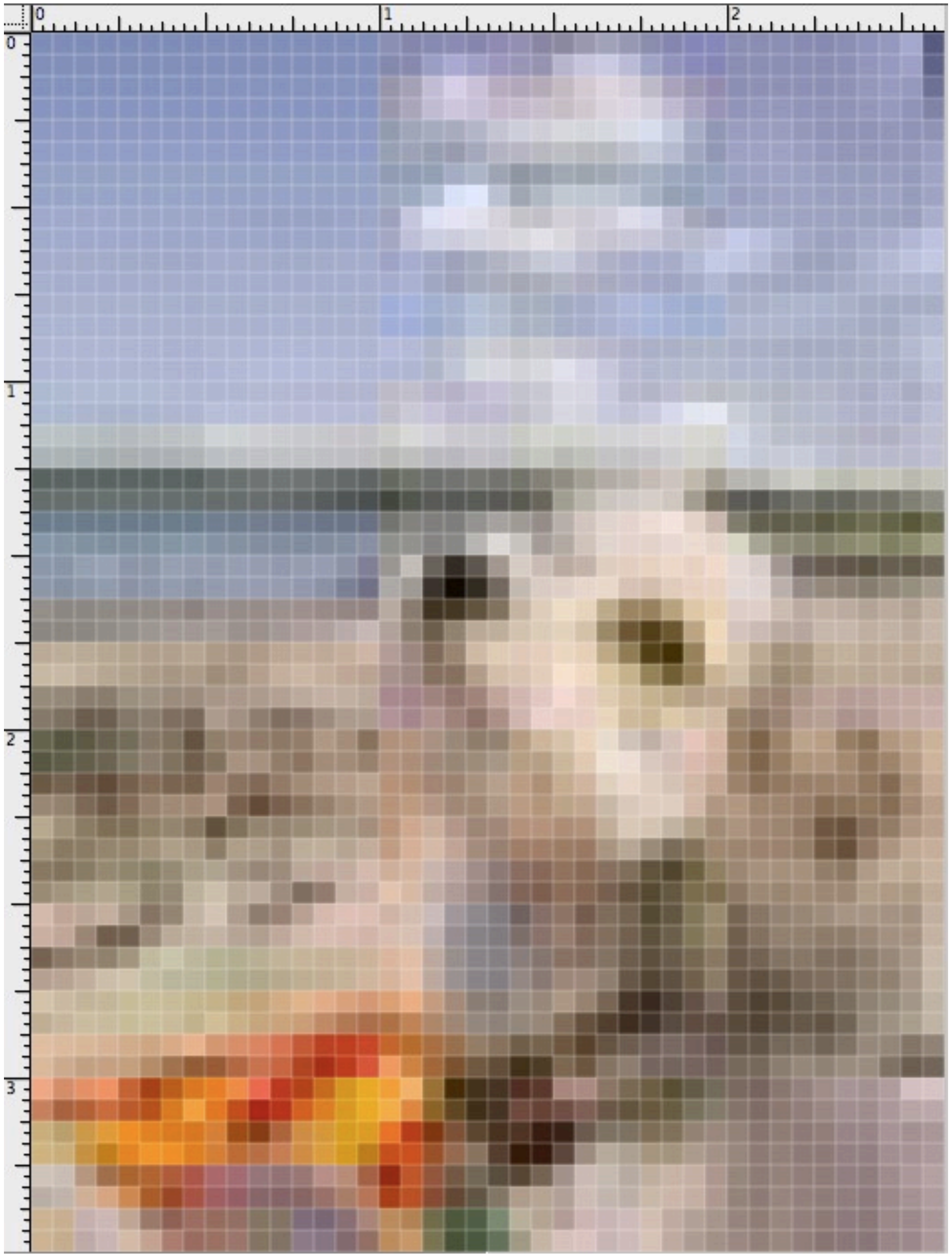
Digital Audio: Digitizing an Analog Signal

Computer Science

Compare with digitizing a scene:
Sample each pixel and digitize it using 8 bits for each of the primary colors.

Digital Audio: Digitizing an Analog Signal

Computer Science



Basic idea: Sample each pixel and digitize it using 8 bits for each of primary colors.

The same idea can be used for audio information:

Digitize the signal by

- sampling: periodically measure the analog amplitude of the signal
- quantizing: **round** the analog measurement to a discrete value

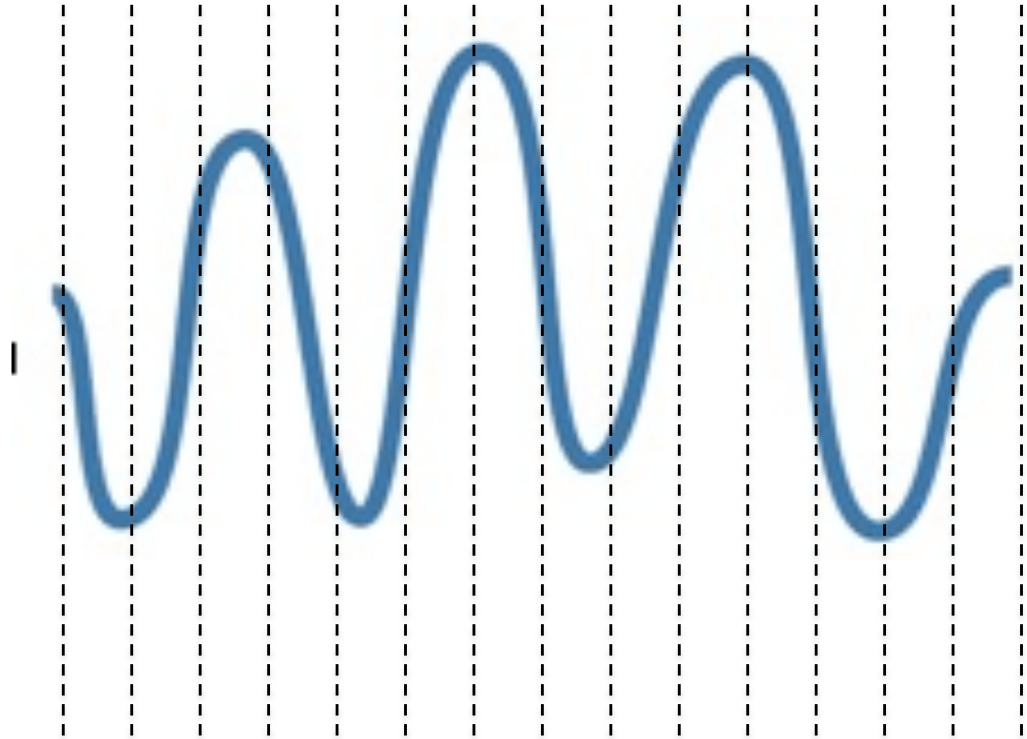
Digital Audio: Digitizing an Analog Signal

Computer Science

Sampling:

Decide on a sample rate (how many times a second to measure the wave);

Example: CD quality sound is measured 44,100 times a second= 44.1 kHz.



Digital Audio: Digitizing an Analog Signal

Computer Science

Quantizing:

Decide on a "bit depth" (how many bits for each measurement);

Example: CD quality sound is 16 bits, giving 2^{16} = 65,536 possible values, from -32768 to 32767

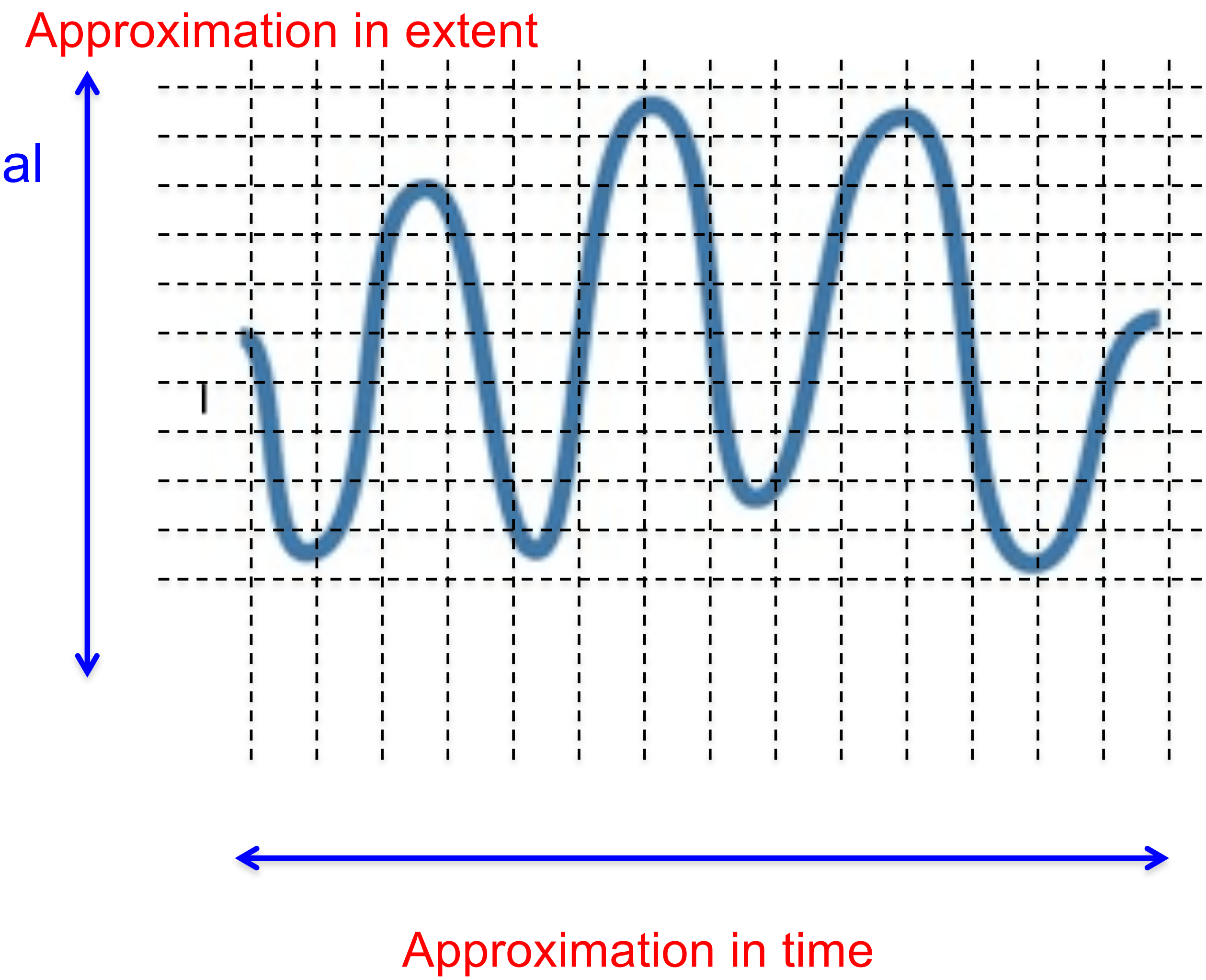
Measure the signal at each sample point, and round to the best approximate value for that bit depth.

Digital Audio: Digitizing an Analog Signal

Computer Science

Result of sampling and quantizing is that the signal is approximated in BOTH time and amplitude.

These two kinds of approximations have different properties and affect the quality of the resulting digital signal in different ways….



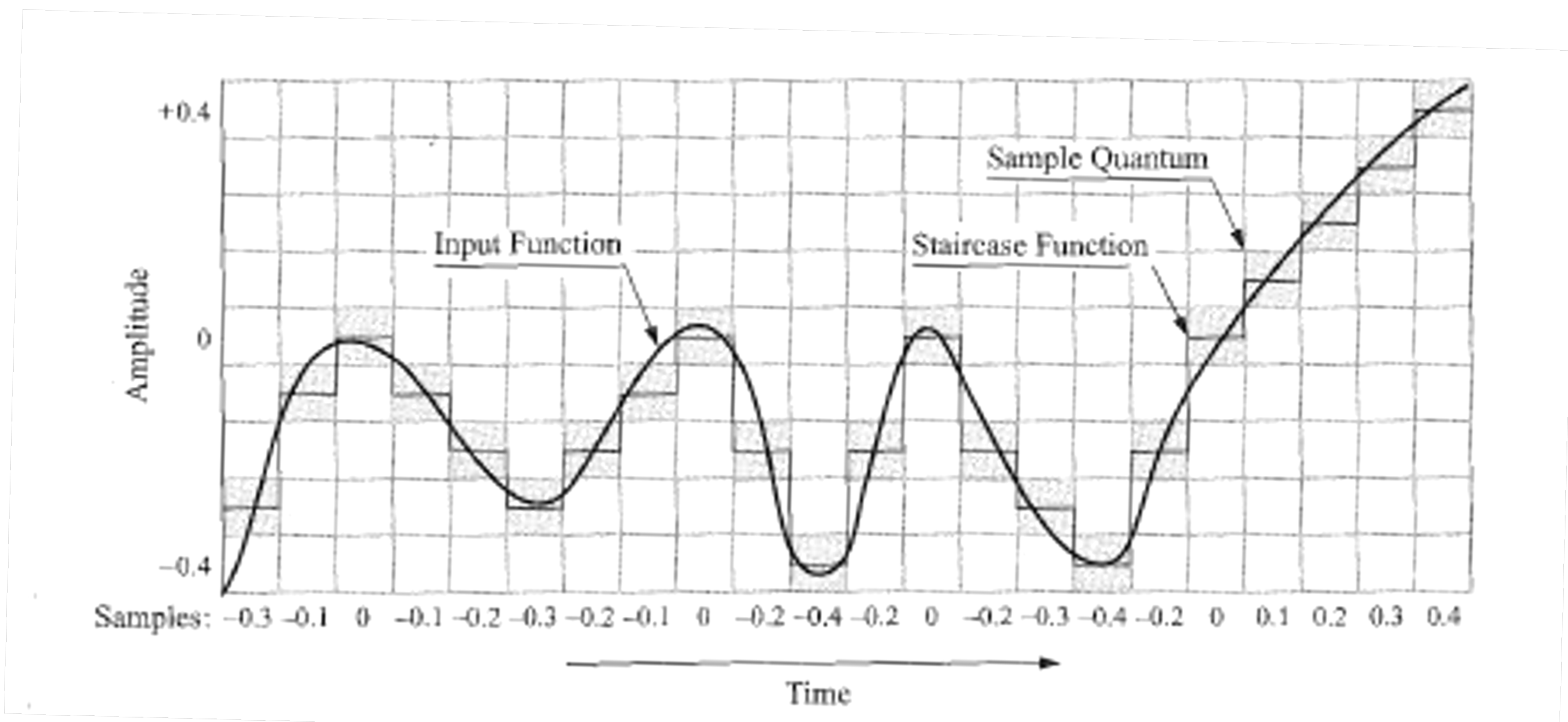
Digital Audio: Digitizing an Analog Signal

Computer Science

So, the two MOST IMPORTANT issues to understand are:

➢How often should we sample (the sample rate) and what are the consequences of this choice?

➢How many bits do we use for the sample type used to record the measurements, and what are the consequences of this choice?

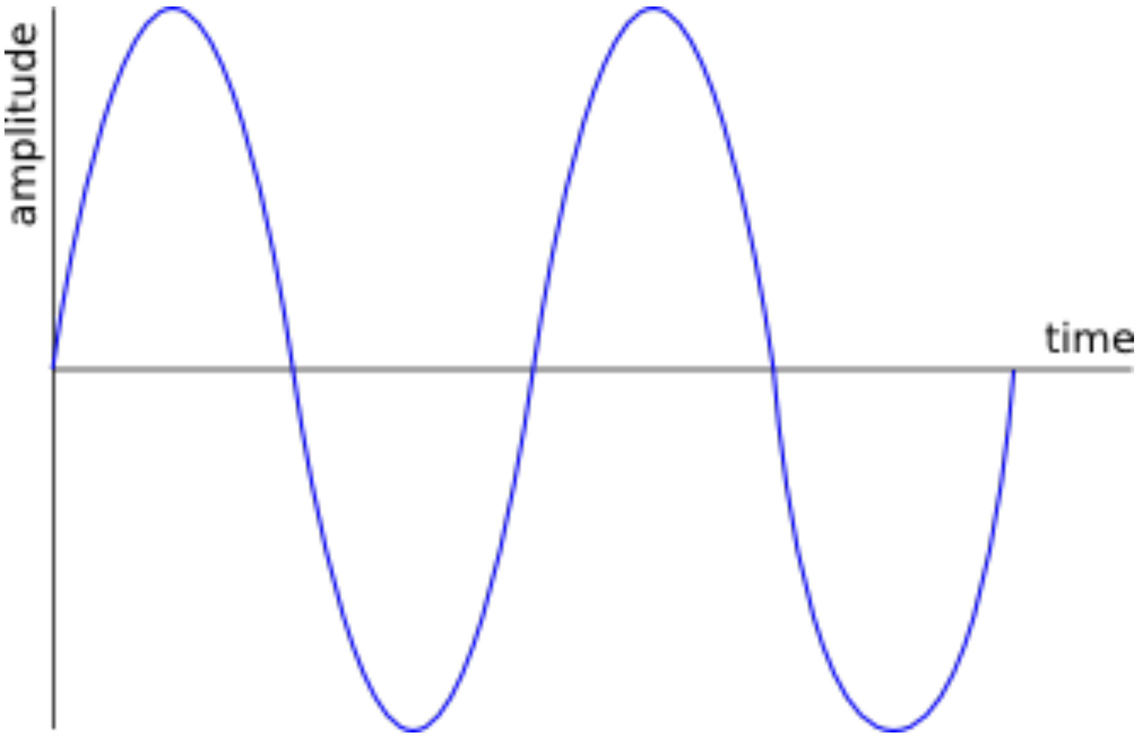


Digital Audio: Choice of sample rate

Computer Science

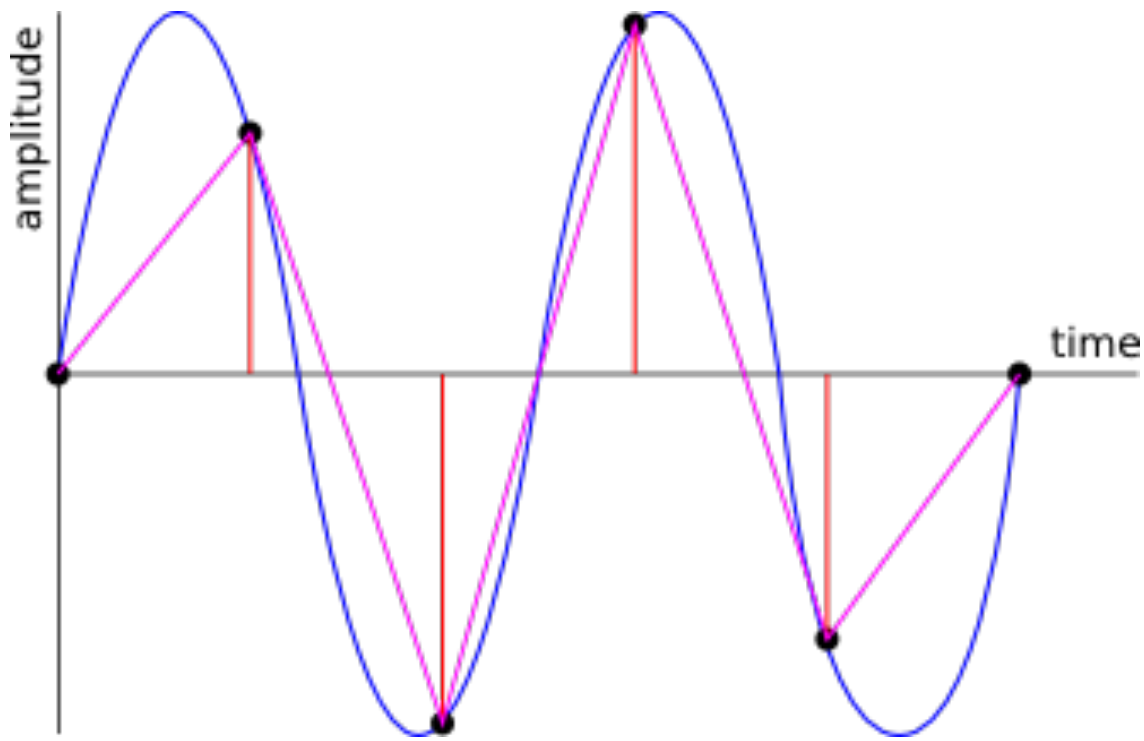
Consider this waveform.

What sampling rate should we choose?



Digital Audio: Choice of sample rate

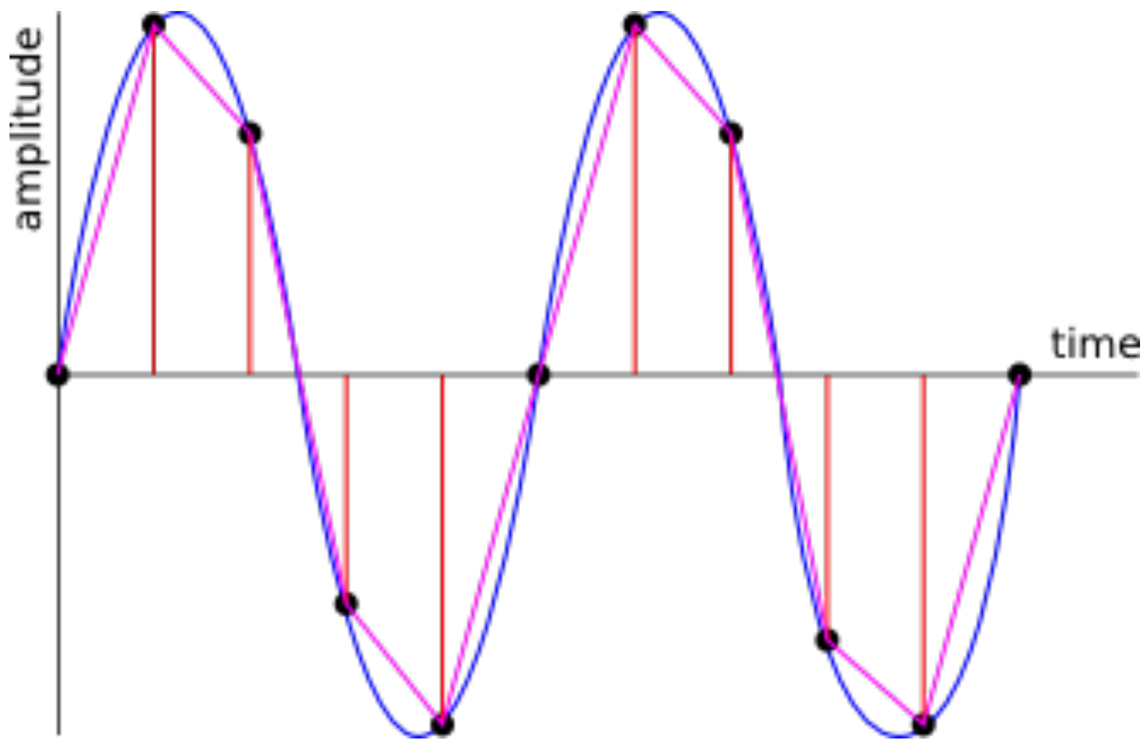
How about this sampling rate? (6 samples)



Digital Audio: Choice of sample rate

Computer Science

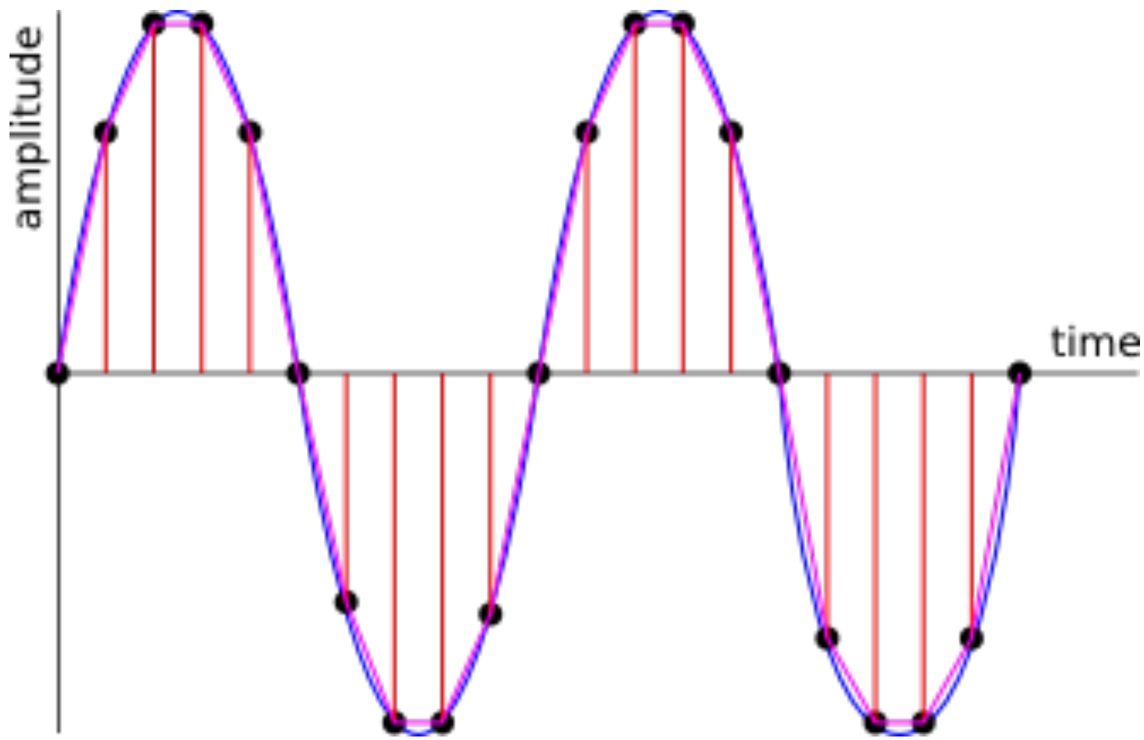
How about this sampling rate? (11 samples)



Digital Audio: Choice of sample rate

Computer Science

How about this sampling rate? (21 samples)



Digital Audio: Choice of sample rate

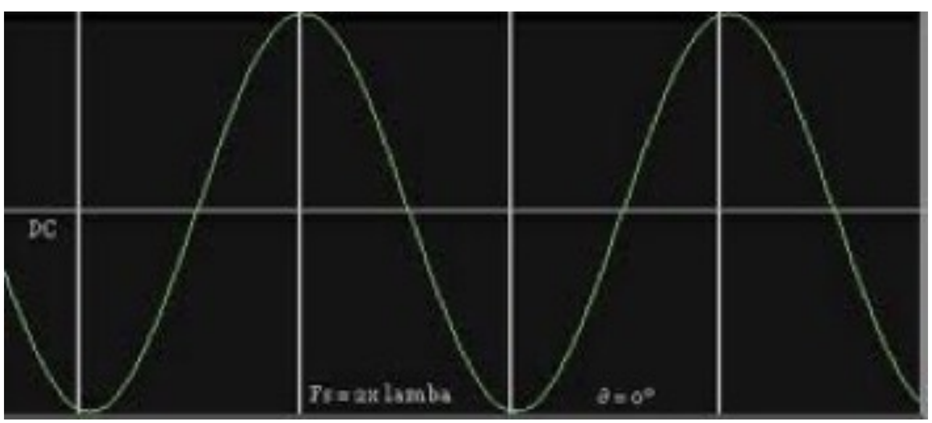
Computer Science

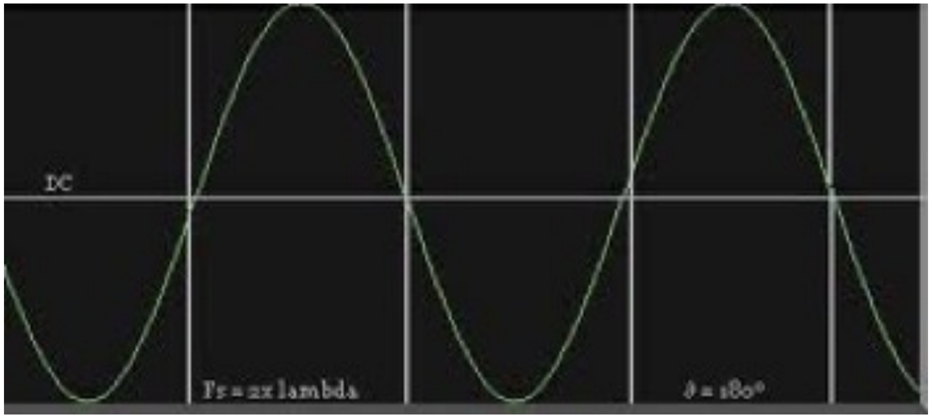
Consider this waveform, and these two sampling strategies.

What's going on here?

Samples: 100 -100 100 .... 0 0 0 0

Can reconstruct the wave.... ?????

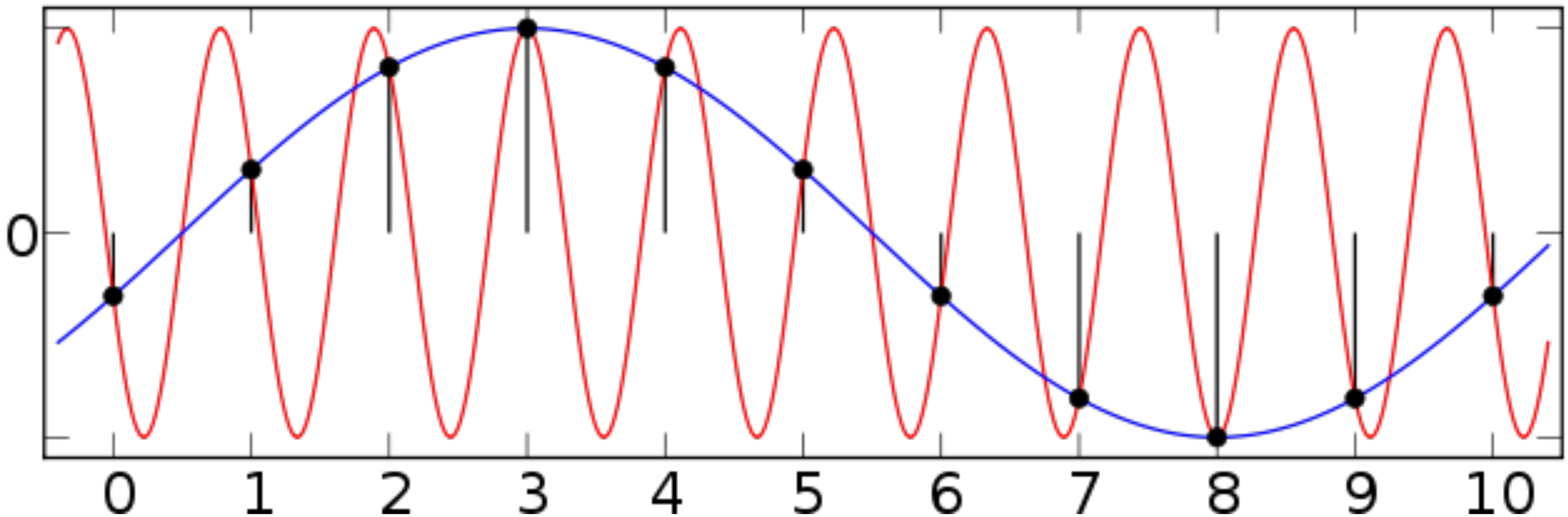




Digital Audio: Choice of sample rate

Boston University Computer Science

Even worse, we might get a completely different signal….



We want to reconstruct the red sine wave, but the simplest explanation is that these samples give us the blue wave!

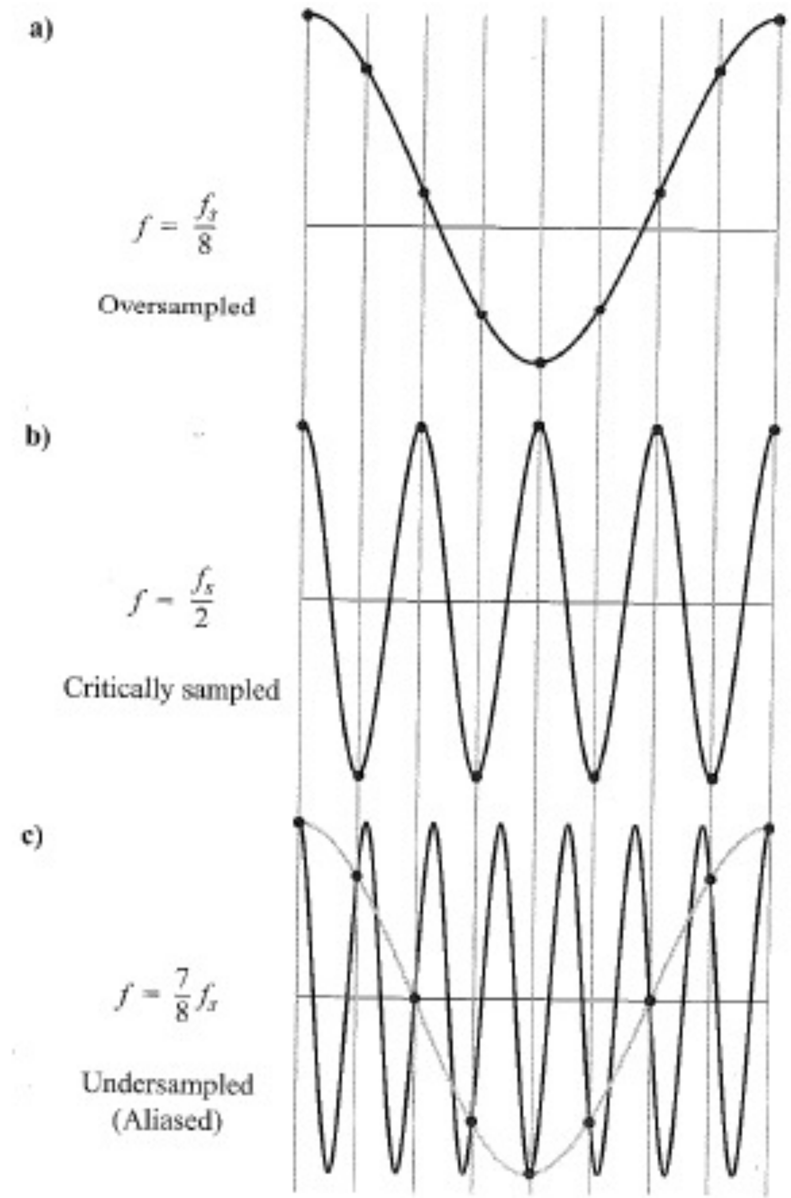
Digital Audio: Choice of sample rate

Computer Science

What sample rate should we use?

Naturally, higher sample rates approximate the analog signal more precisely, but this is a tradeoff with storage space; when do we have “enough” accuracy?

The principal result we need to keep in mind is the Nyquist Sampling Theorem, and the main artifact that results from (bad) choices of sample rate is Aliasing.....



Digital Audio Fundamentals

Computer Science

Nyquist Sampling Theorem

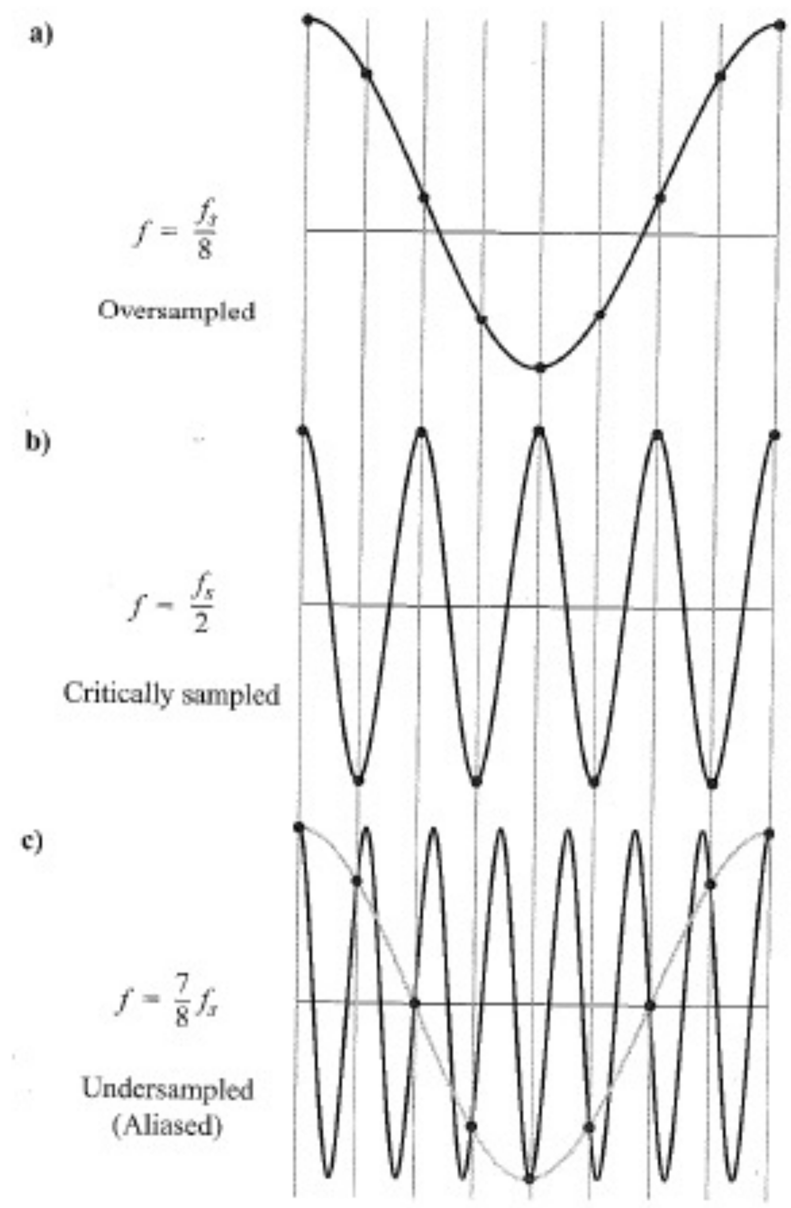
In order to precisely determine the shape of an (analog) sine wave, we must sample it more than twice in its period.

Nyquist Frequency or **Nyquist Limit** =

sample rate / 2

Thus, to accurately determine a sine wave of frequency f, we must sample with frequency > 2*f.

Punchline: Higher sample rates can represent higher frequencies. Lower sample rates limit the upper range of frequencies.. but the NST gives us a precise bound!



Digital Audio Fundamentals

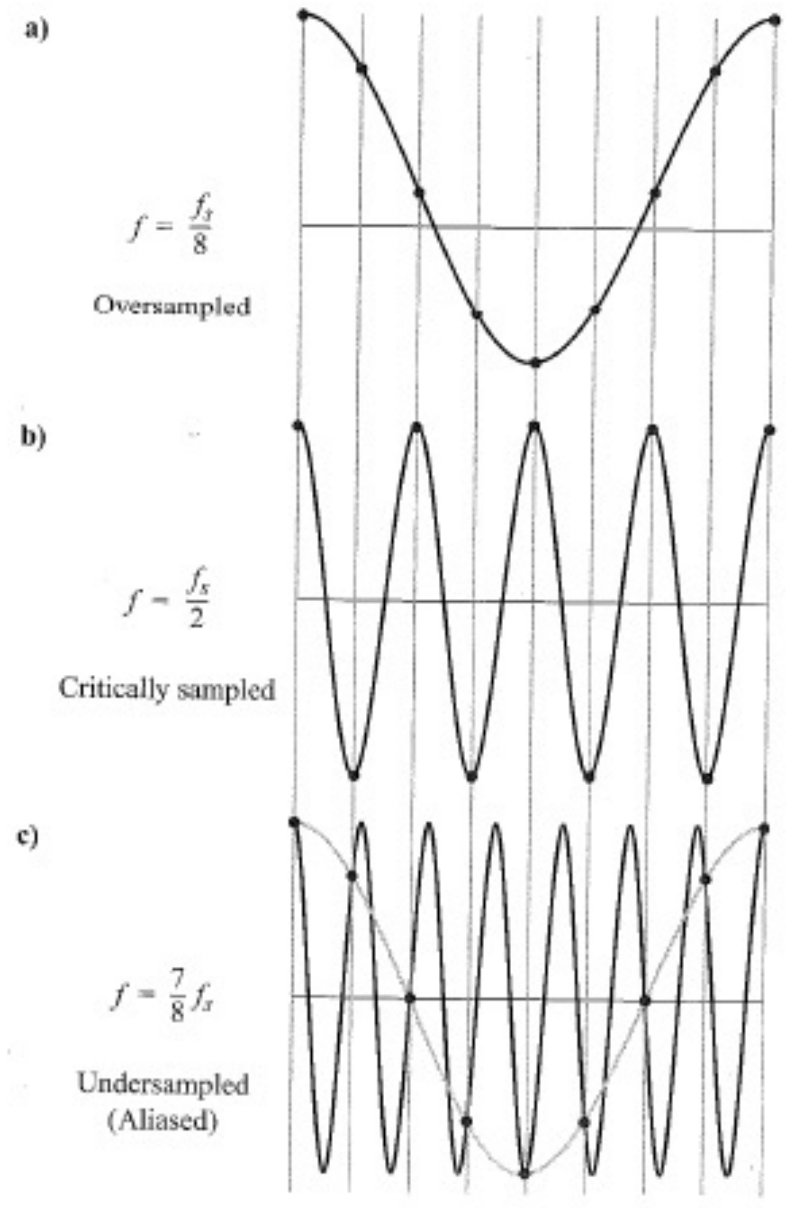
Computer Science

What happens if we reach or exceed the Nyquist Limit?

A wave of frequency above the Nyquist Frequency will be "undersampled" and appear to be a wave of lower frequency.

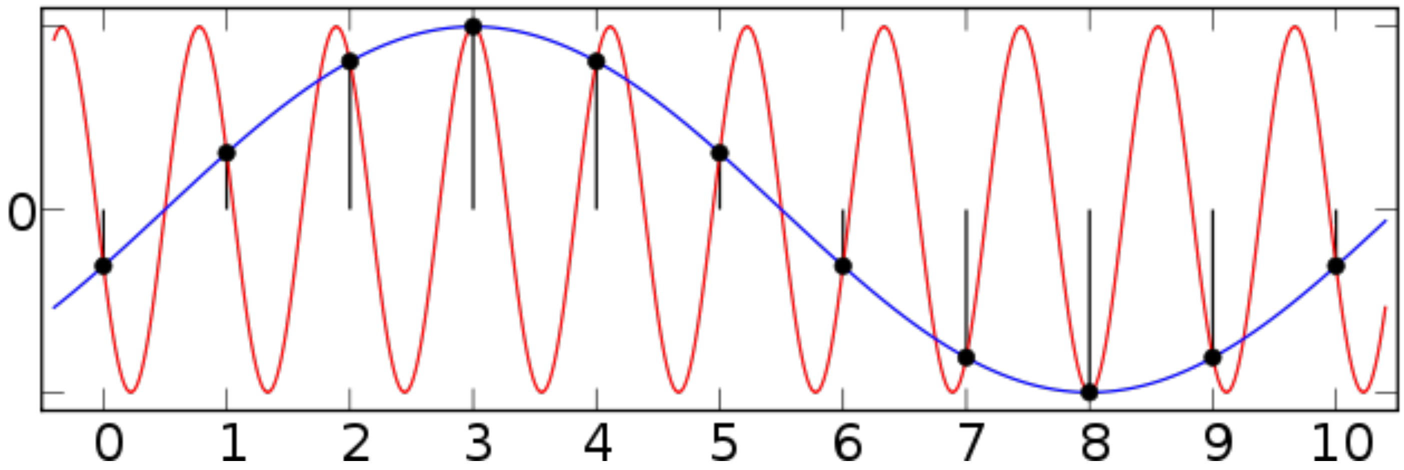
This is called **Aliasing**-- the lower (imaginary) frequencies are called **Aliases** of the higher frequencies.

The frequencies inside the Nyquist Limit are called the **Baseband**.



Digital Audio Fundamentals

Aliasing in Detail...... Suppose we sample at 1 Hz (one sample per second); then a sine wave of frequency of 0.9 Hz (red) and a wave of frequency 0.1 Hz (blue) both fit the samples shown below.

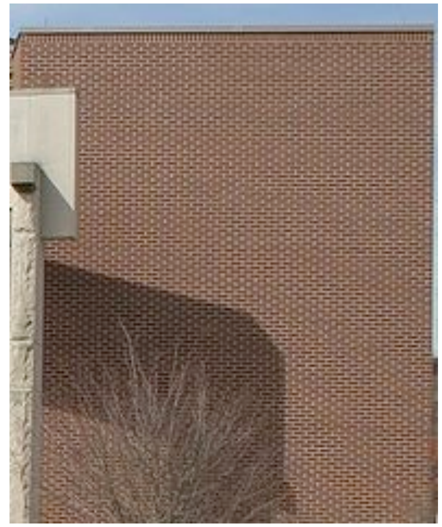


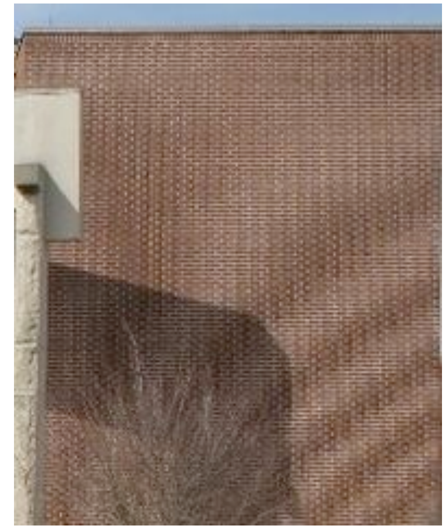
The blue wave is in the baseband (since 0.1 < 0.5 = Nyquist Frequency for 1 Hz) and is being accurately sampled; but if the red wave is the actual frequency, then it is outside the Nyquist Limit (0.9 > 0.5) and it will be HEARD AS THE BLUE WAVE ALIAS.

Digital Audio Fundamentals

Computer Science

DIGRESSION: Aliasing is a common issue in graphics......





But the most dramatic examples occur in situations analogous to ours, where samples (or images) are taken of periodic movement (e.g., rotating wheels or propellers) above the Nyquist Limit:

http://www.youtube.com/watch?v=jHS9JGkEOmA

http://www.youtube.com/watch?v=eTW0rNgMcKk

http://www.youtube.com/watch?v=LVwmtwZLG88&feature=fvwrel

Digital Audio Fundamentals

Computer Science

This causes horrible artifacts in audio files:

Sound example in Wikipedia article on Aliasing

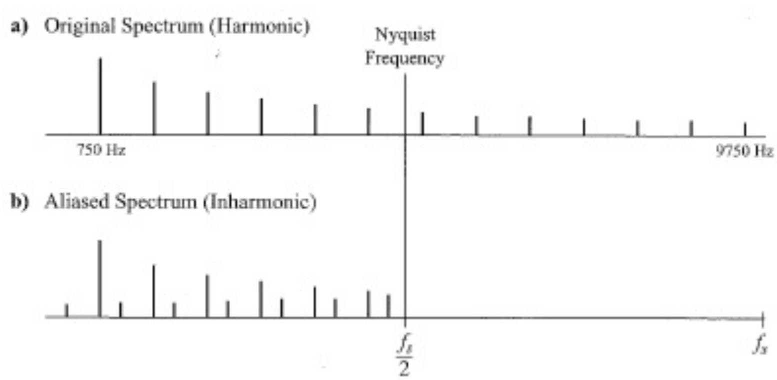


Figure 1.15
Aliased spectrum of harmonic tone.

The Solution: Simply apply a low-pass filter to cut any frequencies at or above the Nyquist Limit before sampling!

Digital Audio Fundamentals

Computer Science

Punchline on Sample Rate:

We must sample at more than twice the rate of the highest frequency we wish to represent; we must filter out frequencies higher than this maximum frequency before sampling, to avoid aliases.

Human hearing range: 20 – 20,000 Hz

Typical Sample Rate for WAV files: 44,100 Hz

A coincidence? I think not!

A lower sample rate will reduce the range of frequencies at the high end.... Let's listen to Higher Love by Steve Winwood at various sample rates....... 44kHz, 22kHz, 11kHz, and 6 kHz.

Sample Type

Second issue to consider: What are the consequences of various sample types? Common choices are:

16-bit integers (used in commercial CDs)
32-bit integers
32-bit floats

Digital Audio Fundamentals

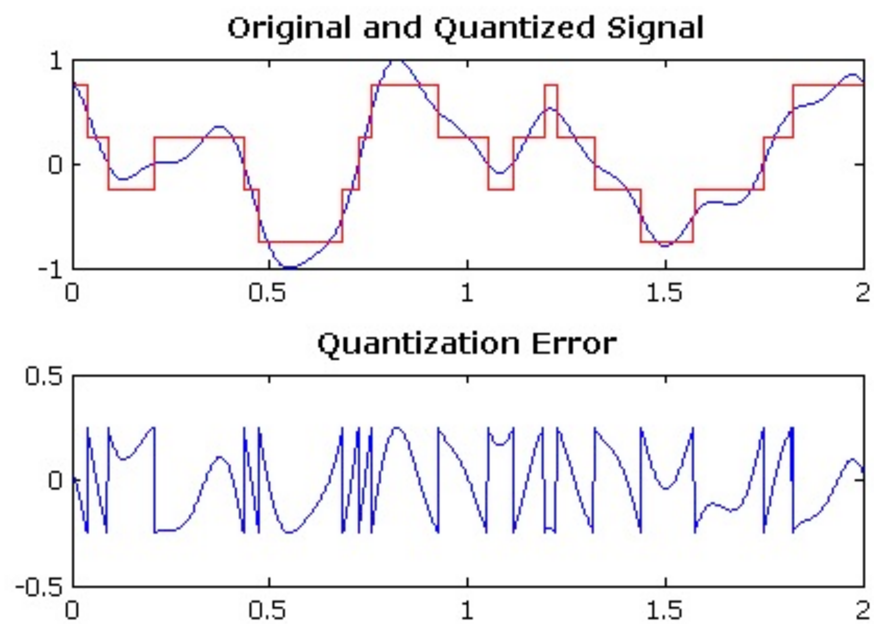
Computer Science

There are TWO consequences of the choice of sample type.

The first is that quantization error always occurs (the difference between the actual signal and the sample value), and this error is perceived as a separate signal ADDED to the original signal. Since it is basically randome, it is heard as NOISE.

The second is that samples with less information content allow for less dynamic range (the contrast between the softest and the loudest sounds): there are simply less gradiations in the amount of amplitude or loudness.

Let's listen to Higher Love again, this time changing the sample type from floats to 32-bit integers, to 16 to 8......



Digital Audio Fundamentals

Computer Science

But this is somewhat of overkill, at least for production quality CDs, which use 16-bit integers and a sample rate of 44.1 kHz.

The usual practice is to manipulate the signals at higher precision (e.g., using 64-bit FP) to handle all the information present optimally, then convert to 16 bit Integer for the final stage. 96 dB is about as much as can be perceived….

Audio Device/Application	Dynamic Range
AM Radio	48 dB
Analog Broadcast TV	60 dB
FM Radio	70 dB
Analog Cassette Player	73 dB
Video Camcorder	75 dB
ADI SoundPort Codecs	80 dB
16-bit Audio Converters	90 to 95 dB
Digital Broadcast TV	85 dB
Mini-Disk Player	90 dB
CD Player	92 to 96 dB
18-bit Audio Converters	104 dB
Digital Audio Tape (DAT)	110 dB
20-bit Audio Converters	110 dB
24-bit Audio Converters	110 to 120 dB
Analog Microphone	120 dB

Sound Environment	Sound Pressure Level (dBA SPL)	Approximate loudness with regard to ordinary conversation
Threshold of hearing	0	Don't hear anything
Broadcast studio interior or rustling leaves	10	1/32nd as loud as conversation
Quiet house interior or rural nighttime	20	1/16th as loud
Quiet office interior or watch ticking	30	1/8th as loud
Quiet rural area or small theater	40	1/4th as loud
Quiet suburban area or dishwasher in next room	50	1/2 as loud
Office interior or ordinary conversation	60	Ordinary Conversation
Vacuum cleaner at 10 ft.	70	Twice as loud
Passing car at 10 ft. or garbage disposal at 3 ft	80	4 times as loud
Passing bus or truck at 10 ft. or food blender at 3 ft.	90	8 times as loud
Passing subway train at 10 ft. or gas lawn mower at 3 ft.	100	16 times as loud
Night club with band playing	110	32 times as loud
Threshold of pain	120	64 times as loud as conversation (twice as loud as night club)

Example: I just bought a “remastering” of some older recordings, and they did the work at 48 bits and 96 kHz.