Multimedia Communication (SW-416)

SOUND

Dr. Areej Fatemah

email: areej.fatemah@faculty.muet.edu.pk

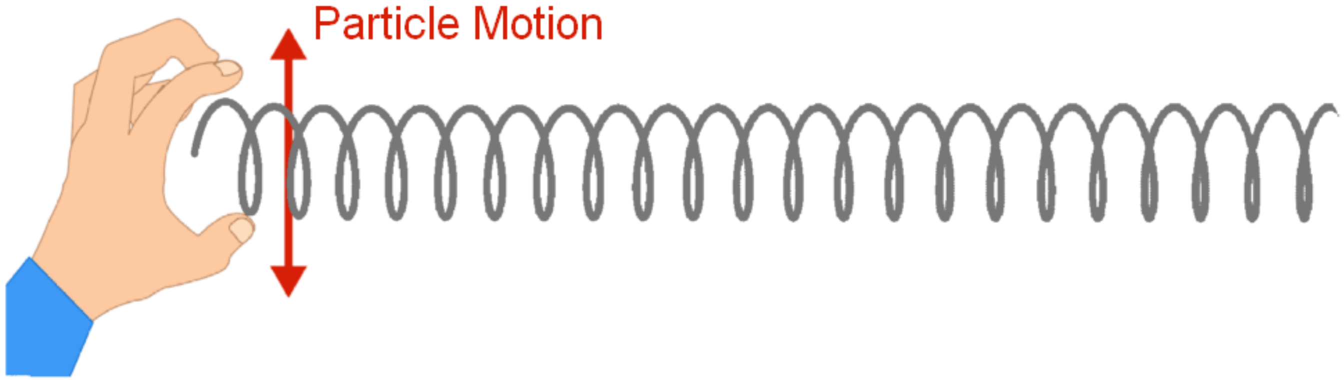
Sound

- Ever wondered why some music/sound that you hear in games or on the Internet sounds better than others? You might even notice that if you listen to something on the TV, the same sound/song sounds better when you stream from the Internet.
- ***This has to do with the amount of data being used to store the sound, whether that's on a CD, your computer or the Internet.***

- Sound, a mechanical disturbance from a state of equilibrium that propagates through an elastic material medium.

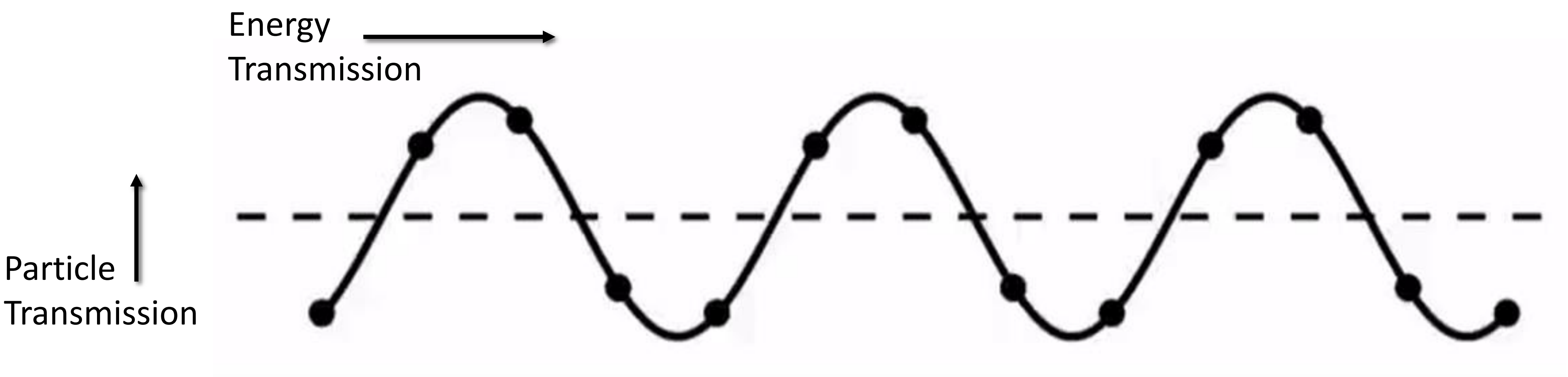
Sound

- The study of sound should begin with the properties of sound waves.
- There are two basic types of wave, transverse and longitudinal, differentiated by the way in which the wave is propagated.
- In a transverse wave, such as the wave generated in a stretched rope when one end is wiggled back and forth, the motion that constitutes the wave is perpendicular, or transverse, to the direction (along the rope) in which the wave is moving.



Sound

A transverse wave is a wave in which particles of the medium move in a direction perpendicular to the direction that the wave moves.

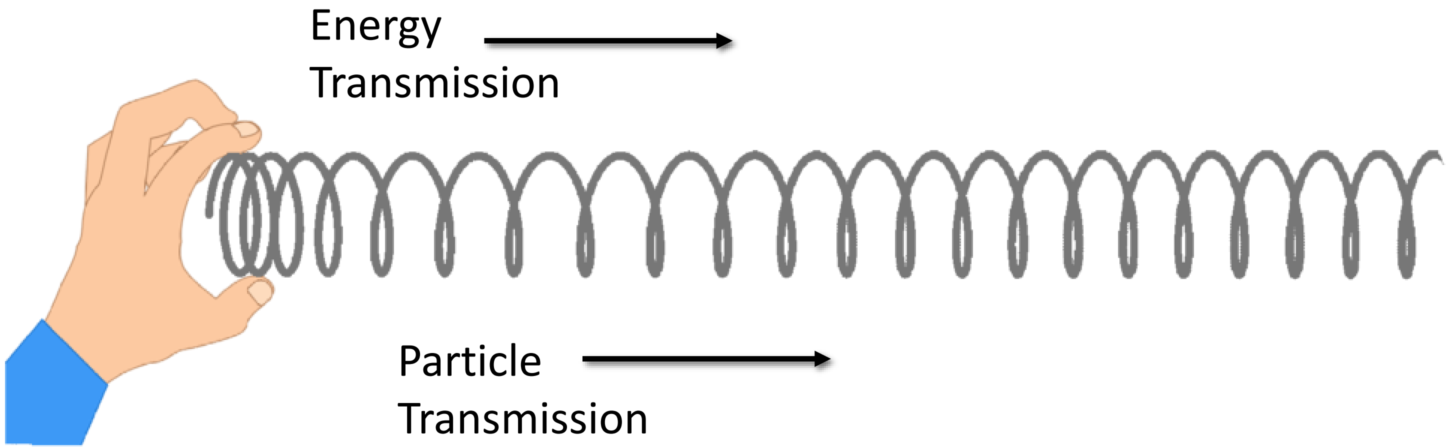


- Examples of transverse waves include seismic (secondary) waves, and the motion of the electric and magnetic fields in an electromagnetic plane wave, which both oscillate perpendicularly to each other as well as to the direction of energy transfer.
- Transverse waves are generated by electromagnetic sources such as light or radio, in which the electric and magnetic fields constituting the wave oscillate perpendicular to the direction of propagation.

Sound

- Sound propagates through air or other mediums as a longitudinal wave, in which the mechanical vibration constituting the wave occurs along the direction of propagation of the wave.
- A longitudinal wave can be created in a coiled spring by squeezing several of the turns together to form a compression and then releasing them, allowing the compression to travel the length of the spring.

A longitudinal wave is a wave in which particles of the medium move in a direction parallel to the direction that the wave moves.

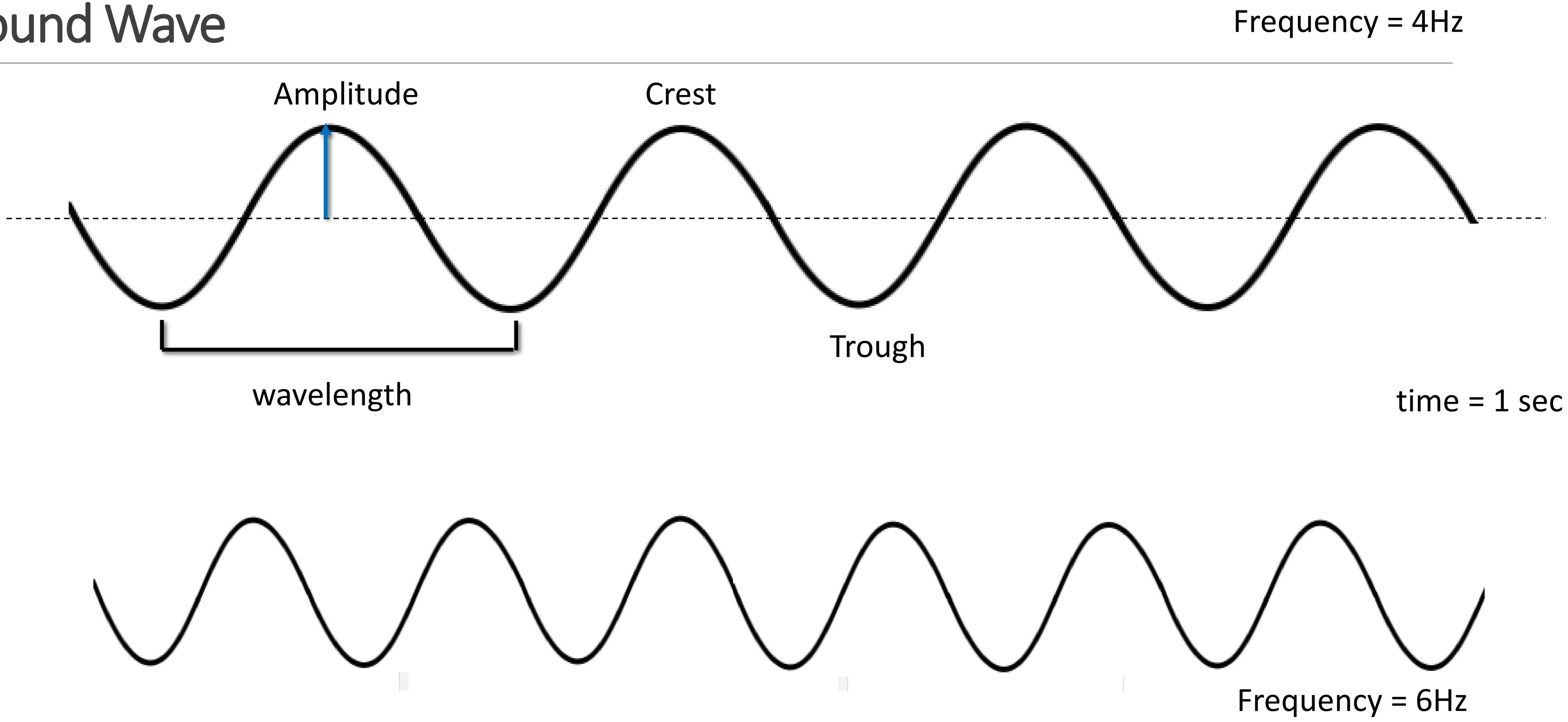


Sound

- Air can be viewed as being composed of layers analogous to such coils, with a sound wave propagating as layers of air “push” and “pull” at one another much like the compression moving down the spring.

- A sound wave thus consists of alternating compressions and rarefactions, or regions of high pressure and low pressure, moving at a certain speed.
- Put another way, it consists of a periodic (that is, oscillating or vibrating) variation of pressure occurring around the equilibrium pressure prevailing at a particular time and place.

Sound Wave


Frequency = 4Hz
Amplitude
Crest
Trough
wavelength
time = 1 sec
Frequency = 6Hz

Sound Waves

- Sound waves can be characterized by the attributes like period, frequency, amplitude, bandwidth, pitch, loudness.

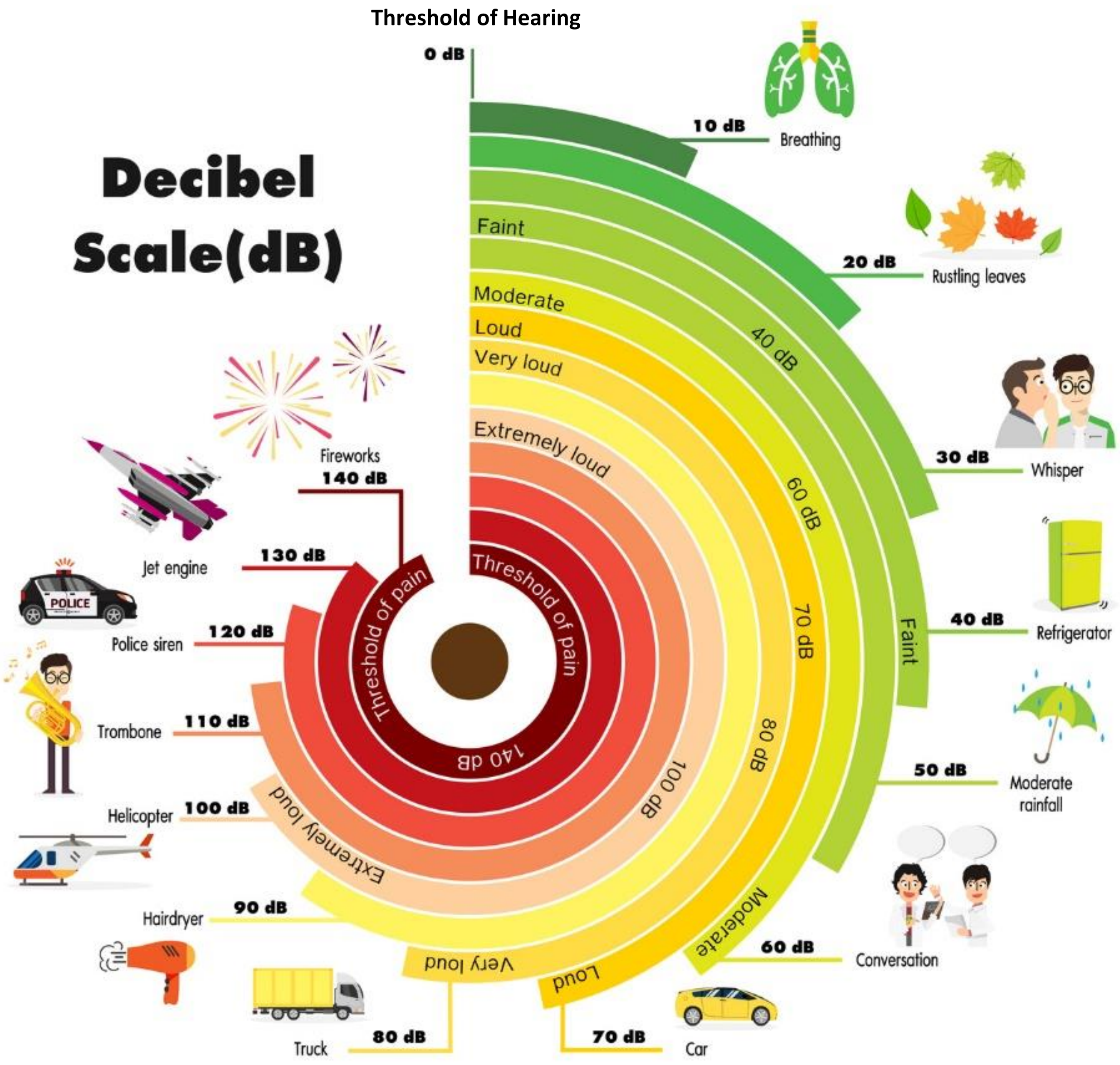
- *Period* is the interval at which a periodic signal repeats regularly.
- *Frequency* is determined by number of cycles of vibration in one second (relates to pitch).
- *Wavelength* is the distance between peaks.
- *Amplitude* is the measure of sound levels. Height of a cycle (relates to loudness).
- *Bandwidth* is the range of frequencies a device can produce, or a human can hear.

Threshold of Hearing and the Decibel Scale

- The amount of energy that is transported past a given area of the medium per unit of time is known as the intensity of the sound wave.
- The greater the amplitude of vibrations of the particles of the medium, the greater the rate at which energy is transported through it, and the more intense that the sound wave is.
- Typical units for expressing the intensity of a sound wave are Watts/meter^2.
- As a sound wave carries its energy through a two-dimensional or three-dimensional medium, the intensity of the sound wave decreases with increasing distance from the source.
  - Since energy is conserved and the area through which this energy is transported is increasing, the intensity (being a quantity that is measured on a *per area* basis) must decrease.

Threshold of Hearing and the Decibel Scale

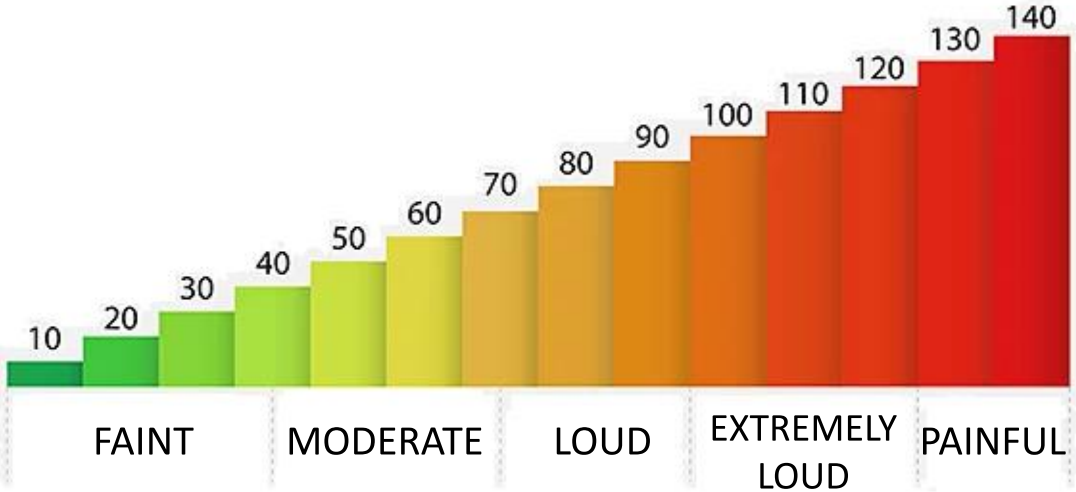
- Humans are equipped with very sensitive ears capable of detecting sound waves of extremely low intensity.
- The faintest sound that the typical human ear can detect has an intensity of $1*10^{-12}$ W/m^2
- This faintest sound that a human ear can detect is known as the threshold of hearing.
  - The most intense sound that the ear can safely detect without suffering any physical damage is more than one billion times more intense than the threshold of hearing.

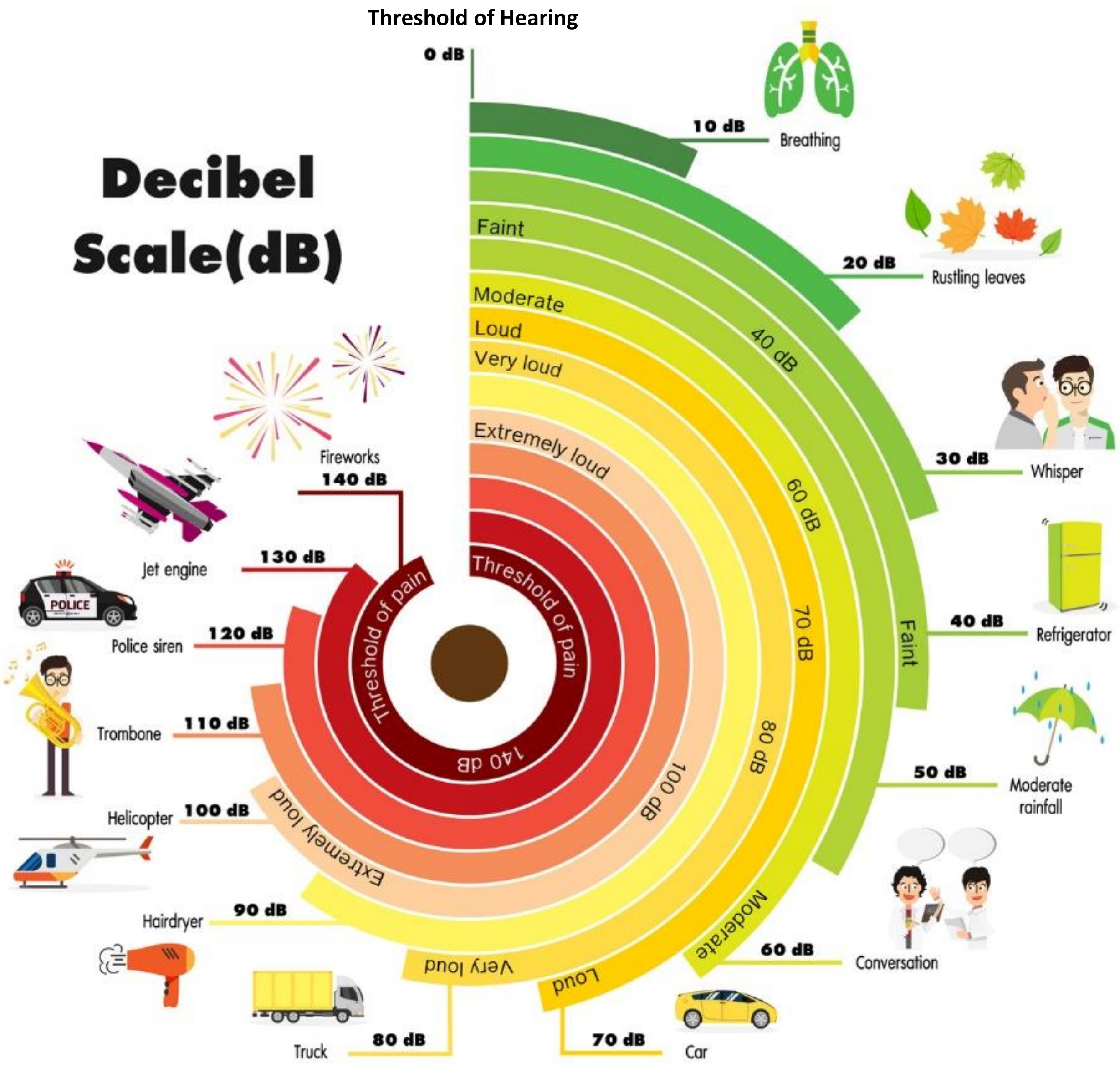


Interpretation of the Decibel Scale

- 0 dB = threshold of hearing (TOH)
- 10 dB = 10 times more intense than TOH
- 20 dB = 100 times more intense than TOH
- 30 dB = 1000 times more intense than TOH

- An increase in 10 dB means that the intensity of the sound increases by a factor of 10

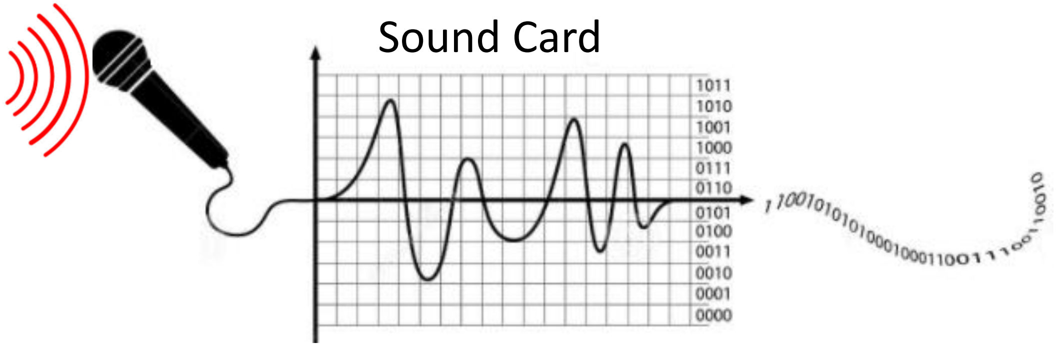




Sound Masking

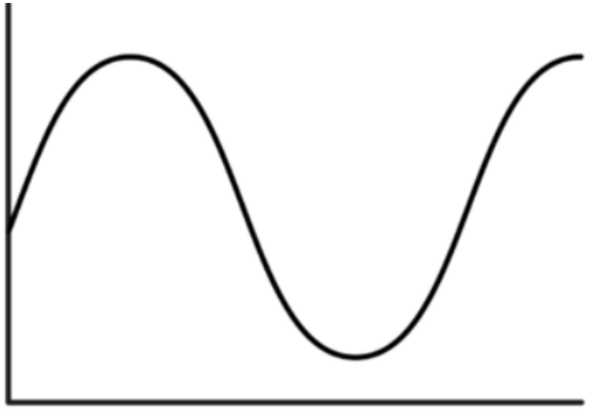
- Masking occurs when the perception of a sound is affected and covered by another, distracting the ear from being able to clearly perceive the simultaneous sounds.
  - Basically, two sounds that cover the same frequencies.
  - Perception of one sound interferes with another

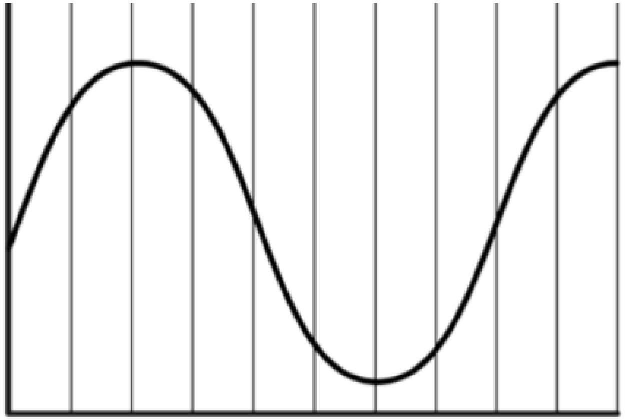
1. Frequency Masking
   - An auditory phenomenon that can occur when two or more sounds sharing similar frequencies are played at the same time. In the range where the frequencies overlap one sound will overlap the other, confusing your perception of either sound.
2. Temporal Masking
   - When exposed to a loud sound, the human ear contracts slightly to protect delicate structures. Causes louder sounds to overpower weaker sounds just before and just after it.



Digitization of Sound

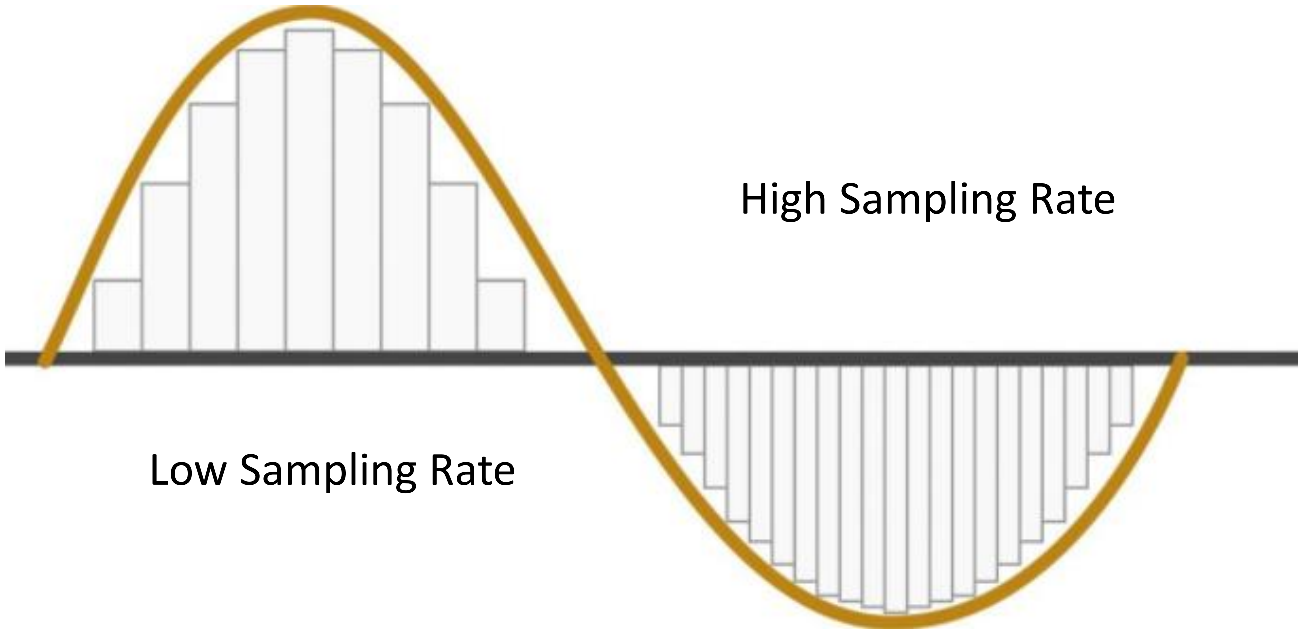
- Digitization is a process of converting the analog signals to a digital signal. There are three steps of digitization of sound.
- **Sampling** - Sampling is a process of measuring air pressure amplitude at equally spaced moments in time, where each measurement constitutes a sample.
  - Sampling is the first step towards transforming an analog audio source into a digital file. Digital audio recording does this by taking samples of the audio source along the soundwaves at regular intervals.





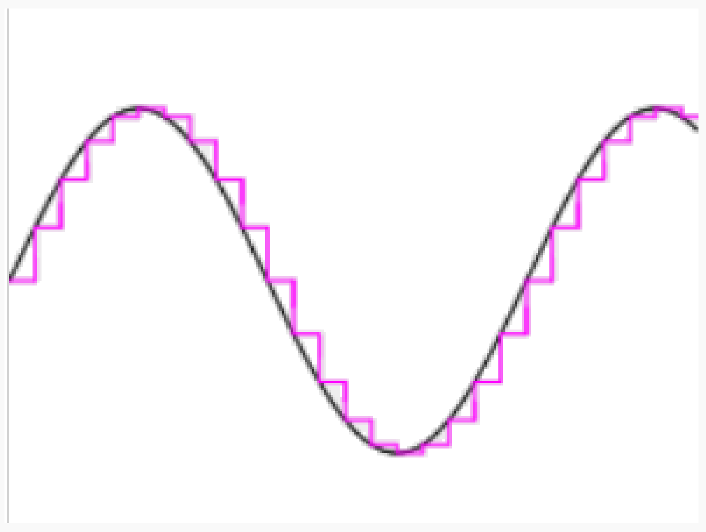
Digitization of Sound
Sampling

- The more samples you take - known as the ‘sample rate’ - the closer the final digital file will resemble the original. A sampling rate is the number of times the analog sound is taken per second.
  - A higher sampling rate implies that more samples are taken during the given time interval and ultimately, the quality of reconstruction is better.
  - sampling rate = 1 / sampling interval

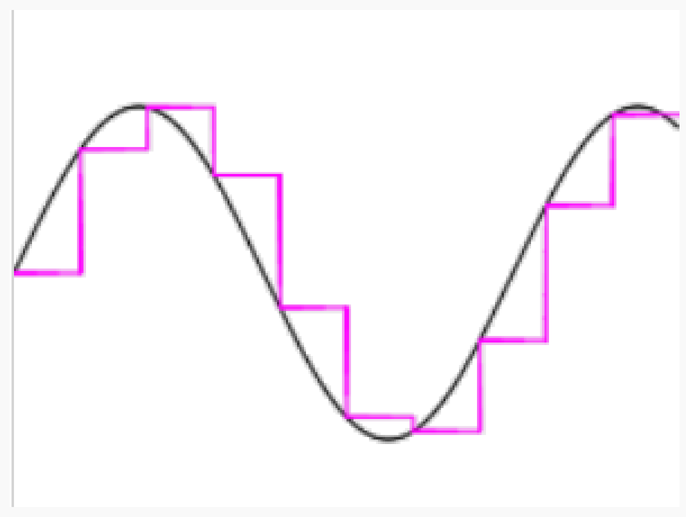


Digitization of Sound
Sampling

- Sampling rate is measured in terms of Hertz, Hz in short, which is the term for Cycle per second or number of samples per second. CDs are usually recorded at 44.1kHz - which means that every second, 44,100 samples were taken.
- Sampling rates most often used in multimedia are 44.1kHz(CD-quality), 22.05kHz and 11.025kHz.



Sampling Interval: 15 ms

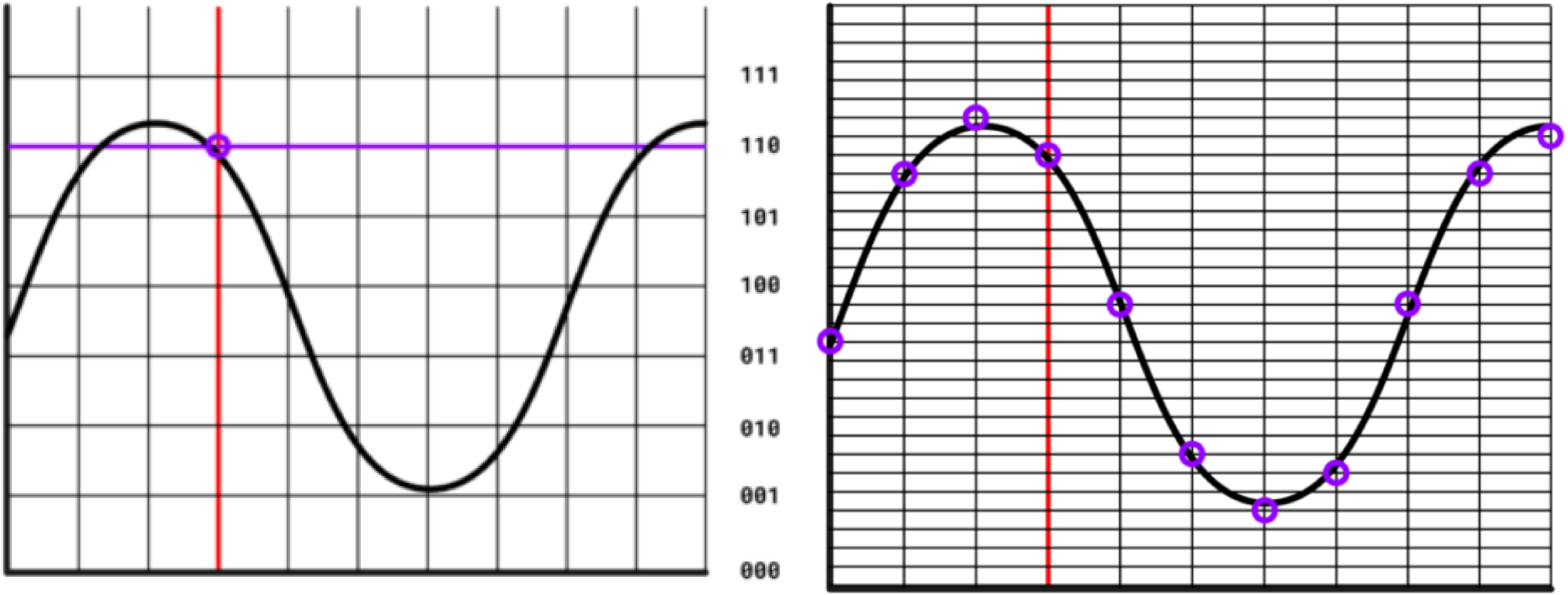


Sampling Interval: 40ms

Digitization of Sound
Quantization

- **Quantization** - Quantization is a process of representing the amplitude of each sample as integers or numbers or representing a point on a continuous scale with a discrete value.
    - The amount of numbers used to represent the value of each sample. Also known as sample size or bit depth or resolution.
- Commonly used sample sizes are either 8 bits, 16 bits or 24 bits.
    - The larger the sample size, the more accurately the data will describe the recorded sound.
- An 8-bit sample size provides 256 equal measurement units to describe the level and frequency of the sound in that slice of time. A 16-bit sample size provides 65,536 equal units to describe the sound in that sample slice of time.
    - The value of each sample is rounded off to the nearest integer (quantization) and if the amplitude is greater than the intervals available, clipping of the top and bottom of the wave occurs.

Digitization of Sound
Quantization



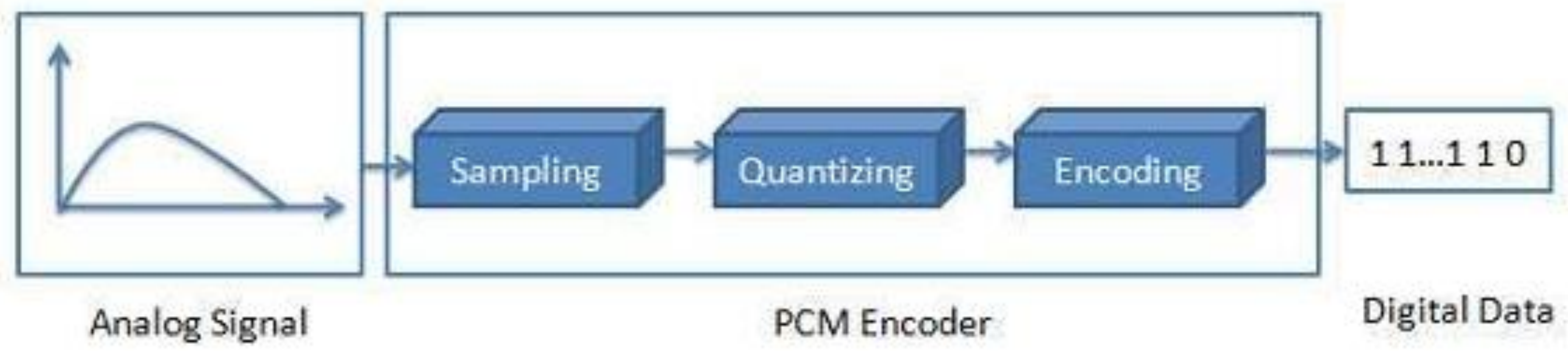
Digitization of Sound
Quantization

- Although there is a difference between the measured value and the stored value, a high enough sample resolution (using more bits) lets the computer get very close to the actual signal level. It’s quite common to use 16 or 24 bits, allowing the computer to represent 65536 or 16777216 different levels.

- This is a similar concept to the one covered in images
- Recall that bit depth was defined as the storage space each pixel needs to represent the available range of different colours; more variation of colours produce a more detailed image. Sample resolution is the matching concept for sound files: a higher sample resolution produces a more detailed sound recording. This is why sample resolution is also called **audio bit depth**

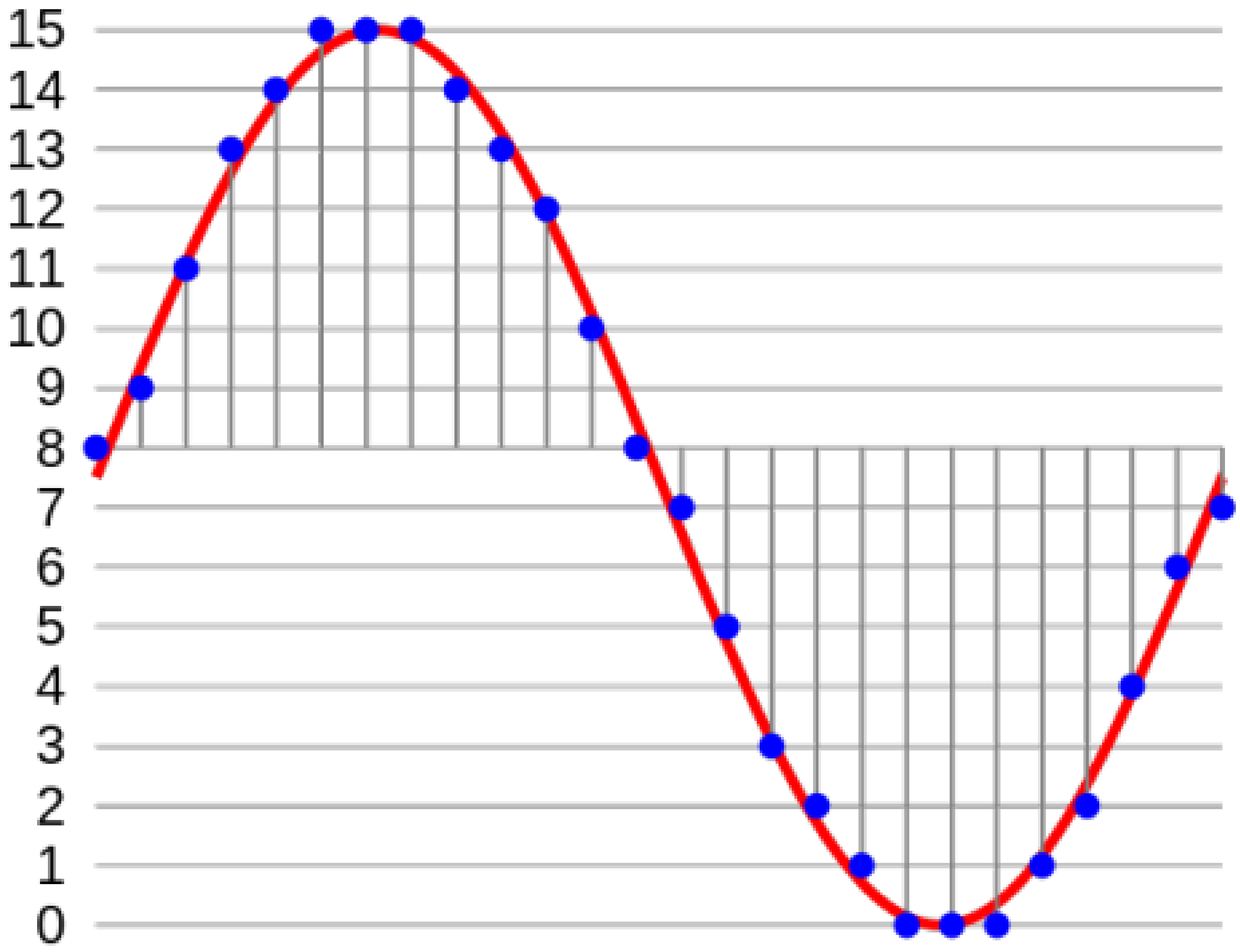
Digitization of Sound

- **Encoding** - Encoding converts the integer base-10 number to a base-2 that is a binary number. The output is a binary expression in which each bit is either a 1(pulse) or a 0(no pulse).



Digitization of Sound
Quantization

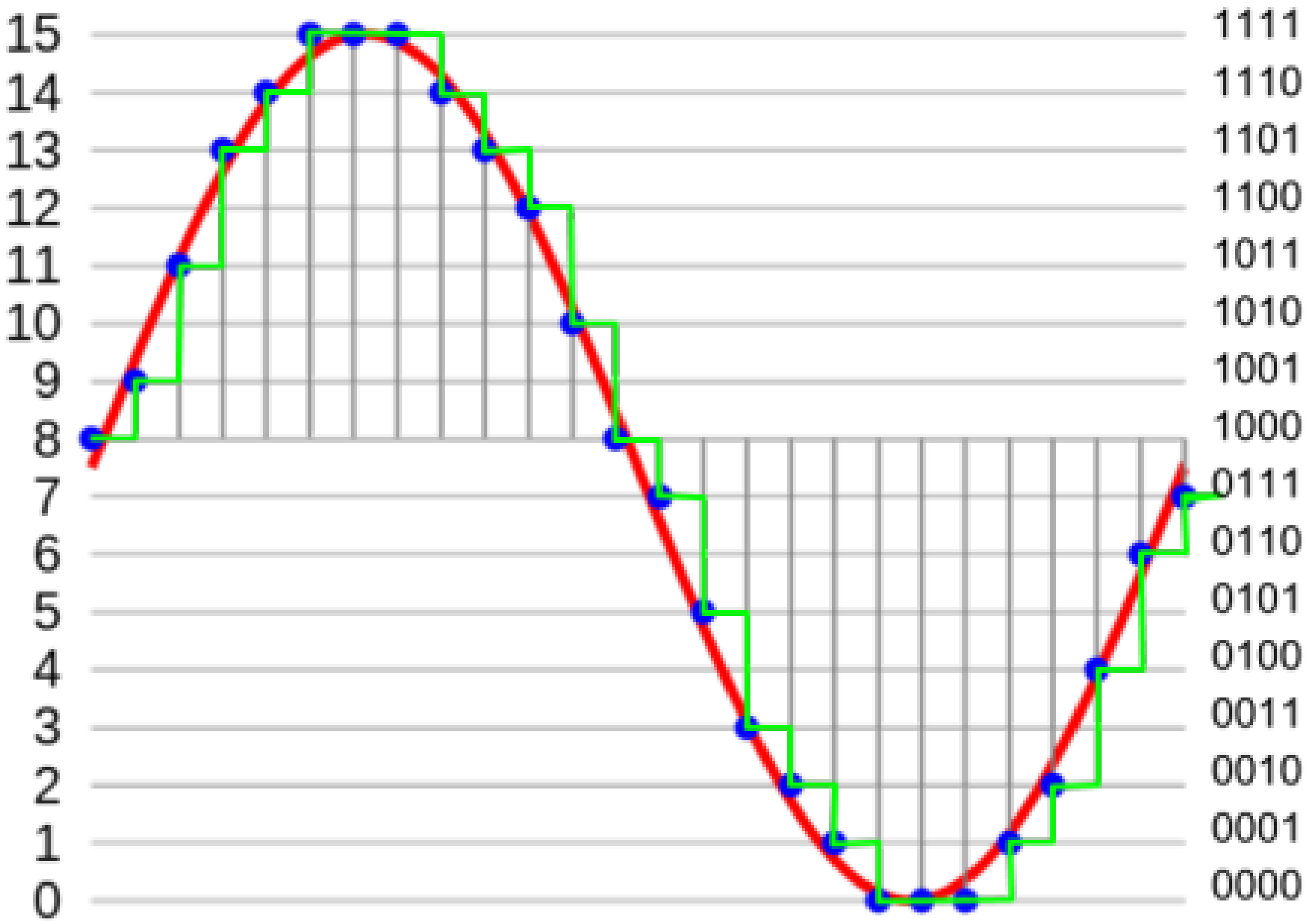
- Example:
- The red line is the wave that represents the audio data we are attempting to store.
- Notice that there are 16 possible values that we can assign to the wave (from 0 through to 15). These 16 values can be represented in 4 bits (since 4 bits allow us to store 16 unique values), so we say that the data we are storing would have a resolution of 4-bits.



Digitization of Sound
Quantization and Encoding

- This wave, would thus be stored using the following (decimal) values:
- 8, 9, 11, 13, 14, 15, 15, 15, 14, 13, 12, 10, 8, 7, 5, 3, 2, 1, 0, 0, 0, 1, 2, 4, 6, 7
- Or, in binary representation: 1000, 1001, 1011, 1101, 1110, 1111, 1111, 1111, 1110, 1101, 1100, 1010, 1000, 0111, 0101, 0011, 0010, 0001, 0000, 0000, 0000, 0001, 0010, 0100, 0110, 0111

If we then count up the number of bits used to store this wave, we would have 26 values x 4 bits = 104 bits = 13 bytes. If we were to then reproduce the wave, it would end up looking like the green line on the image below:



Digitization of Sound

- Using the same wave, what happens when you do each of the following?

    1. Double the sample rate, but keep the bit resolution at 4-bits

    2. Increase the bit resolution to 5-bits, but use the original sample rate

    3. Halve the sample rate and use a 5-bit resolution

Digitization of Sound

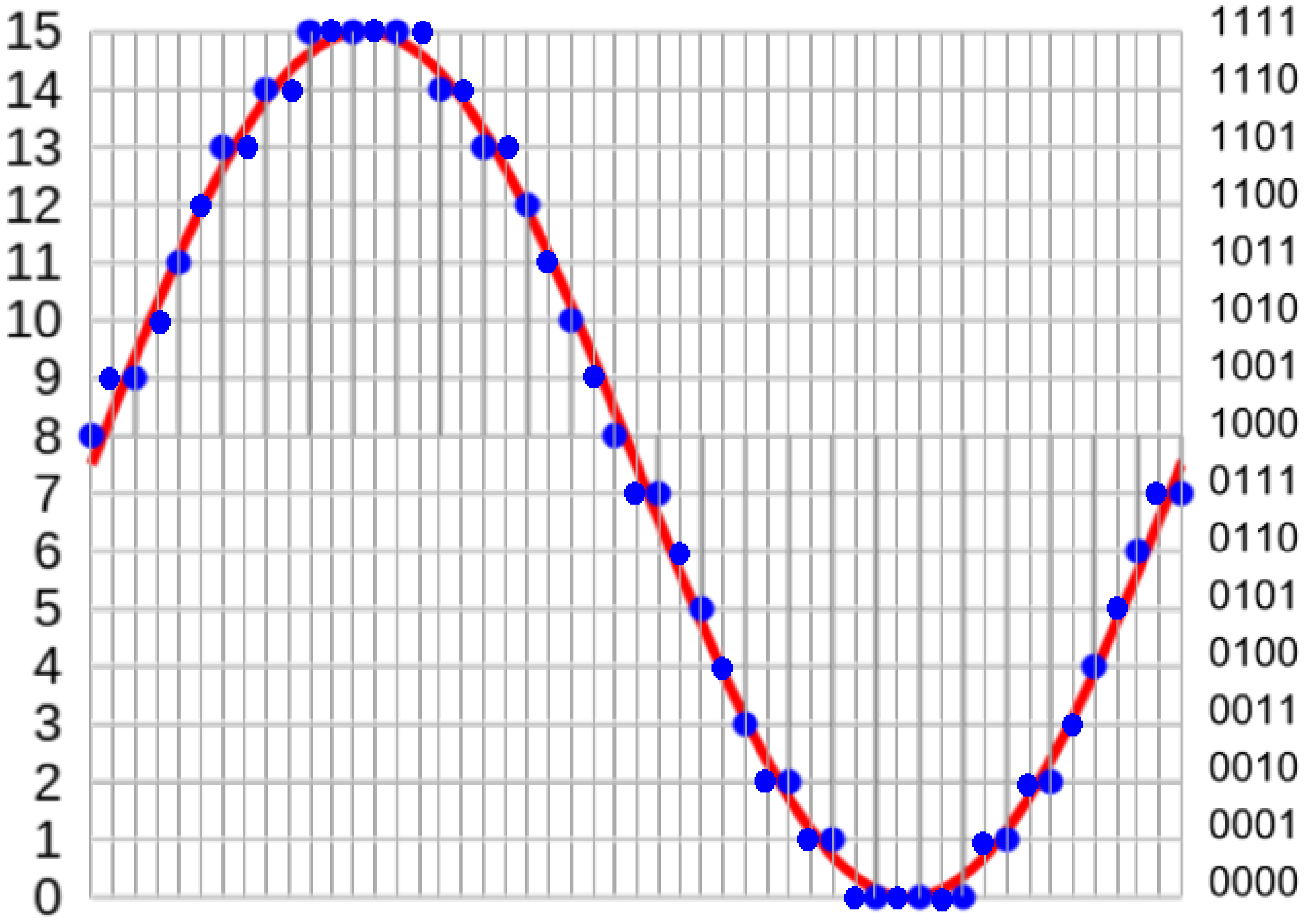
1. Double the sample rate, but keep the bit resolution at 4-bits

Stored Values:

Decimal

Binary

Bytes



Digitization of Sound

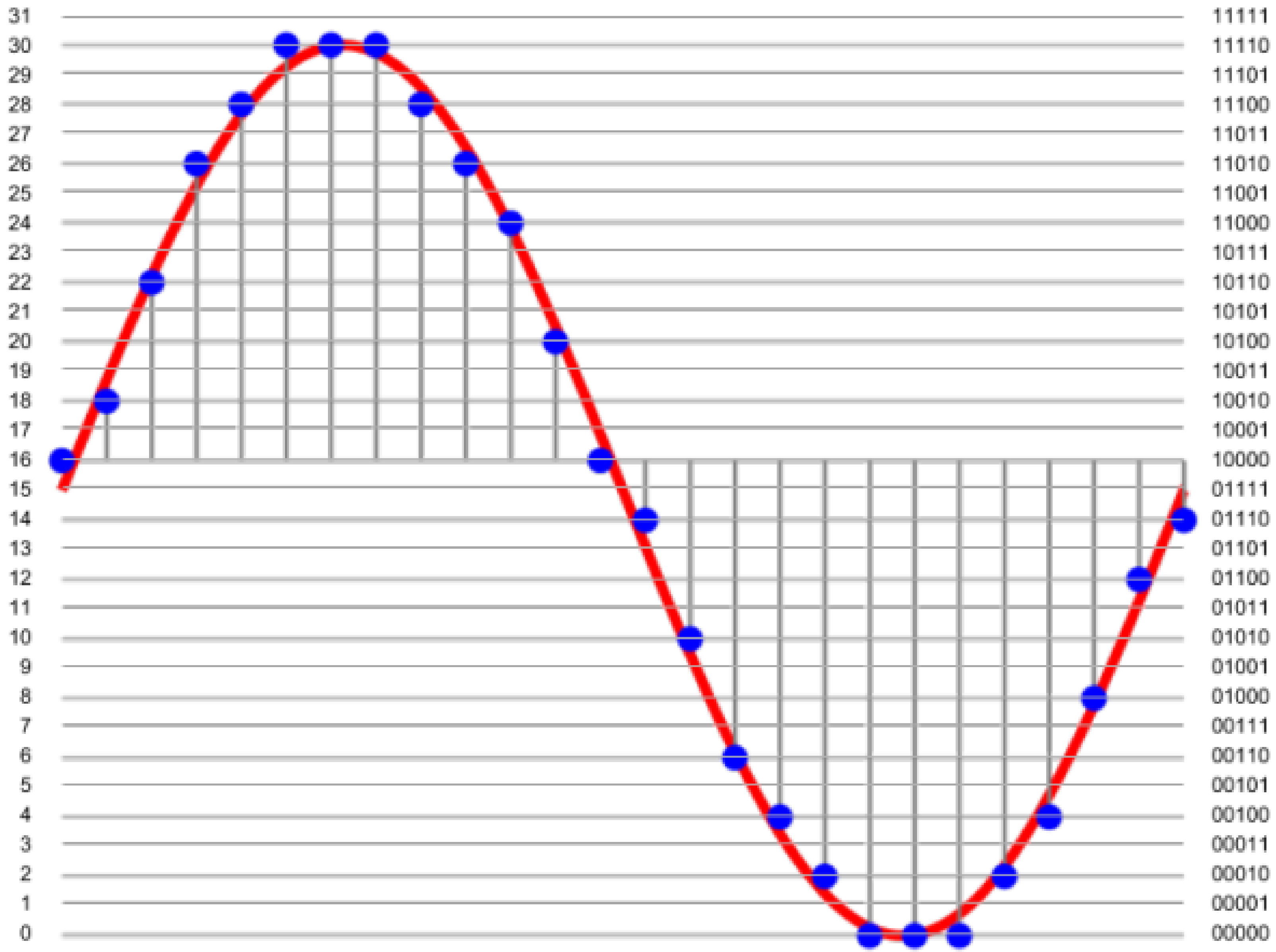
2. Increase the bit resolution to 5-bits, but use the original sample rate

Stored Values:

Decimal

Binary

Bytes



Digitization of Sound

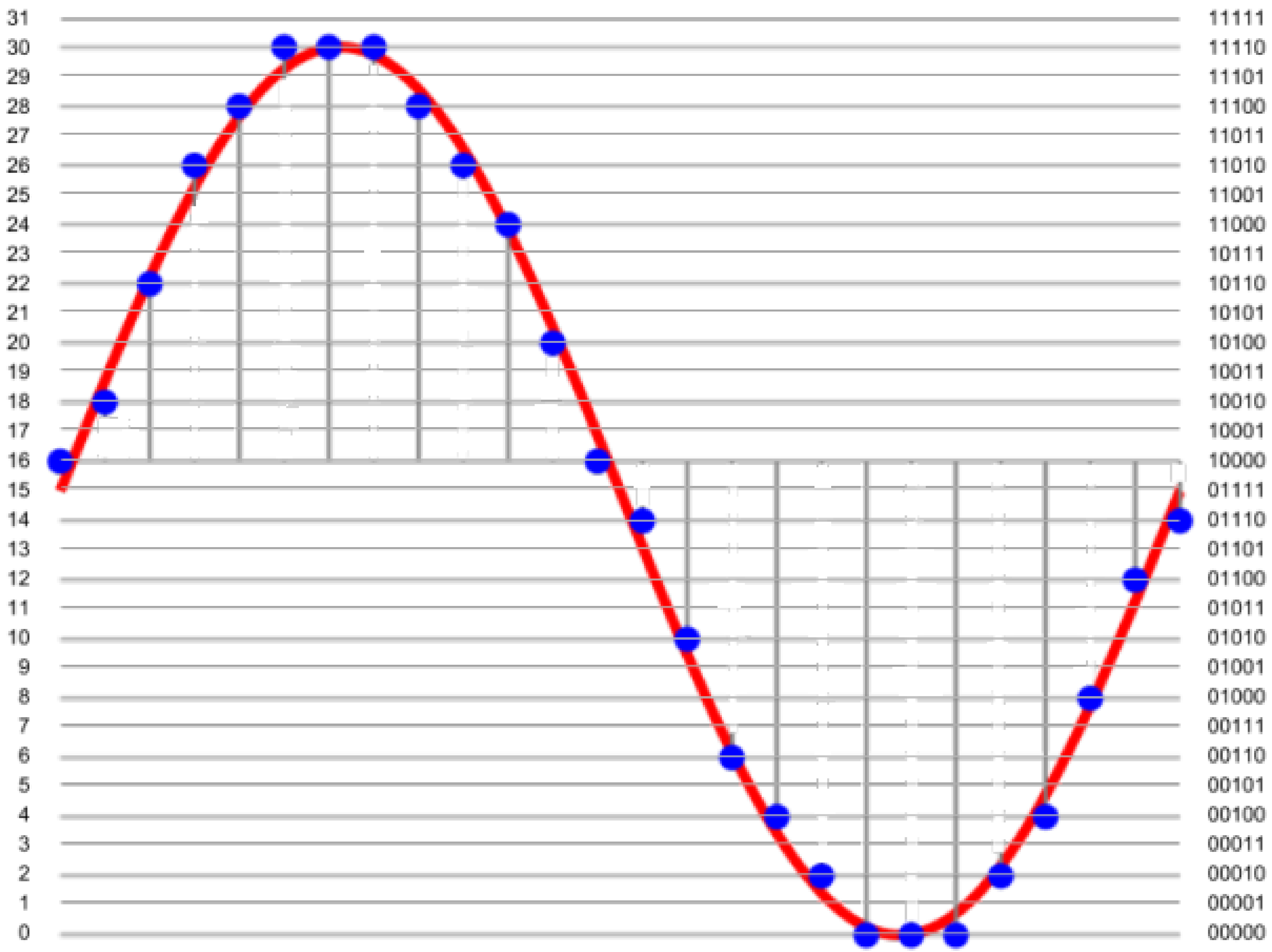
3. Half the sample rate and using 5-bit resolution

Stored Values:

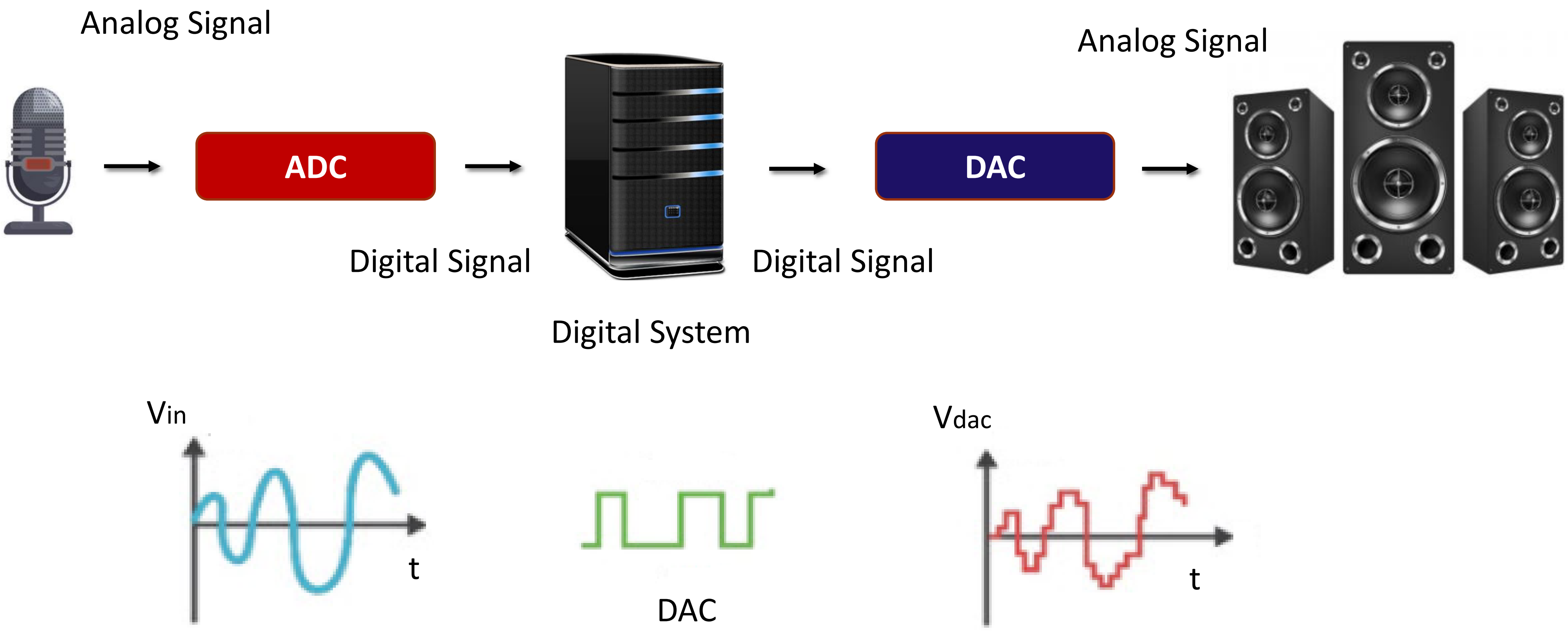
Decimal

Binary

Bytes

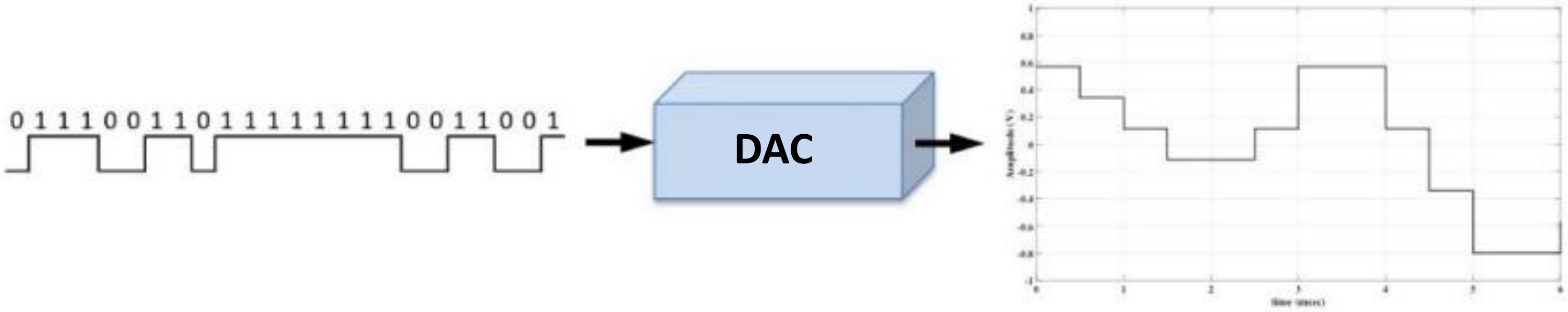


Digitization of Sound



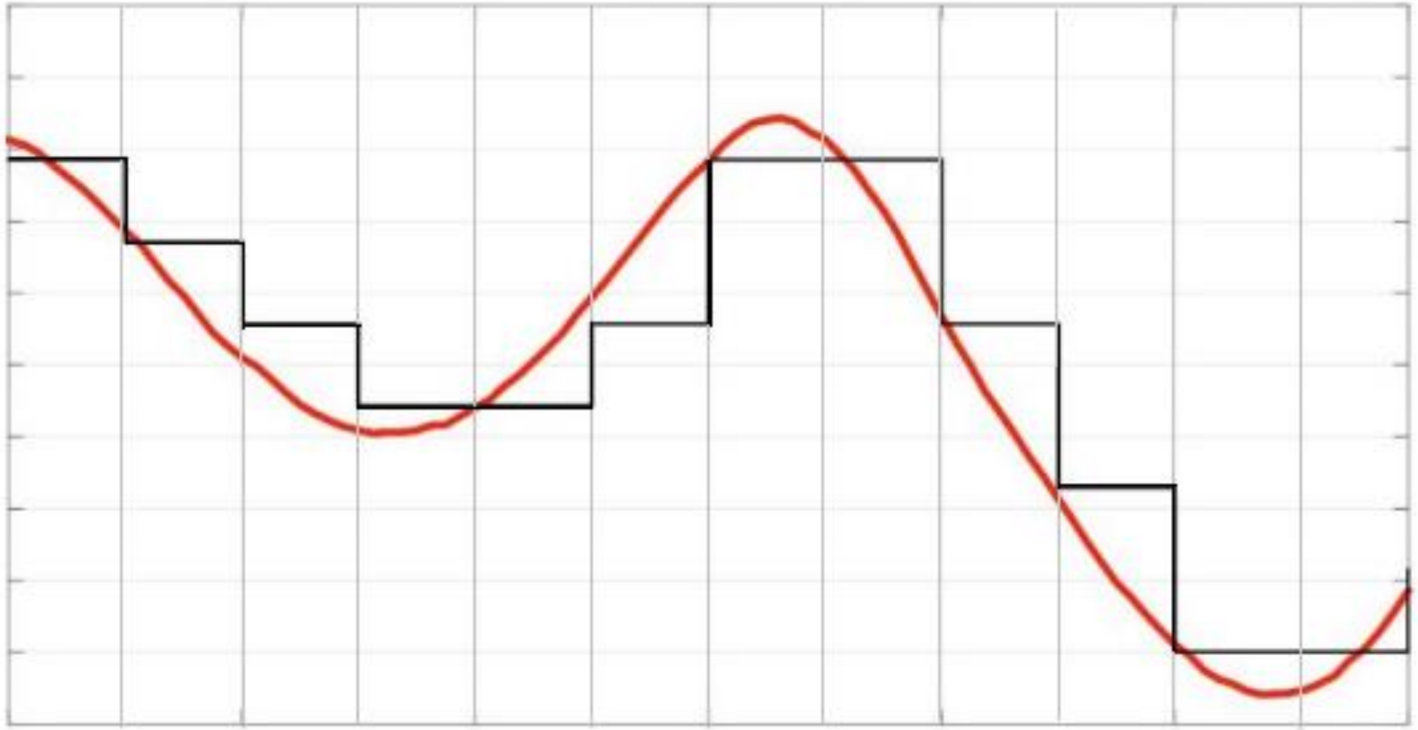
Conversion from Digital to Analog (DAC)

- How do we recover the analog information after it has been converted to digital?
- As seen in the previous slide, the receiver converts the n-bit digital stream back into an analog signal. This process is called digital-to-analog (DAC) conversion.
- It is very similar to being the reverse of the analog-to-digital conversion process.
- The analog signal is reconstructed by converting the n-bit digital stream into the appropriate quantization levels, and this voltage is "held" for one sample period, creating a stair-step type signal shown below.



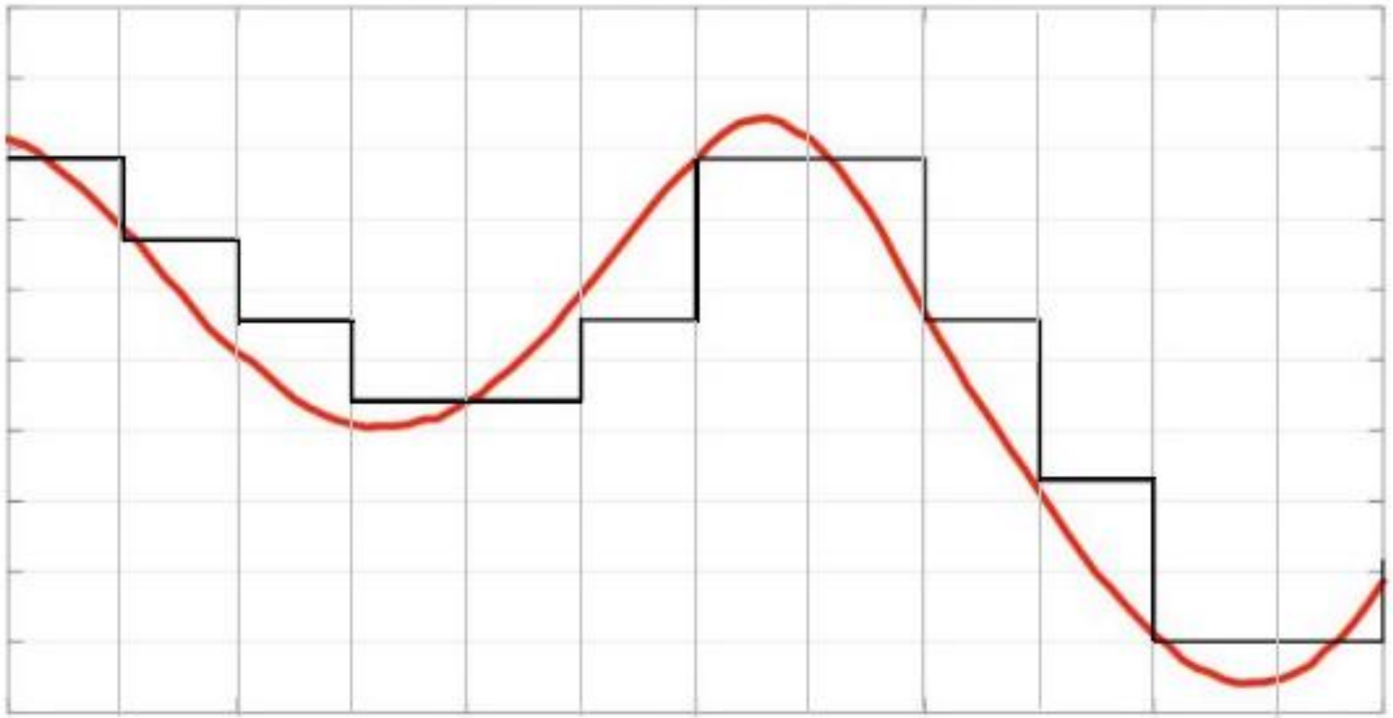
Conversion from Digital to Analog (DAC)

- So at this point we've regenerated the original signal.
- How does it compare with the original?
- The original analog signal is also shown in the continuous line along with all of the sample points that have been considered
- The thick black line shows the digital signal.
- **Is it close?**



Conversion from Digital to Analog (DAC)

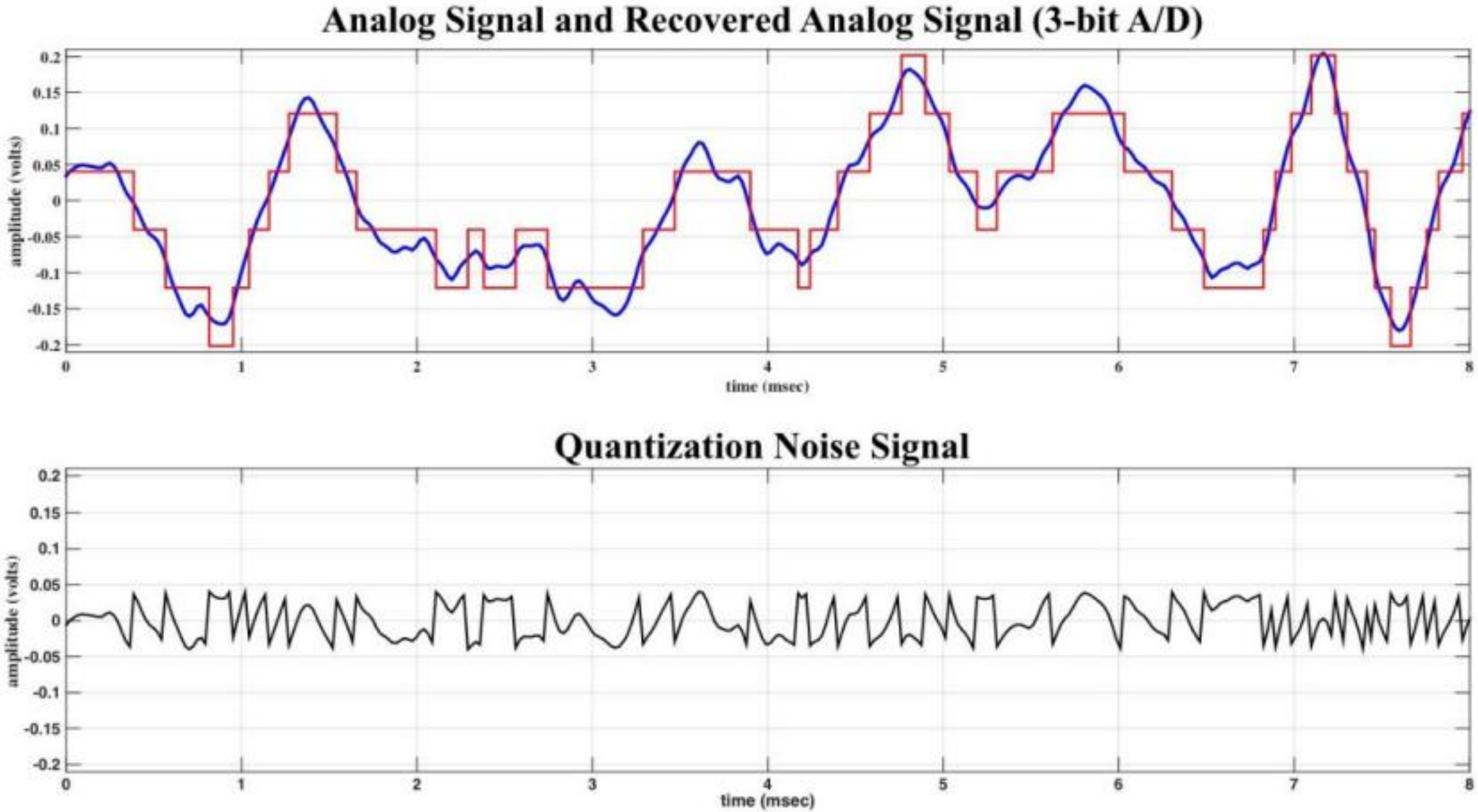
- It follows the same general shape.
- Even if we perform filtering to smooth out the reconstructed signal to remove its staircase appearance (which is typical) it will still not quite be the same as the original red signal.
- Why?
- Is that the best we can do?



Quantization Error (also called Quantization Noise)

- There is always error introduced with the A/D process.
- The error is the difference between the original analog signal and the reconstructed (stair-step) signal after A/D and D/A.
- The figure on the next slide is a portion of a music signal that has been quantized with 3 bits. The upper plot shows the original analog signal along with the recovered analog signal from the A/D process. The bottom plot is the quantization error, which is created by subtracting the recovered signal from the original analog signal at each instance of time.

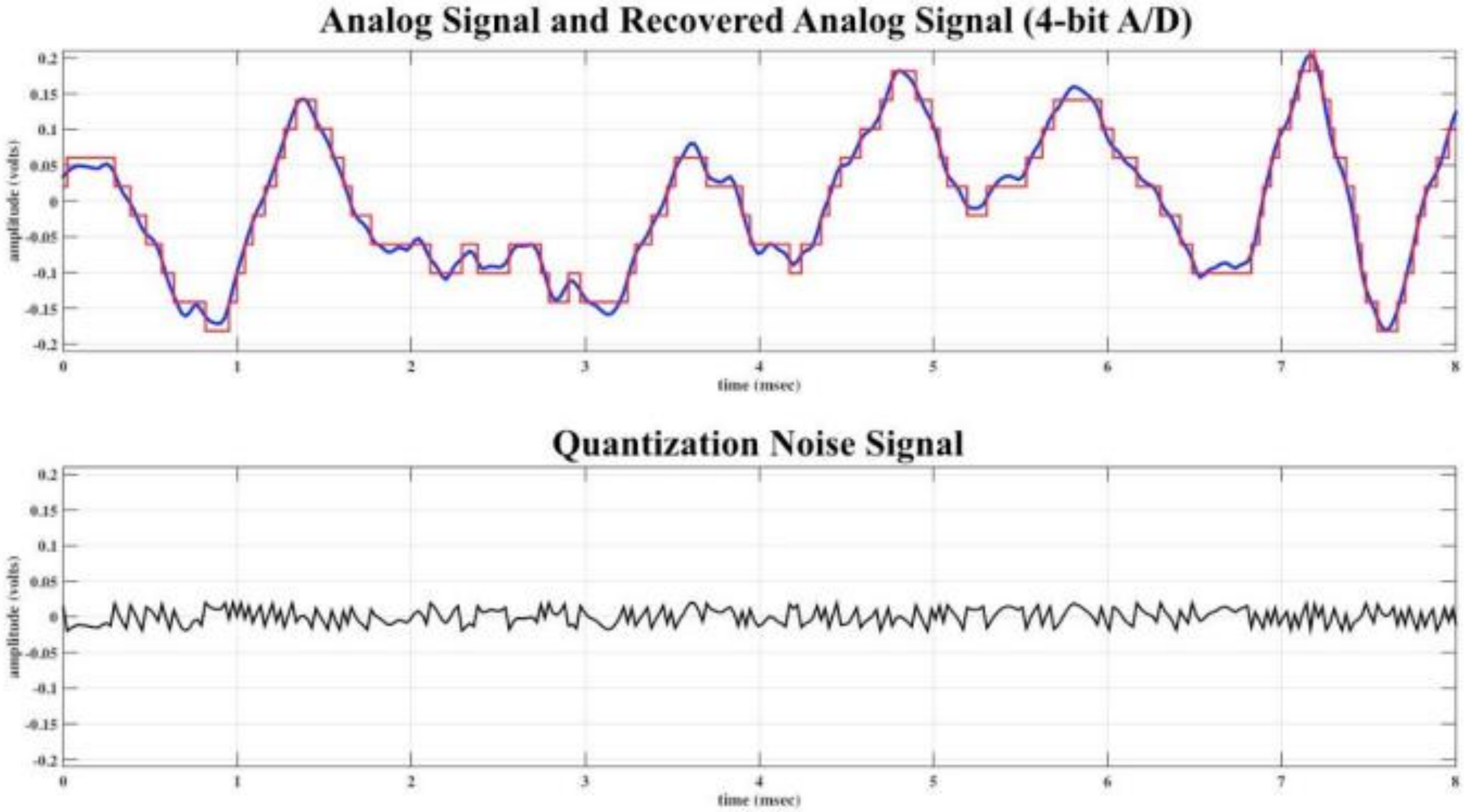
Quantization Error (also called Quantization Noise)



Quantization Error (also called Quantization Noise)

- The quantization error manifests as noise in the reconstructed analog signal.
- For digital audio signals (music or voice), it can sound like static, which is why it is also called quantization noise.
- The greater the quantization error, the louder the static, making it harder to hear the voice or music.
- So how do we reduce the quantization error and its associated noise?
- Quantization error can be reduced by increasing the number of bits n for each sample. This will make the quantization intervals smaller, reducing the difference between the analog sample values and the quantization levels. The figure on the next slide shows the same analog signal quantized with 4-bits per sample.

Quantization Error (also called Quantization Noise)



Quantization Error (also called Quantization Noise)

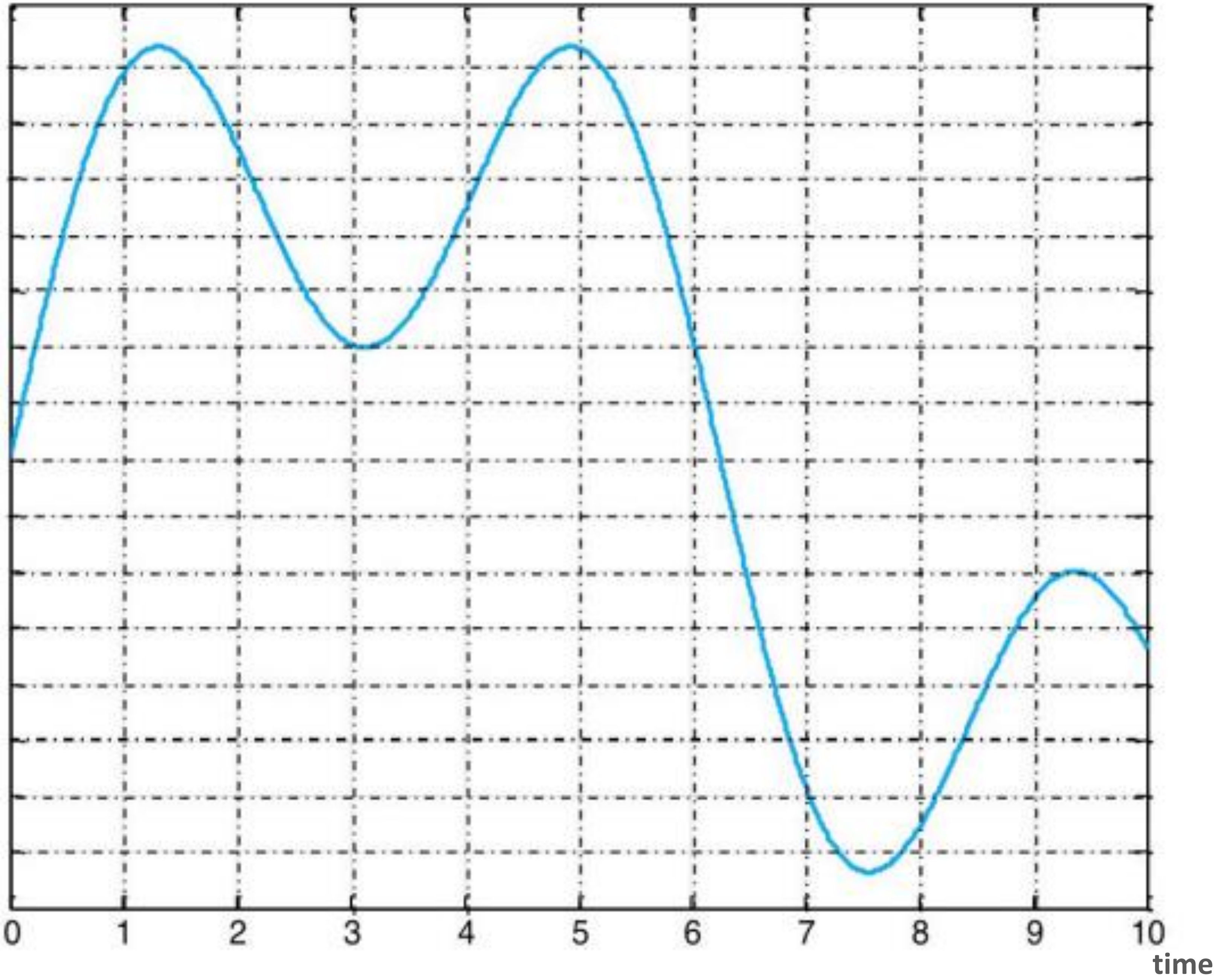
- *It is worth noting that increasing the sampling frequency will not reduce quantization noise, only increasing the number of quantization levels will do this.*

- We of course can’t use an infinite number of bits, so some quantization noise is always inevitable, but the nice thing about the human ear/brain - sticking with the example of audio signals - is that beyond a certain number of bits for each sample, the associated quantization noise becomes imperceptible. We just need enough bits to make the recovered signal “good enough”.

Task

Consider the following analog wave. The wave is to be quantized using 4 bits.

1. Circle the sample points (first sample is at time t = 0 sec).
2. Indicate the quantization intervals.
3. What is the stream of binary bits generated after the A/D conversion is complete?
4. What is the resulting size from this conversion?



MIDI

- MIDI is a protocol that enables computer, synthesizers, keyboards, and other musical or (even) multimedia devices to communicate with each other. It has been widely accepted and utilized by musicians and composers since its conception in the 1982/1983 time frame.
- MIDI compatible instruments and other MIDI devices are connected to each other via standard MIDI cables. MIDI files contain no audio data, only performance instructions for the computer's synthesizer or audio file databank.
- www.midisoft.com
- www.midi-classics.com