An Explanation Of Sampling Rate And Resolution

Presented by____

bobby prince

MUSIC™

-- continuing the

Free Exchange

of Helpful

Information

www.bpmusic.com

An Explanation Of Sample Rate And Resolution

Bobby Prince
bobby prince Music
www.bpmusic.com

The Physics of Sound

Sound is produced by vibration. Vibrations disturb the air to produce sound waves that travel out in all directions from the source of the sound. When the waves reach someone's ear, they set up vibrations that cause electrical signals to be sent to the brain. These electrical signals are perceived as sound.

Sound depends on three things: a vibrating source to set up sound waves, a medium (such as air) to carry the waves, and a receiver to detect them. The air waves are a result of the vibrating source moving rapidly in different directions. This movement pushes and pulls at the air next to the surface of the vibrating source. When the vibrating source moves in one direction, it crowds together those air molecules that are close to its surface. These molecules push outward against other molecules, continuing enough to set a **compression wave** into motion. The wave travels outward from the vibrating source, becoming weaker and weaker until it dies away. When the vibrating source begins to move in the opposite direction, it spreads the air molecules apart more than normal, creating a different type of wave. The series of these two types of waves traveling outward from the vibrating source make up what is called sound.

As you probably know from high school, Mr. Wizard, or Bill Nye the Science Guy, sound waves cannot travel through a vacuum. Of course this doesn't stop us from using them in the virtual vacuums of games and movies to help meet our **goal** above :)

Sound is one of the first sensations an unborn child experiences. Sound brings us our first lessons of life while we are still in the womb. We don't have to learn to detect the vibrations of sound. We take it all for granted from our earliest experiences.

The Three Basic Properties of Sound

1. **Pitch** -- this is the rate at which vibrations are produced, usually expressed in Hz (hertz, or cycles per second). One cycle is a complete vibration back and forth. The number of Hz is the frequency of the tone. The higher the frequency of a tone, the higher its pitch.

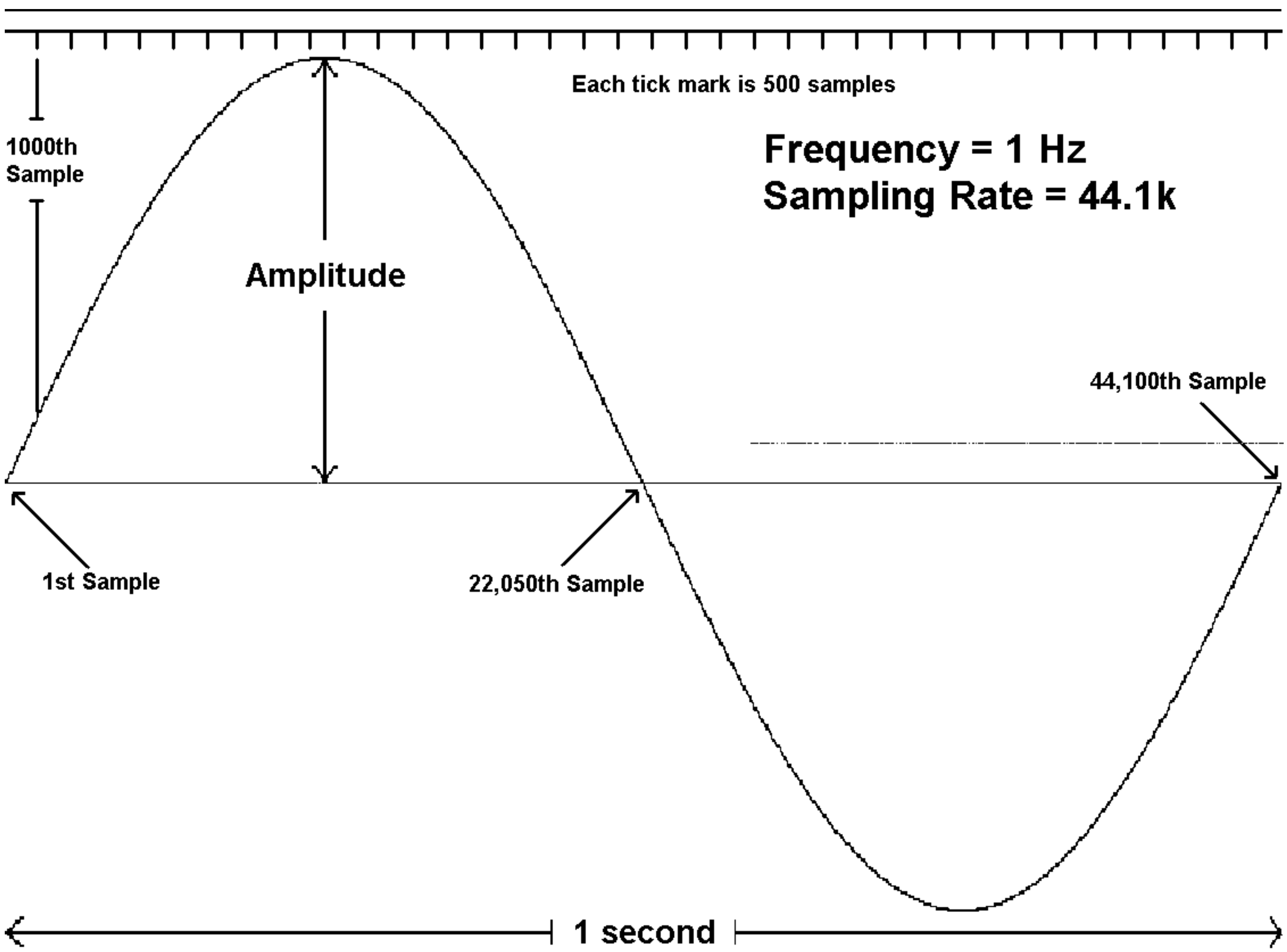

Figure 1

In Figure 1 we have a one Hertz vibration -- there is one complete vibration back and forth in one second. This wave reflects a slow pressure change in the one second period. You could not hear this sound, but at proper intensity and with the proper equipment, you could feel the pressure waves on your skin .

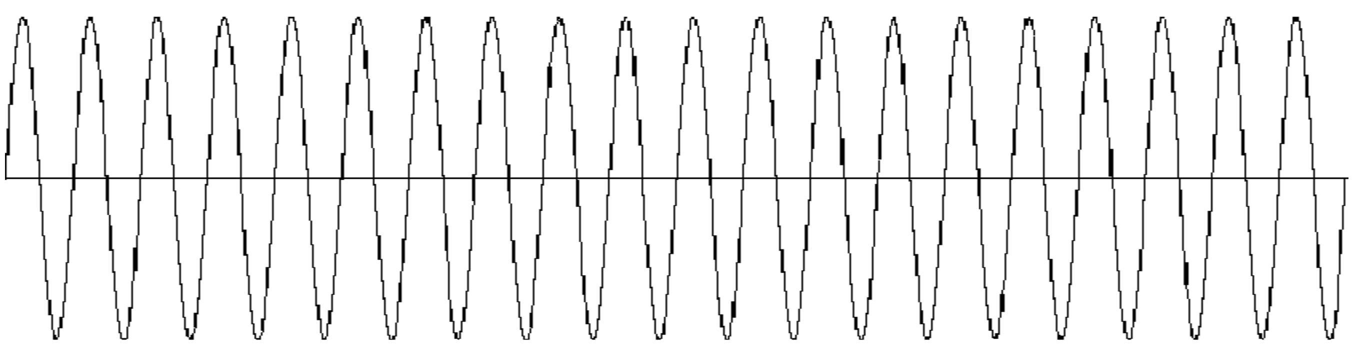

Figure 2

In Figure 2 we have faster pressure changes. This is a 20 Hz vibration. Notice that the individual vibrations are closer together. Also notice that the amplitude or intensity of this wave is less than the one Hz wave in Figure 1.

2. **Intensity** -- the intensity of a sound depends upon the strength, or **amplitude**, of the vibrations producing the sound. This has nothing to do with its pitch. Stronger vibrations produce a louder tone. A vibration of greater amplitude compresses the molecules of the air more forcefully and gives them greater energy. When a series of such strong compression waves enters the ear, the brain interprets it as a loud tone. The loudness of sounds is measured in decibels (dB). On the scale used, 0 indicates the softest audible sound. The rustle of leaves is rated as 20 dB, average street noise as 70, and nearby thunder as 120.

3. **Tone Quality** -- the quality, or timbre, of a sound is more complicated than pitch or intensity. The tone of a dove cooing is much more pleasant than the screech of a bluejay. This is true even if both sounds are at the same pitch and intensity. Why? Neither bird sound is a simple tone. Each contains multiple tones of varying pitch and intensity. These are called overtones. The difference is that the dove's combinations of tones harmonize well with one another while the jay's combinations of tones are extremely discordant. Overtones give sounds special character. Without overtones every vibrating source would sound very much alike.

Things can get a lot more complicated than the simple sine wave.

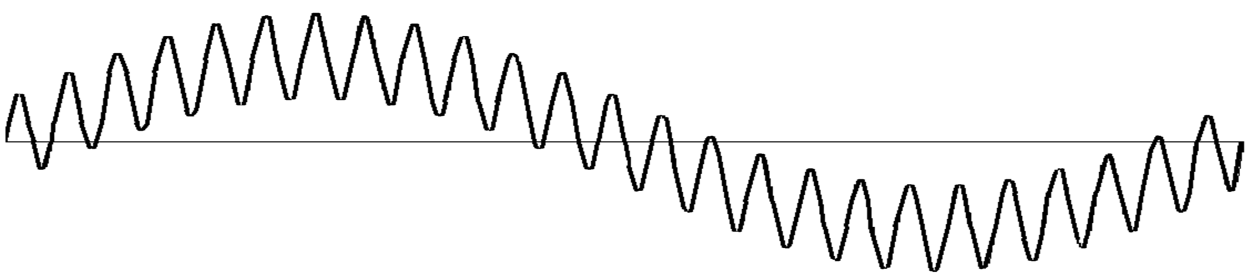

Figure 3

In Figure 3, we have a basic one Hz sine wave, but notice that it consists of 25 smaller sine waves. This definitely changes the tone quality of the basic wave.

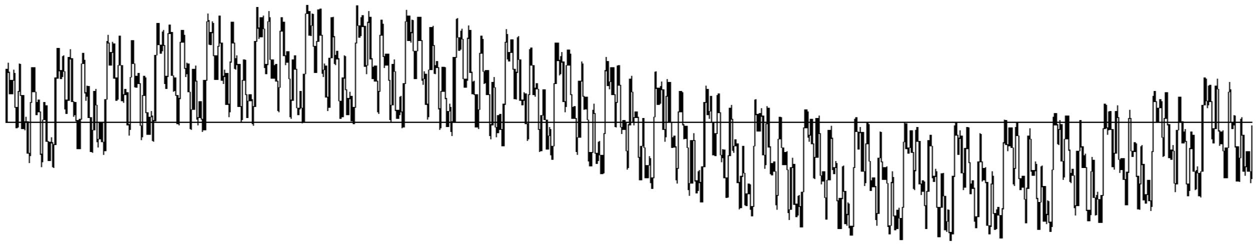

Figure 4

And Figure 4 is our same basic sine wave with multiple overtones. These overtones further change the tone quality of the basic wave. To make matters even more complicated, the basic tone and the multiple overtones can change over time, causing the sound to change with them. A good example of this is the sound of a cymbal.

Going back to the Figure 1, notice the amplitude moves from zero to positive back through zero to negative and back to zero. If you look at this as the air pressure mentioned above, the left half of this wave is the compression wave and the right half is the wave where the air molecules are more spread out than normal. Basically you could say that this is a chart of the air pressure over a one second interval. It is a picture of the positive and negative pressure waves that make up sound. The amplitude reflected here is the volume of the sound. The greater the amplitude (amount of pressure change) the greater the volume.

Recording and Storing Sounds

In the early days of movie making it was discovered that the human eye is fooled into seeing lifelike moving pictures if the individual frames pass in front of the eye at 24 frames per second. Similarly, the audio community experimented with sound to see how many snapshots of a sound must be taken per second to fool the ear into hearing lifelike sounds from digital signals. It was decided that 44,100 audio snapshots would do the same trick for the ears as 24 frames per second does for the eyes. These snapshots are called **samples** and they are the equivalent of a video frame in a movie. The audio equivalent of the number of video frames per minute is the **sampling rate**. You have probably heard the term sample used for a complete sound. It is also used to describe the smallest part of a digital sound.

Why did audio engineers decide upon a sampling rate of 44.1 kHz for audio CD's?

Very few people can hear any fewer than 16 Hz or any more than about 20 kHz (thousand cycles per second). The question then is what sampling rate is required to record a 20 kHz sound? This question is answered in the Nyquist formula: the sampling rate of a sound must be twice the frequency of the highest sound to be sampled. You might ask what will happen if you record something at too low a sampling rate, disregarding the Nyquist formula. The answer is that the average person would not think that the sound was lifelike.

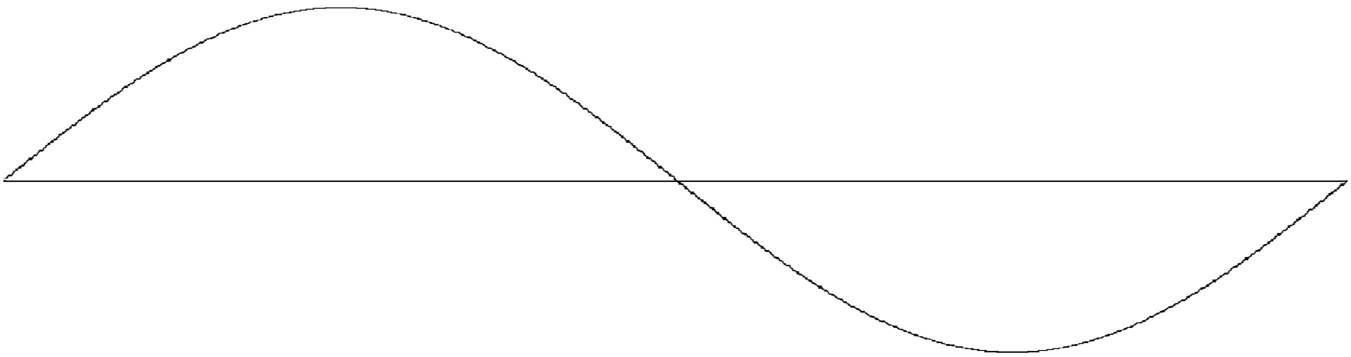

Figure 5

If you record a sound with a microphone to a cassette tape, the microphone acts like your ear and converts the sound made into electrical signals. Those electrical signals are converted to electromagnetic signals that are stored on the cassette tape. The signals recorded onto the tape are **analog**. This means that the recorded sound is represented by a continuously variable electromagnetic signal. Figure 5 is a representation of a one second analog sine wave. If I give you a plastic ruler that has amplitude levels in the place of inches and ask you to tell me what the amplitude is at **any** time in this one second sound, you could answer **exactly and accurately** because the amplitude levels reflected here were measured and stored continuously.

Referring back to Figure 1, we see that the amplitude levels were sampled 44,100 times during the one second of recording. Another way of saying this is that data was collected on pressure levels every 1/44,100th of a second. If I were to ask you to tell me the amplitude at any time other than a multiple of 1/44,100th of a second, you could not answer exactly or accurately because there would be no data for that point in time. You would have to look at the data on each side of the time and interpolate from there.

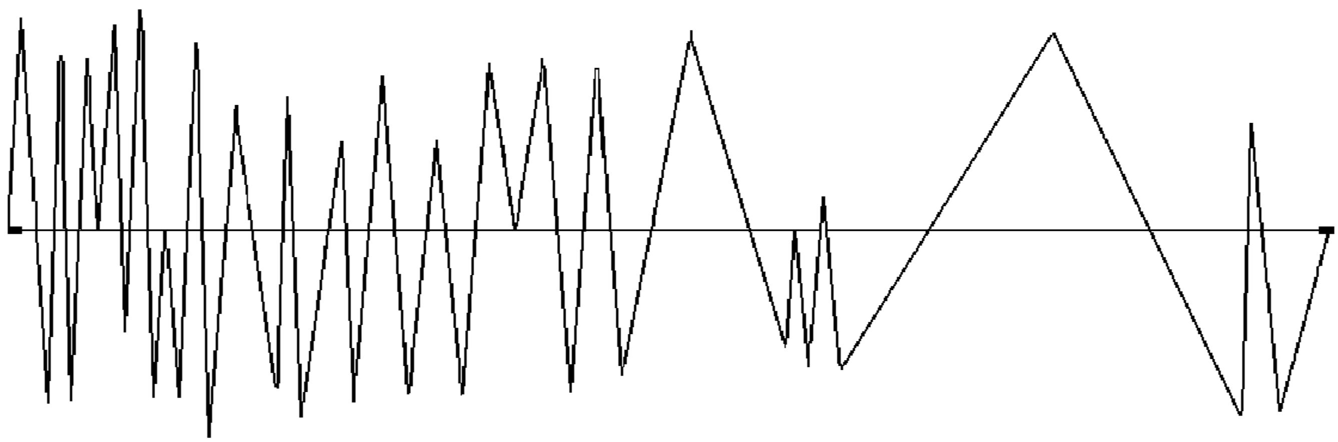

Figure 6

Figure 6 is a fairly complex analog waveform. To digitize it, we have to take snapshots of it at regular intervals. Let's sample this waveform at 100 Hertz. Under the Nyquist formula, a 100 Hz sampling rate would be good for digitizing frequencies of 50 Hz and below. If we place a grid with 100 measuring points over the waveform, our resulting waveform looks like this:

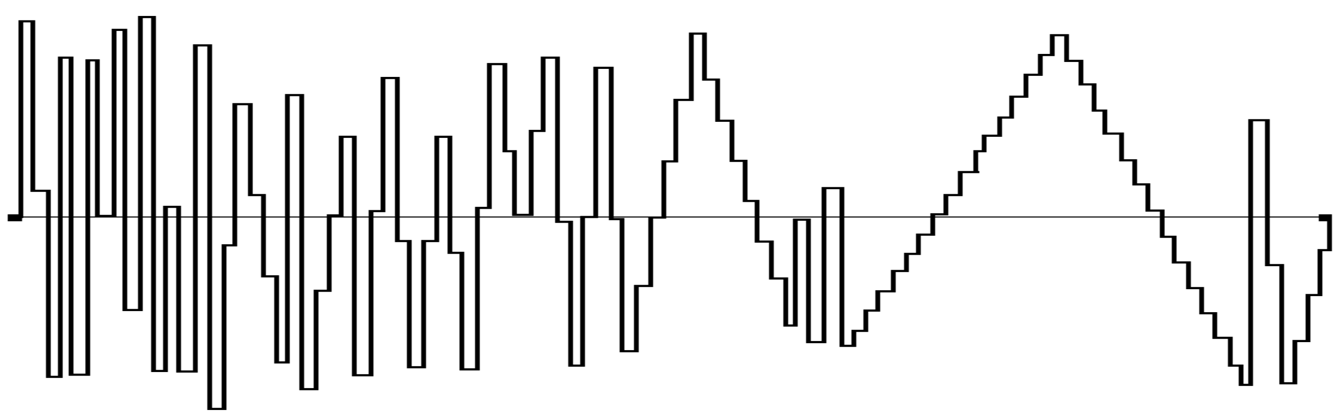

Figure 7

If we overlay this waveform and the original, we can see that most of the analog wave is represented well. Now let's cut the sampling rate in half and see what happens. This will require that we take measurements of the waveform in Figure 6 at fifty even points along the one second line.

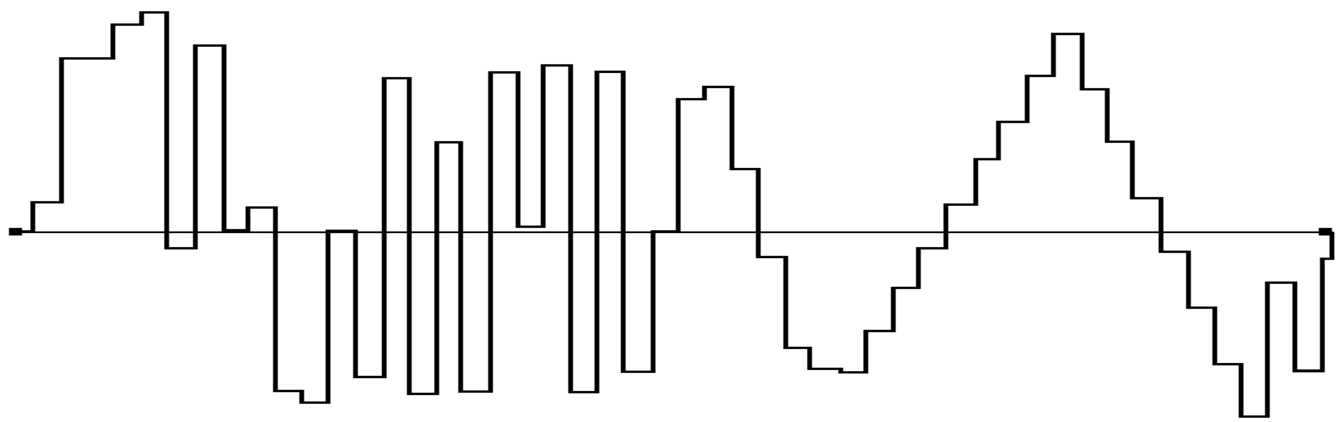

Figure 8

Figure 8 is the resulting digital waveform. Comparing it to the 100 Hz sample (Figure 7) and the original analog waveform (Figure 6), it is easy to see that it is not a very accurate representation when the frequency is high. When the frequency is low, it is as accurate a representation as the 100 Hz sample. Notice though that a lot of information is lost as a result of the lower sampling rate. Can you see Nyquist's formula at work here?

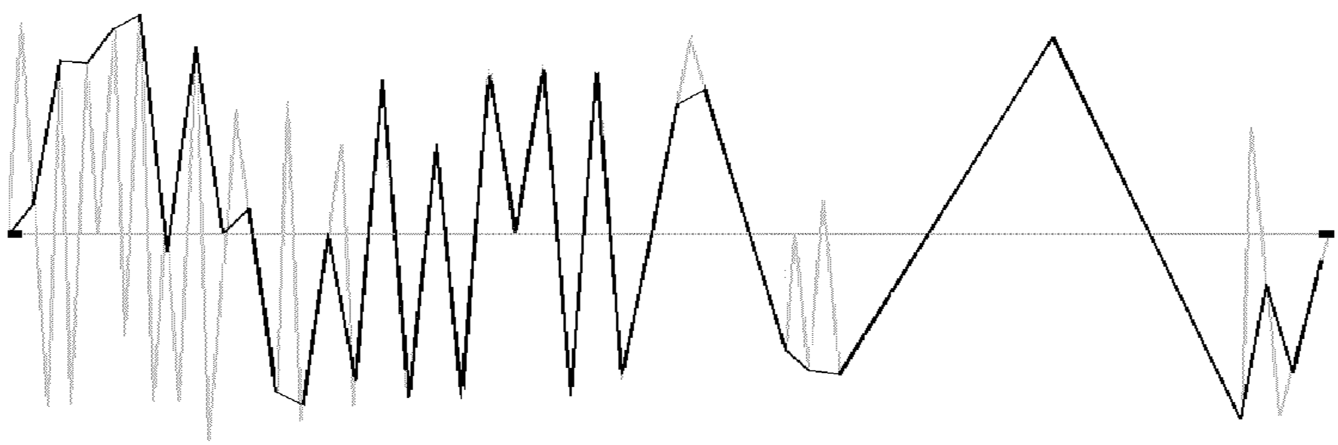

Figure 9

In Figure 9 you can see that when the frequency is 25 Hz or less (one cycle of the waveform takes up at least 1/25th of the time line), the 50 Hz sampling rate (black) records the same data as the 100 Hz rate (gray). The problems occur when the frequency goes above 25 Hz. At that point, the 50 Hz sampling rate reflects a completely erroneous waveform.

How do we store the digital samples of the one second sound?

We take snapshots of the amplitudes at regular intervals and store the values. Remember that we are looking at values which actually represent air pressure at any one point in time. So a digital sound file is merely a method of storing air pressure!

How much air pressure can be stored in one byte? As you know, one byte (8 bits) can have a value from 0 to 255 decimal. For this reason, 8-bit digital audio can store 256 discrete air pressure levels. 16 bit digital audio can store air pressure much more precisely, allowing values from 0 to 65535, or 65536 discrete air pressure levels.

Let's store the amplitude of Figure 1 in low resolution using a grid overlay to decide on the values. This grid has five values, each being represented by one of the horizontal lines shown. We will round off the analog values of the sine to the nearest horizontal line.

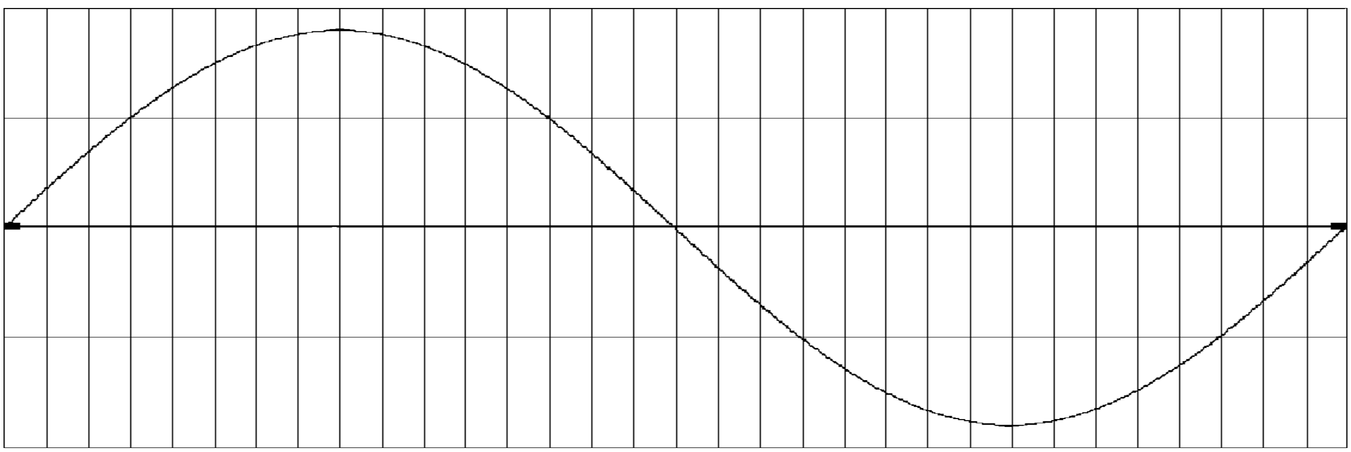

Figure 10

The resulting digital representation is as follows in Figure 11:

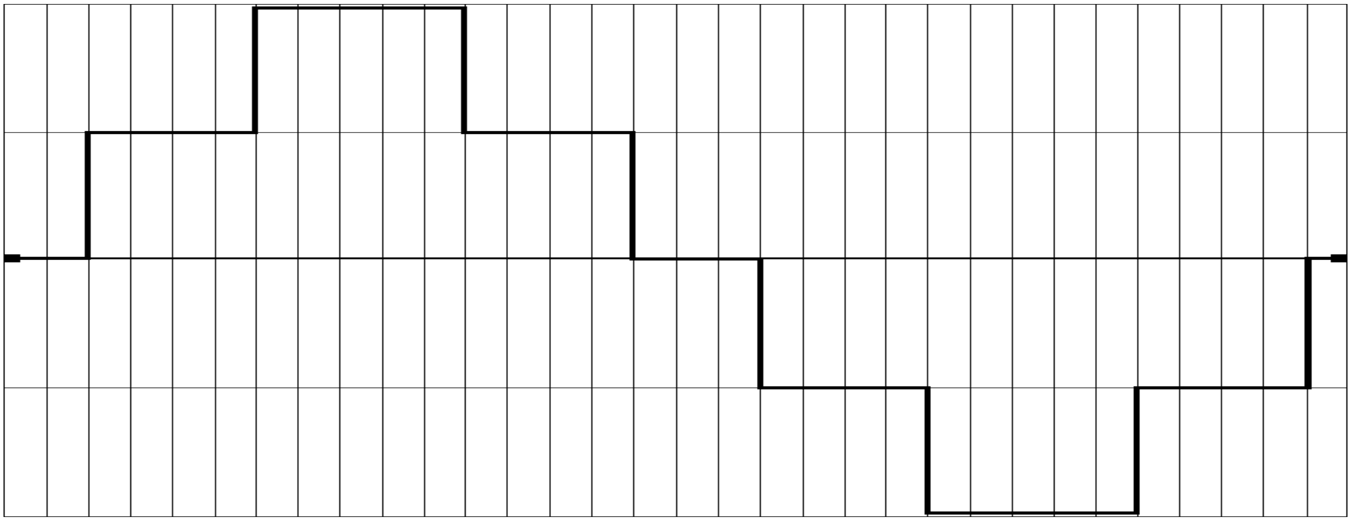

Figure 11

As you can see, this is not a very accurate reflection of the correct amplitudes. Now let's use a grid with the same sampling rate but a higher amplitude resolution.

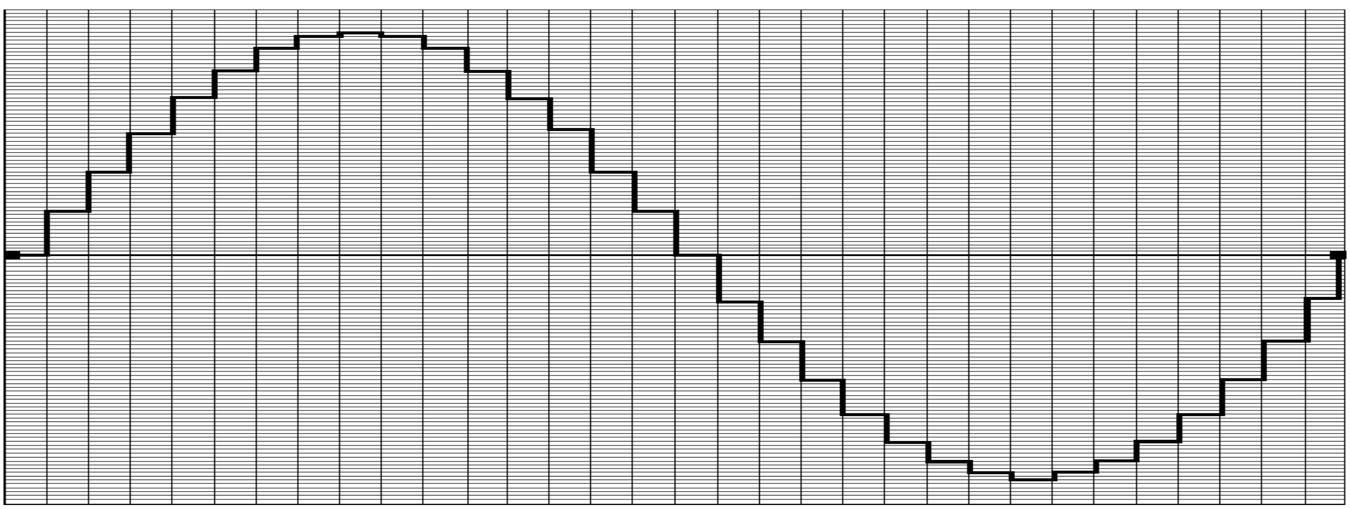

Figure 12

Figure 12 has 32 times more amplitude lines than Figure 11 (16 bit would have 256 times more than 8 bit). Notice that the higher resolution in Figure 12 allows for a much more accurate representation of our sine wave. Resolution is very important at low sampling rates. Let's look now at these same resolutions at twice the sampling rate. First, low resolution:

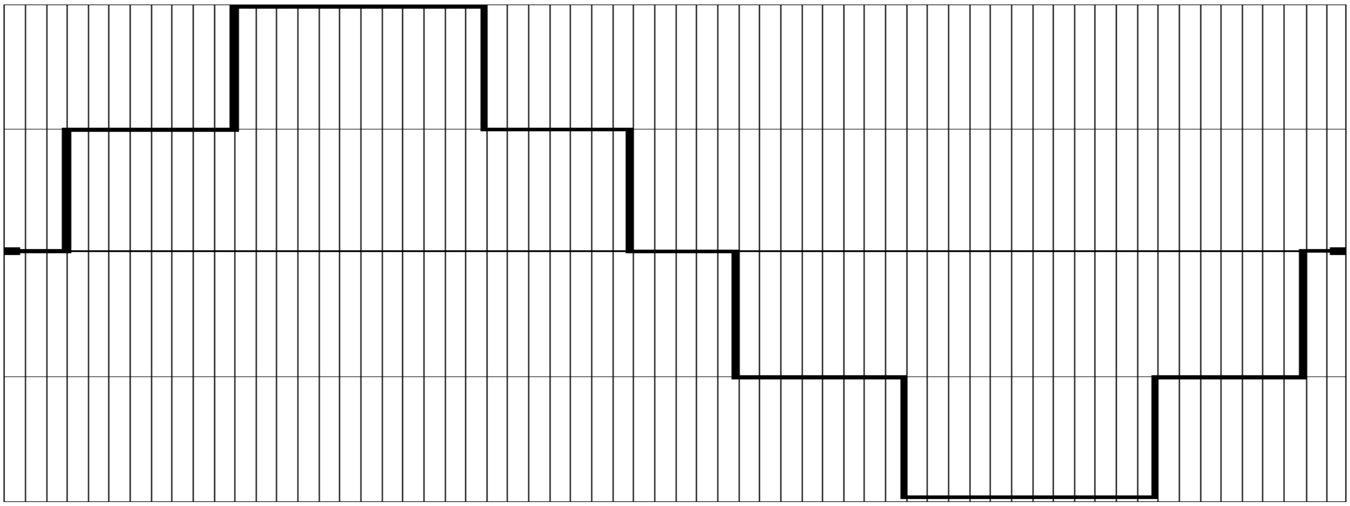

Figure 13

There is virtually no difference between Figure 11 (low resolution, low sampling rate) and Figure 13 (low resolution, high sampling rate). Why? Remember the Nyquist formula -- here the frequency is much lower than half of either sampling rate. Both **sampling rates were greater than the Nyquist formula requires.** Let's raise the resolution and use this same higher sampling rate.

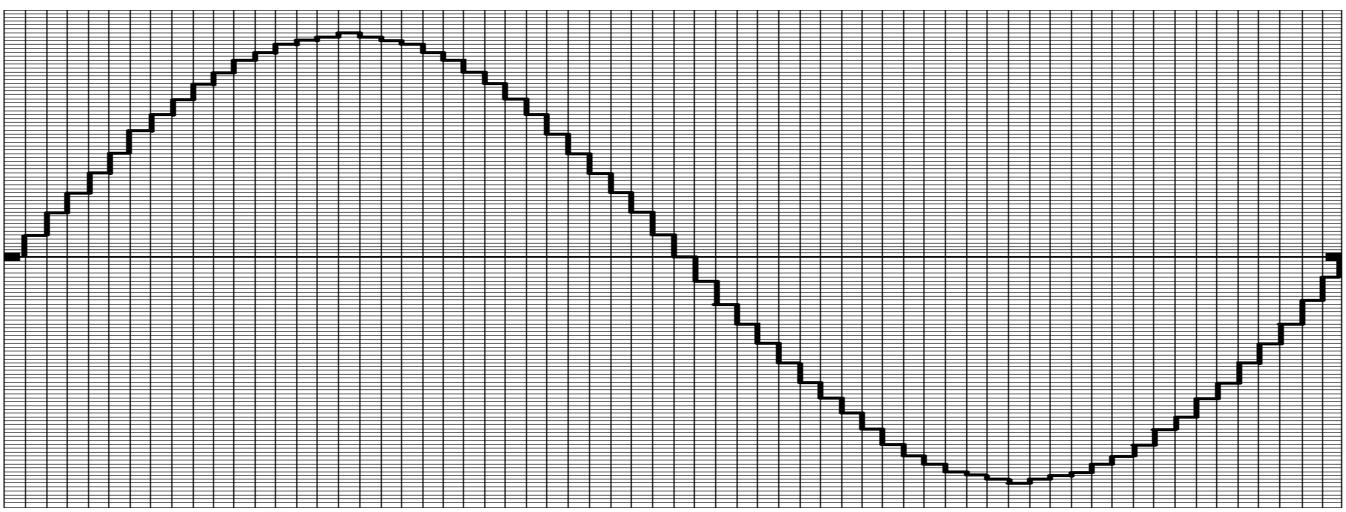

Figure 14

Figure 14 is a very accurate digital representation of the analog sine wave. Resolution is very important at high sampling rates, too.

Another way to visualize the difference in 8 and 16 bit audio is to consider a home stereo that has a volume knob with "detents" in it. You know the type that clicks as you turn it? Well, imagine that you buy a tuner/amplifier that has a knob with positions (clicks) from 0 or, no volume to 65535, or full volume. The volume change from one click to an adjacent one would be indistinguishable. Now take the same knob and give it positions (clicks) from 0, or no volume to 255, or full volume. Now you have only 256 discrete volume settings. You can see how coarse your volume changes will be from click to click -- you can hear each change.

How much space is required to store the different types of samples?

Audio CDs have stereo samples at 44.1 kHz, 16 bit resolution. That resolution requires 2 bytes per sample. At 44,100 samples per second, each taking 2 bytes (16 bits) of precision, we need 88,200 bytes per second for storage. Since stereo is two tracks we have to double that: 176,400 bytes per second. Take that times 60 seconds per minute and you get 10,584,000 bytes per minute. So one minute of stereo 44.1k 16 bit sound requires almost 10 megabytes of storage (uncompressed).

Storage Requirements for One Minute of Sound

Type:	Mono	Mono	Stereo	Stereo
Resolution:	8 bit	16 bit	8 bit	16 bit
Sampling Rate				
44.1k	2646k	5292k	5292k	10584k
22.05k	1323k	2646k	2646k	5292k
11.025k	661.5k	1323k	1323k	2646k
8k	480k	960k	960k	1920k
7k	420k	840k	840k	1680k
6k	360k	720k	720k	1440k
5k	300k	600k	600k	1200k

Figure 15

The table in Figure 15 indicates the different storage requirements for different types, sample rates and resolutions of one minute of digital sound. The decision one makes regarding these variables should be based upon best balance of storage requirement and sound quality. If the sound effects we want to use have a large dynamic range (amplitudes from very low to very high), we would probably want to use 16 bit resolution. If all sounds are going to be about the same volume, 8 bit should suffice.

The decision regarding sampling rate only requires that we look to the Nyquist formula. The more we disregard that formula, the more we reduce sound quality.