MUSC 208 Winter 2014
John Ellinger
Carleton College

# Six Basic Properties of Sound

Sound waves create pressure differences in the air. These pressure differences are analogous to ripples that appear when a small stone is thrown in water. The troughs and valleys of the wave spread out from the center in all directions. A slice through the waves into the center reveals sinusoidal motion and its associated amplitude and frequency. Our eardrum is very sensitive to pressure differences in the sound wave. Through a complex process in the inner ear and neural pathways, we hear the sound. The range of human hearing is approximately 16 Hz to 16000 Hz. As we get older the upper range of hearing diminishes.

Sounds we hear have six basic properties.

1. Frequency
2. Amplitude
3. Timbre
4. Duration
5. Envelope
6. Location

These six properties of sound are studied in the fields of music, physics, acoustics, digital signal processing (DSP), computer science, electrical engineering, psychology, and biology. This course will examine these properties from the perspective of music, MIDI, and digital audio.

## Frequency

Frequency is measured in Hertz (Hz). One Hz is one cycle per second. Human hearing lies within the range of 20Hz - 20,000Hz. As we get older the upper range of our hearing diminishes. Human speech generally falls in the range from 85 Hz - 1100 Hz. Two frequencies are an octave apart if one of them is exactly twice the frequency of the other. These frequencies are increasing by one octave: 100Hz, 200Hz, 400Hz, 800Hz, and 1600Hz.

The higher the frequency, the higher the pitch. The lower the frequency, the lower the pitch. High pitches are written higher on the musical staff than low pitches. High notes on the piano are on the right, and low notes are on the left as you're facing the piano.

| Low frequency = low pitch | High frequency = high pitch |
|---|---|
|  |  |
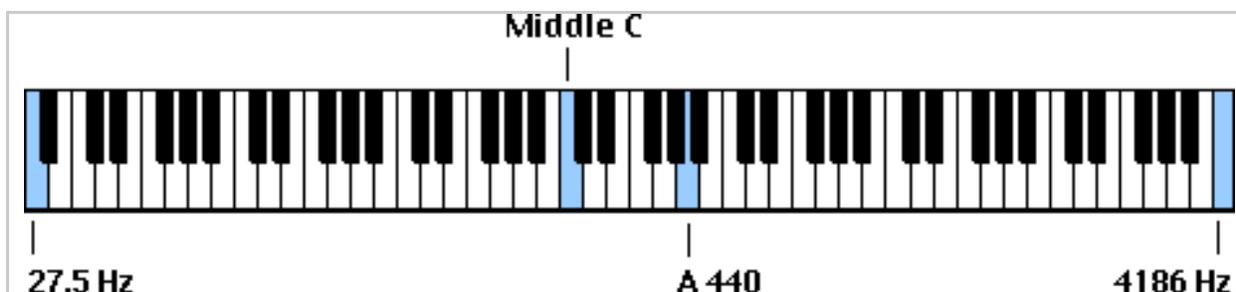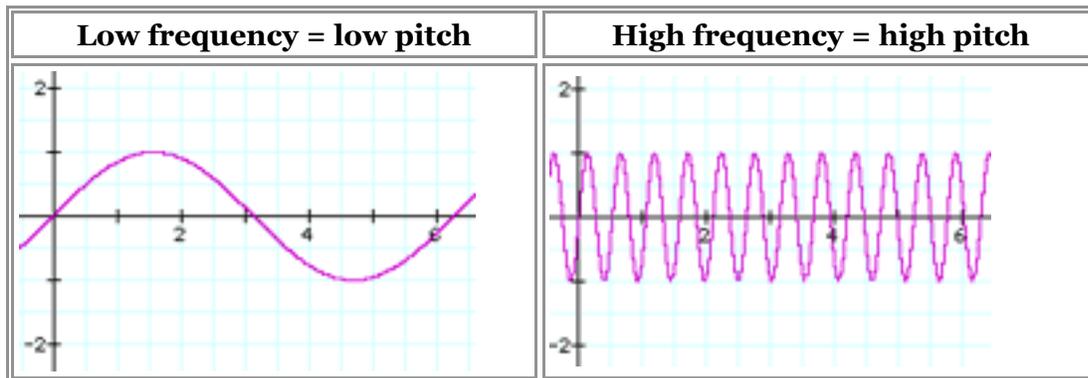
Frequency and pitch are often thought to be synonymous. However, there is a subtle distinction between them. A pure sine wave is the only sound that consists of one and only one frequency. Most musical sounds we hear contain a mix of several harmonically related frequencies. Nonetheless, musical frequencies are referred to as pitch and are given a names like "Middle C" or "the A above Middle C".

## How Long Is A Sound Wave?

You can calculate the distance between crests of a sound wave if you know its frequency. The speed of sound approximately is 791 mph or 1160.1 feet/second. In one second the note A440 goes through 440 cycles and has traveled 1160 feet. Therefore the distance between crests from one cycle to the next is 1160.1 ft/sec ÷ 440 cycle/sec = 2.64 feet/cycle. The distance between crests for the lowest (27.5 Hz) and highest (4186 Hz) notes on the piano are 42 feet and 3.3 inches respectively.

## Frequency In MIDI

MIDI represents frequency as a MIDI note number corresponding to notes on the piano. The 88 keys on the piano are assigned MIDI note numbers 21-108, with middle C as note number 60. You can move up or down an octave in MIDI by adding or subtracting 12 from the MIDI note number. MIDI is capable of altering the frequency of a note using using the "pitch bend" command. A MIDI Tuning Extension permits microtonal tuning for each of the 128 MIDI note numbers. However, not all synthesizers support micro tuning.

## Frequency In Digital Audio

Physics and digital signal processing (DSP) deal with very large frequency ranges. The audio spectrum is a very small part of that range, a relatively narrow band between 20 - 20,000 Hz. The Audio band spans 20000 ($2 \times 10^4$) Hz out of a total bandwidth of $10^{20}$ Hz. Because of the extreme differences between audio waves and gamma waves, a logarithmic scale was used for the graph.
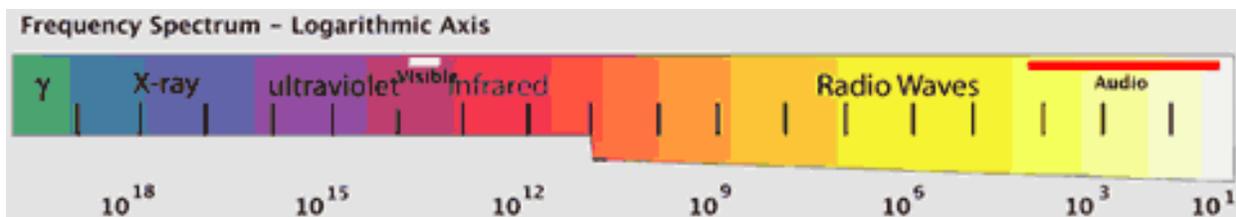
*https://upload.wikimedia.org/wikipedia/commons/7/72/Radio_transmition_diagram_en.png, modified JE*

The frequency of a sound wave is measure in Herz, Hz. One Hz is one cycle per second. The more cycles per second the higher the frequency.

| Low frequency = low pitch | High frequency = high pitch |
|---|---|
|  |  |
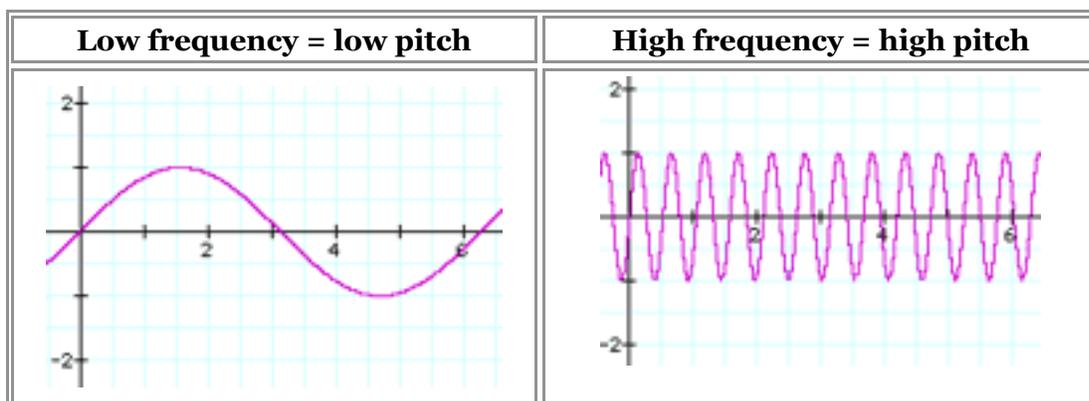
The period of a sound wave is the time it takes to complete one cycle, measured in seconds. Frequency and period are reciprocals of each other.

$$frequency = \frac{1}{period} \quad \text{and} \quad period = \frac{1}{frequency}$$

The "purest" pitched tone is a sine wave. The formula for a sampled sine wave is shown below where n is a positive integer sample number, f is the frequency in Hz, SR is the sampling rate in samples per second, and θ is the phase in radians.

$$Sinewave(n) = Sin(2\pi f \frac{n}{SR} + \theta)$$

Many complex tones can be generated by adding, subtracting, and multiplying sine waves. Any frequency is possible and any tuning system is possible.

**Pitch, Frequency and Tuning**

Pitch and frequency are often thought to be synonymous. However, there is a subtle distinction between them. Frequency refers to all wavelengths in the electromagnetic spectrum. Pitch is only used to describe frequencies we can hear. A pure sine wave is the only musical tone that consists of exactly one frequency. All other musical tones are a mixture of several frequencies. These frequencies are related by integer multiples of a fundamental frequency forming a pattern

called the harmonic series.

The frequency spectrum used in music is a discrete system, where only a select number of specific frequencies are used. For example, the piano uses 88 of them from 27.5 Hz to 4186 Hz.

 The tuning reference for all modern musical instruments is based on a frequency commonly called A-440. It's the note A above middle C on the piano with a frequency of 440 Hz.

**Tuning and Equal Temperament**

The modern system of tuning is called Equal Temperament and divides the octave 12 equal parts. Any two frequencies that are an octave apart have a frequency ratio of 2:1. The frequency ratio between each of the twelve neighboring notes within an octave is $2^{\frac{1}{12}} = 1.0595$.

The frequency of the note A# one half step above A-440 is $440 * 2^{\frac{1}{12}} = 466.16$ Hz.

The frequency of the note G two half steps below A-440 is $440 * 2^{\frac{-2}{12}} = 392$ Hz.

In general the frequency of any note N half steps away from A-440 is $440 * 2^{\frac{\pm N}{12}}$ Hz.

**In Tune and Beats**

Two notes are in tune when the have the same frequency. If two notes are out of tune but close in frequency it's difficult to tell which one is higher. A musician tuning one note to another does not listen to pitch, she listens for beats. When two notes with slightly different frequencies are played together the peaks and valleys of the individual wave forms add together, sometimes amplifying and sometimes attenuating each other. These changes in amplitude intensity can be heard and are called beats. The rate of beating is equal to the difference between the two frequencies.

# Amplitude

The amplitude of a sound is a measure of it's power and is measured in decibels. We perceive amplitude as loud and soft. Studies in hearing show that we perceive sounds at very low and very high frequencies as being softer than sounds in the middle frequencies, even though they have the same amplitude. Amplitude refers to the displacement of the waveform above or below the x axis.
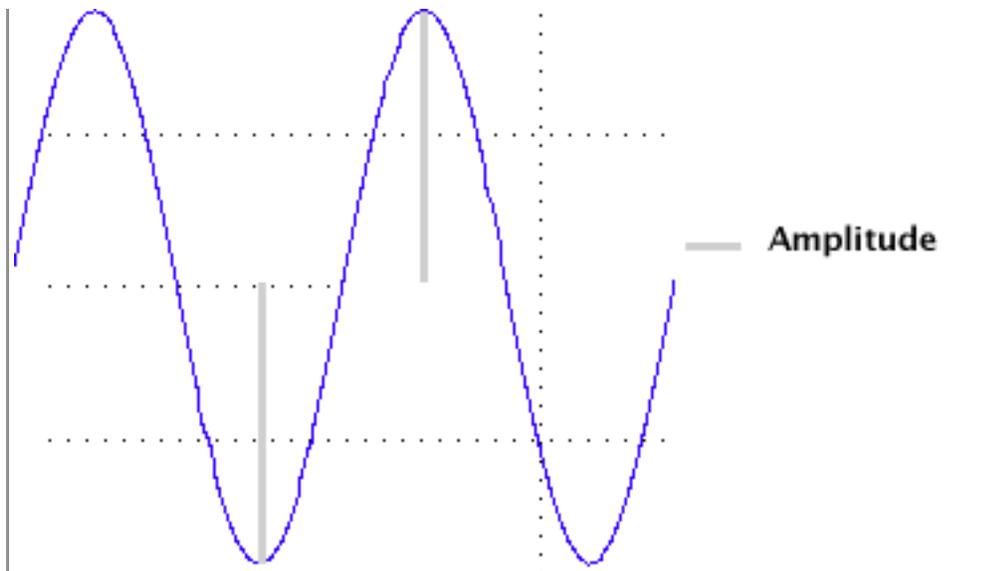
## Amplitude In Music

The musical term for amplitude is dynamics. There are nine common music notation symbols used used to represent dynamics. From extremely loud to silence they are:

| Symbol | Name | Performed |
|:---:|:---:|:---:|
| *fff* | fortississimo | silence<br>as loud as possible |
| *ff* | fortissimo | very loud |
| *f* | forte | loud |
| *mf* | mezzo forte | medium loud |
| *mp* | mezzo piano | medium soft |
| *p* | piano | soft |
| *pp* | pianissimo | very soft |
| *ppp* | pianississimo | as soft as possible |

The pianoforte is the ancestor of the modern piano and was invented in Italy around 1710 by Bartolomeo Cristofori. It was the first keyboard instrument that could play soft (piano) or loud (forte) depending on the force applied to the keys, thus its name. If it had been invented in England we might know it as the softloud.

### Amplitude In MIDI

The MIDI term for amplitude is velocity. MIDI velocity numbers range from 0-127. Higher velocities are louder.

### Amplitude In Digital Audio

The amplitude of a sound wave determines its relative loudness. Looking at a graph of a sound wave, the amplitude is the height of the wave. These two sound waves have the same frequency but differ in amplitude. The one with the higher amplitude sounds the loudest.

| Low amplitude = soft sound | High amplitude = loud sound |
|:---:|:---:|
|  |  |
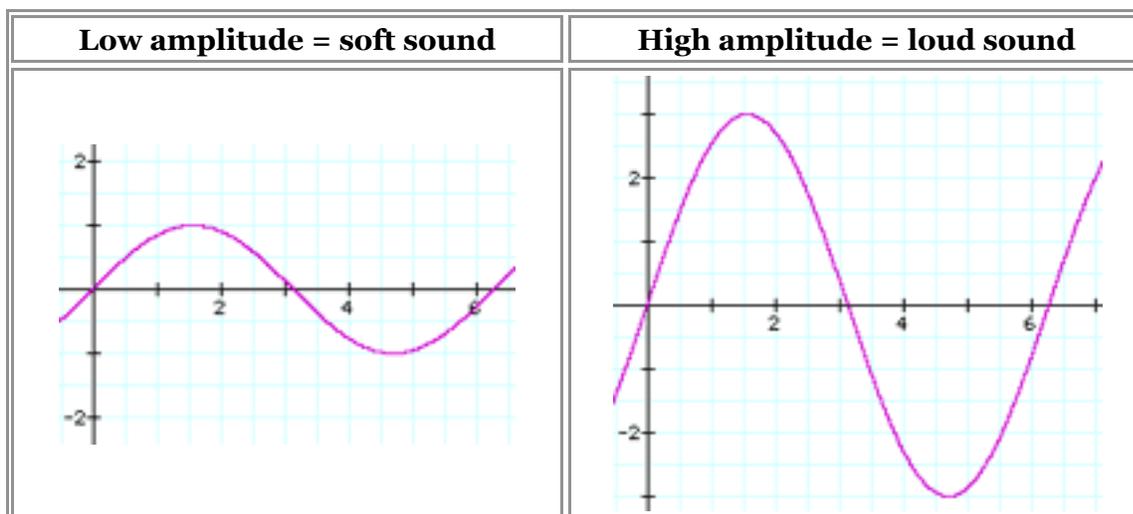
Amplitude is measured in decibels. Decibels have no physical units, they are pure numbers that express a ratio of how much louder or softer one sound is to another. Because our ears are so sensitive to a huge range of sound, decibels use a logarithmic rather than a linear scale according to this formula.

One of the Amplitudes is used as a reference level to which all other amplitudes are compared. The reference level usually chosen represents the softest sound that can just barely be heard by a person with good hearing. Because either Amplitude 1 or Amplitude 2 can be chosen as the reference level decibel scales can run in two directions.

When the amplitude is changed by a factor of ten the decibel change is 20 dB. Every 10 decibel change represents a power of ten increase in sound intensity. For example, the intensity difference between the softest symphonic music (20 dB) and loudest symphonic music (100 db) differs by a factor of 100,000,000 (10 to the 8th power). This table shows some relative decibel

6

levels. Read 10^3 as $\mathbf{10^3}$.

| Decibels | Magnitude | Description |
|---|---|---|
| Logarithmic Scale | Linear Scale | |
| 160 | 10^16 | |
| 150 | 10^15 | |
| 140 | 10^14 | Jet takeoff |
| 130 | 10^13 | |
| 120 | 10^12 | Threshold of pain <br> Amplified rock band |
| 110 | 10^11 | |
| 100 | 10^10 | Loudest symphonic music |
| 90 | 10^9 | |
| 80 | 10^8 | Vacuum Cleaner |
| 70 | 10^7 | |
| 60 | 10^6 | Conversation |
| 50 | 10^5 | |
| 40 | 10^4 | |
| 30 | 10^3 | |
| 20 | 10^2 | Whispering <br> Softest symphonic music |
| 10 | 10^1 | |
| 0 | 10^0 | Threshold of hearing |

Digital audio often reverses the decibel scale making 0 dB the loudest sound that can be accurately produced by the hardware without distortion. Softer sounds are measured as negative decibels below zero. Software decibel scales often use a portion of the 0 dB to 120 dB range and may choose an arbitrary value for the 0 dB point. This dB scale is found in the Logic Pro software. The dB scale on the left goes from 0 dB down to -60 dB. The 0.0 dB setting on the volume fader on the right corresponds to -11 dB on the dB scale.
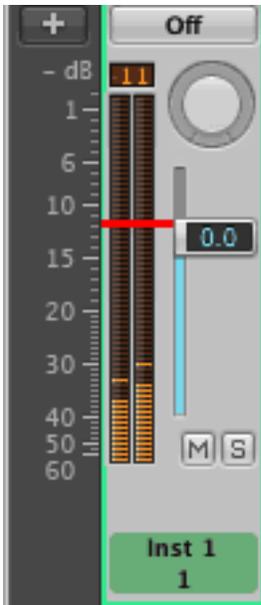
When recording digital audio, you never want the sound to stay under well under 0 dB. Notice the flat tops on the signal on the left at the 0 dB mark. That's referred to as digital clipping and it sounds terrible. The screen shots were captured using the open source cross platform software Audacity. (http://audacity.sourceforge.net/)
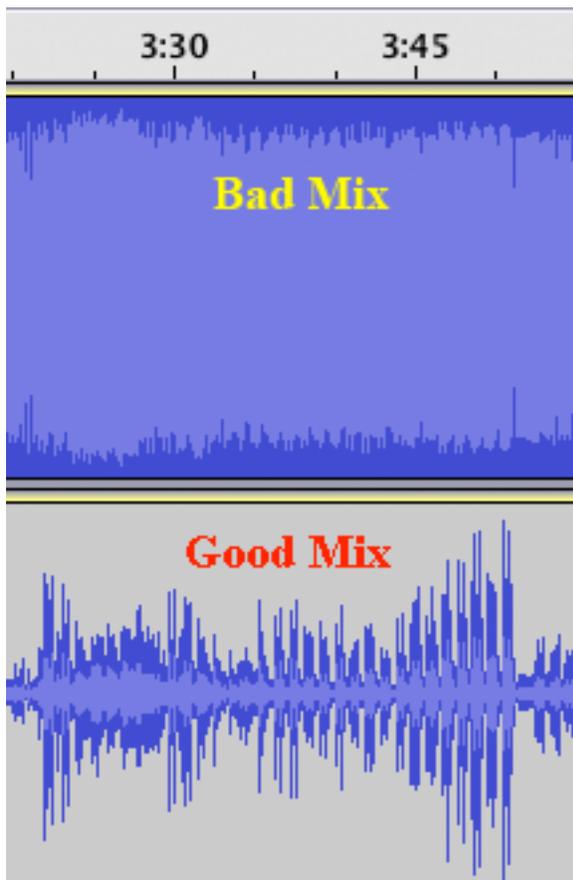
Decibels can be confusing because of the difference between amplitude and power. Decibel measurements are described as Sound Power Level, Sound Pressure Level, and Sound Intensity Level. When converting between amplitude and decibels in digital audio where the amplitude values are real numbers between 0 and 1, these two formulas are used:

$$dB = 20\,Log_{10}(amplitude) \quad \text{and} \quad amplitude = 10^{\frac{dB}{20}}$$

## Duration

Sound exists in time. Without the repeating patterns of compression and rarifaction of air molecules we'd never hear a sound.

When we talk about duration we're talking about time. We need to know two events related to the time of a sound, when did it start and how long did it last. In music and digital audio, time usually starts at zero. How time is tracked is usually a variation on chronological time or proportional time. Here are some examples.

When we talk about duration we're talking about the time between two events, when did it start and when did it end. In music and digital audio, time starts at zero. The starting and ending times of other events can be measured in absolute terms or relative terms. For example, an audio CD is sampled at 44,100 samples per second. We know that sample 22050 is exactly one half second after the sound started, absolute time. We also know that the duration between sample numbers 441 and 882 is 100 milliseconds, relative time.

### Duration In Music

Muscial time is relative. A metronome supplies a steady source of beats separated by equal units of time. One musical note value (often a quarter note) is chosen as the beat unit that corresponds to one click of the metronome. All other note values are a proportional to that. Metronome clicks can come at a slow, medium, or fast pace. The speed of the metronome clicks is measured in beats per minute and is called the tempo. Whether the tempo is slow or fast, the rhythmic proportions between the notes remains the same. The actual duration of each note expands or contracts in proportion to the tempo.

### Duration In MIDI

Time is not defined in the MIDI standard. The software uses the computer clock to keep track of time, often in milliseconds, sometimes in microseconds. Here's a recipe (pseudo code) to play two quarter notes at a tempo of 60.

1. Check the song tempo. It's 60 beats per minute so each quarter note lasts 1000 milliseconds.
2. Set the computer clock to 0.
3. Send a Note On message (NON) for the first note.
4. Keep checking the clock and wait until the clock reads 1000 milliseconds.
5. Send a Note Off message (NOF) to turn off the first note.
6. Send a NON message to turn on the second note.
7. Keep checking the clock and wait until the clock reads 2000 milliseconds.

8.    Send a NOF message to turn off the second note.

## Duration In Digital Audio

Duration in digital audio is a function of the sampling rate. Audio CD's are sampled at a rate of 44,100 samples per second with a bit depth of 16. Bit depth refers to the range of amplitude values. The largest number that can be expressed in sixteen bits is $2^{16}$ or 65,536. To put audio sampling in perspective, let's graph of one second of sound at the CD sampling rate and bit depth. Start with a very large piece of paper and draw tick marks along X axis, placing each tick exactly one millimeter apart. You'll need 44,100 tick marks which will extend about 44.1 meters (145 feet). Next you need place tick marks spaced one millimeter apart on the Y axis to represent the 16 bit amplitude values. Let's put half the ticks above the X axis and half below. You'll need 32.768 meters (about 107 feet) above and below the X axis. Now draw a squiggly wave shape above and below the X axis, from the origin to the end of the 44100th sample while staying within the Y axis boundaries. Next carefully measure the waveform height in millimeters above or below the X axis at every millimeter sample point point along the 145 foot X axis. Write these numbers down in a single file column. You've just sampled one second of a sound wave at the CD audio rate. It probably took longer than one second; definitely not real time sampling.

Digital sampling is done with specialized hardware called an Analog to Digital Converter (ADC). The ADC electronics contain a clock running at 44100 Hz. At every tick of the clock, the ADC reads the value of an electrical voltage at its input which is often a microphone. That voltage is stored as a 16 bit number with values from −32767 to +32767. Bit depths of 24 are also used in today's audio equipment. A bit depth of 24 can use 16,777,216 different values to represent the amplitude.

The counterpart to the ADC is the DAC (Digital to Analog Converter) that converts the sample numbers back into an analog signal that can be played through a speaker. Most modern computers have consumer quality ADC and DAC converters built in. Professional recording studios use external hardware ADC's and DAC's that often cost thousands of or tens of thousands of dollars.

# Timbre

Timbre (pronounced **TAM**-burr) refers to the tone color of a sound. It's what makes a piano sound different from a flute or violin. The timbre of a musical instrument is determined by its physical construction and shape.

## Timbre in Music

Timbre in music is specified as text in the score, like Sonata for flute, oboe and piano. Timbre changes may be specified in the score as special effects like growls, harmonics, slaps, scrapes, or by attaching mutes to the instrument.

## Timbre In MIDI

Timbre in MIDI is changed by pushing hardware buttons or sending a MIDI message called the

Patch change command.

## Timbre In Digital Audio

Sounds with different timbres have different wave shapes. However it's not the amplitude shape that determines the timbre it's the  relative strength of the individual frequency components present in the sound as it evolves over time. The method of obtaining a frequency spectrum is based on the mathematics of the Fourier transform. The Fourier theorem states that any periodic waveform can be reproduced as a sum of a series of sine waves at integer multiples of a fundamental frequency with well chosen amplitudes and phases. The pure mathematics sum would have an infinite number of terms. However, in digital audio a reasonable number works quite well.

$$anyWave = \sum_{n=1}^{\infty} a_n Sin(2\pi nf + \theta_n)$$

or

$$anyWave = a_1 Sin(2\pi 1f + \theta_1) + a_2 Sin(2\pi 2f + \theta_2) + a_3 Sin(2\pi 3f + \theta_3) + a_4 Sin(2\pi 4f + \theta_4)...$$

## The Fast Fourier Transform (FFT)

The FFT is arguably the single most important tool in the field of digital signal processing. It can convert a sampled waveform from the time domain waveform into the frequency domain and back again. By adjusting and manipulating individual components in the frequency domain it is possible to track pitch changes; create brand new sounds; create filters that modify the sound; morph one sound into another; stretch the time without changing the pitch; and change the pitch without affecting the time. Within the last 10 years, desktop computers have become fast enough to process these DSP effects in real time.

Here are the time domain and frequency domain waveforms of several different instruments playing the pitch A440. The magnitude, spacing, and frequencies in the frequency spectrum largely determine the timbre. FFT pictures from Audicity Plot Spectrum menu item.

| | Time Domain | Frequency Domain FFT |
|---|---|---|
| **Piano** |  |  |
| **Violin** |  |  |
| **Flute** |  |  |
| **Oboe** |  |  |
| **Trumpet** |  |  |

| | | |
|---|---|---|
| **Electric Guitar** | | |
| **Bell** | | |
| **Snare Drum** | | |
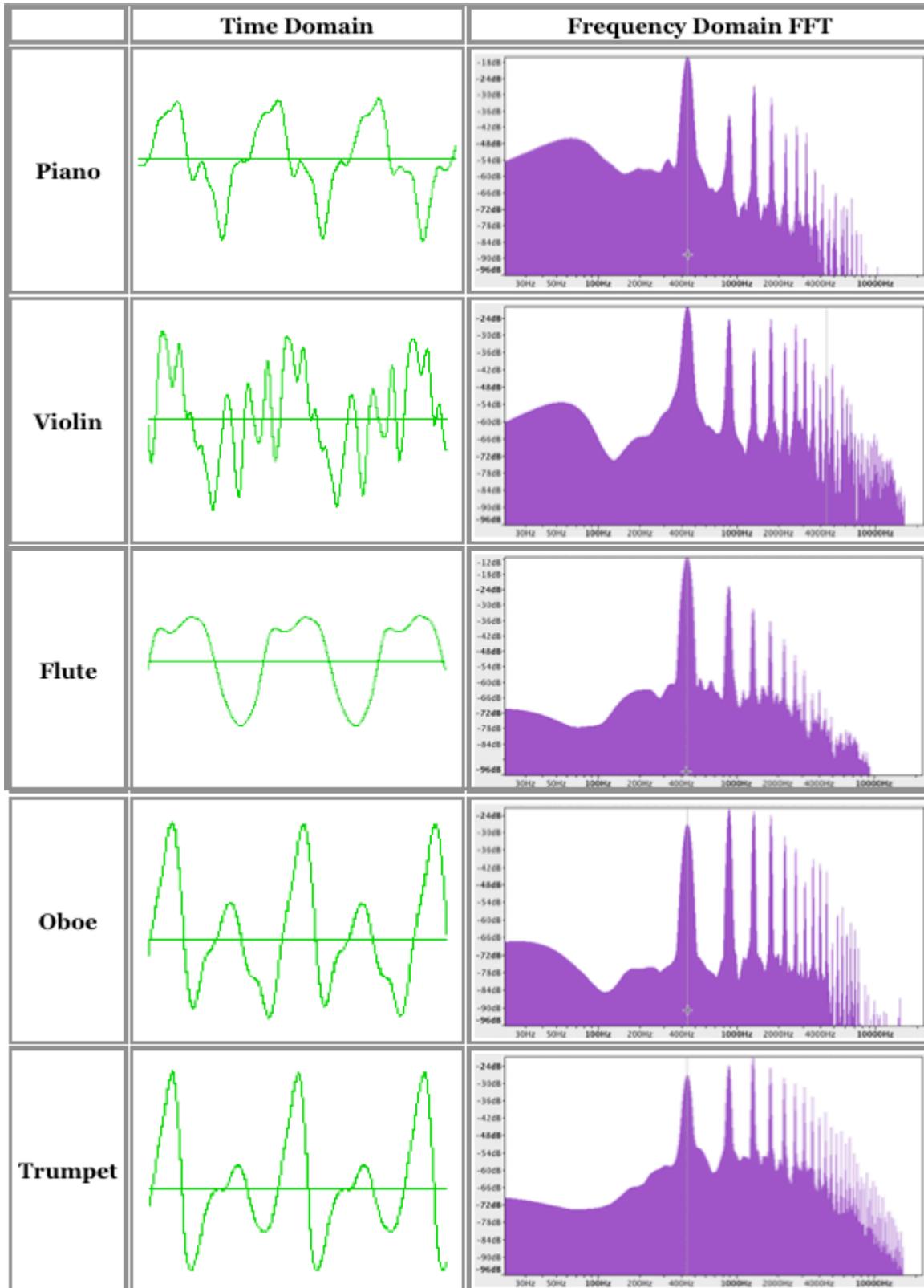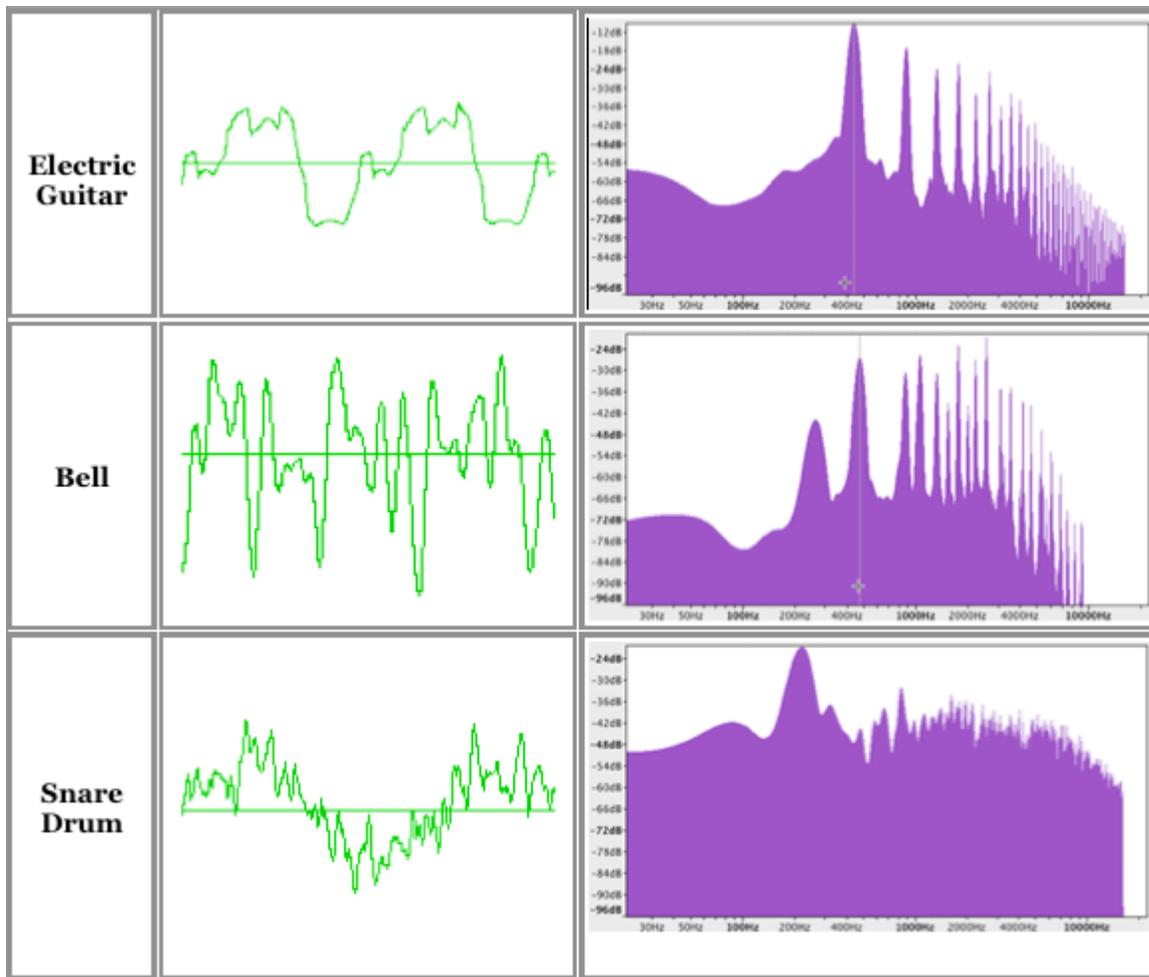
# Envelope

Envelope refers to the shape or contour of the amplitude of the sound as it evolves over time. A simple envelope consists of three parts: attack, sustain, and decay. An acoustic guitar has a sharp attack, little sustain and a rapid decay. A piano has a sharp attack, medium sustain, and medium decay. Voice, wind, and string instruments can shape the individual attack, sustain, and decay portions of the sound.

### Envelope In Music

The performers touch or breath can shape the envelope of a musical sound. Another term for this is articulation

### Envelope In Midi

A sound envelope in MIDI may be built in to the sound itself or it may be controlled through an Attack, Decay, Sustain, Release (ADSR) envelope. ADSR envelopes can sometimes be adjusted by tweaking knobs and buttons on the hardware or in the software. An ADSR envelope can be

simulated in MIDI with volume or expression control messages.

**Envelope In Digital Audio**

The envelope of a sound is the outline of the waveform's amplitude changing over time as seen on an oscilloscope or in sound editing software.
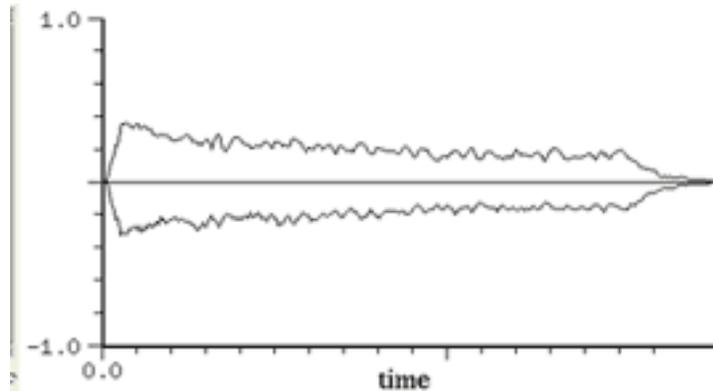


You can mathematically "envelope" a waveform by creating a second wave (amplitude values 0-1) that represents the envelope. The second wave must be the same duration as the first wave. When you multiply the samples of the original wave by the envelope wave, you create a third waveform conformed to the shape of the envelope.
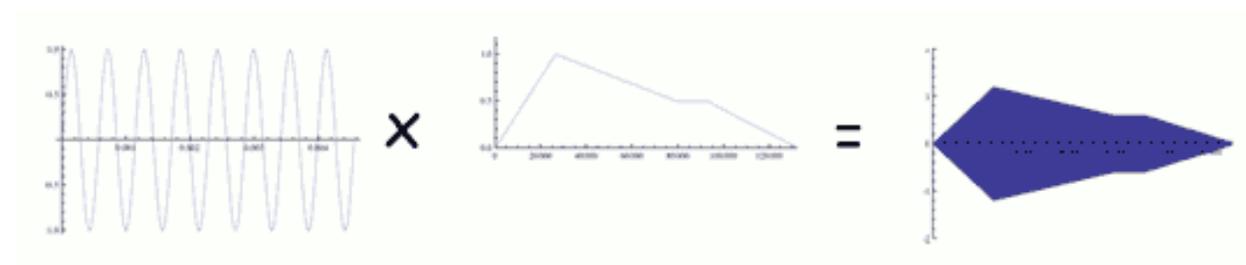


# Location

Location refers to the listener's perception of where the sound originated. Sound is perceived in three dimensional space based on the time difference it reaches our left and right eardrums.

**Location In Music**

The location of the sound is usually not specified in the score but is determined by the standing or seating arrangement of the performers.

**Location In Midi**

MIDI restricts the location of sound to the two dimensional left right stereo field. The MIDI Pan control message can be used to position the sound from far left to far right or anywhere in between. If the sound is centered, equal volumes appear in the left and right speaker. If more signal is sent to the left speaker than to the right speaker, the sound will be heard coming from the left.

**Location In Digital Audio**

Specialized binaural audio recording techniques can create a three dimensional sound sound image when heard over headphones. Binarual audio is not the same as stereo and loses it's effect when played over speakers.

A single waveform combined with multiple delayed and decaying versions of itself can simulate the reverberation characteristics of an acoustic space.