# Fundamentals of Audio

# Objectives

- Understanding sound wave and digitization
- Adding sound to multimedia project.

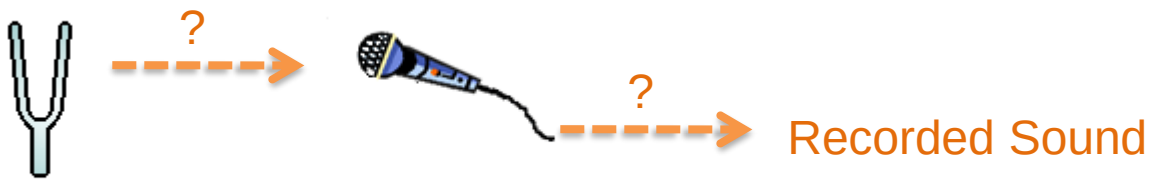
Fundamentals of Digital Audio

# THE NATURE OF SOUND WAVE

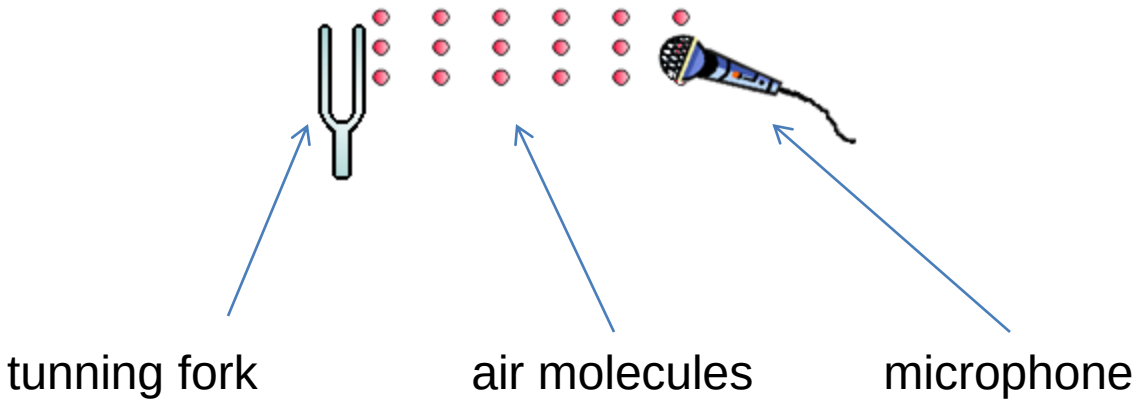
# Sound

- A wave that is generated by vibrating objects in a medium such as air
- Sound waves vary in sound pressure level (amplitude) and in frequency or pitch
- ‘Acoustics’ is the branch of physics that studies sound
- Examples of vibrating objects:
  - vocal cords of a person
  - guitar strings
  - tuning fork

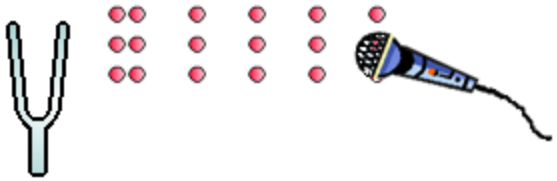
So how is vibration turned into sound we can hear or record with a microphone?



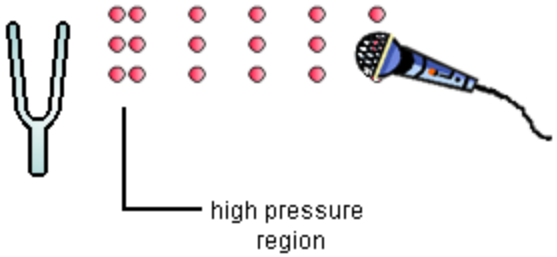
An illustration of how the propagating sound wave formed by changes of the air pressure reaches the microphone



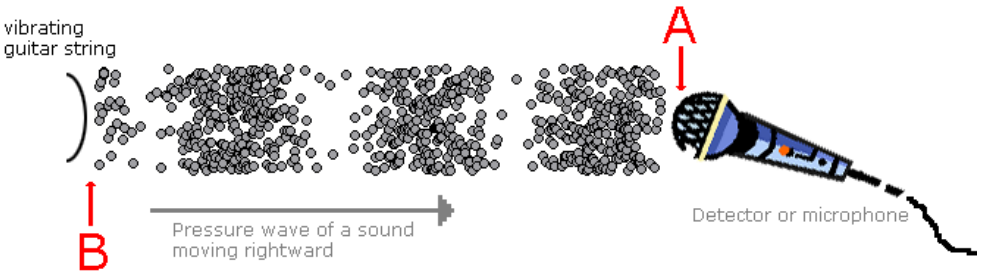
An illustration of how the propagating sound wave formed by changes of the air pressure reaches the microphone



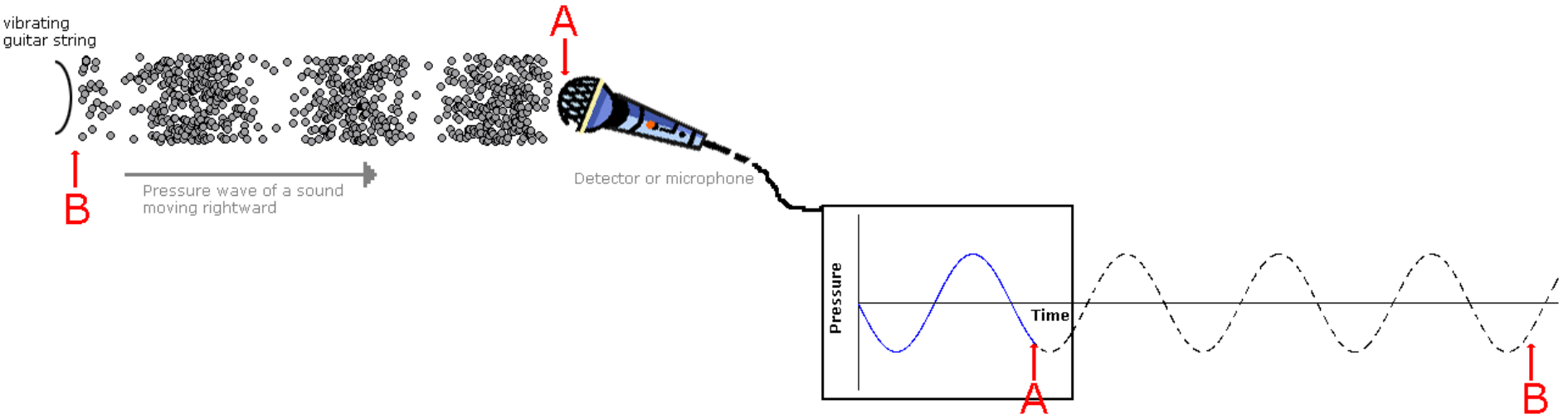
# Let's step through the process slowly



The changes of pressure in the propagating sound wave reaching the recorder are captured as changes of electrical signals over time.



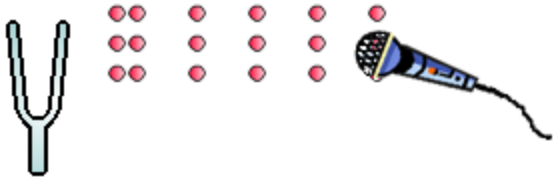
The sound wave can be represented graphically with the changes in air pressure or electrical signals plotted over time—a *waveform*.



# Be careful...

- NOT to interpret sound as a wave that has crests and troughs

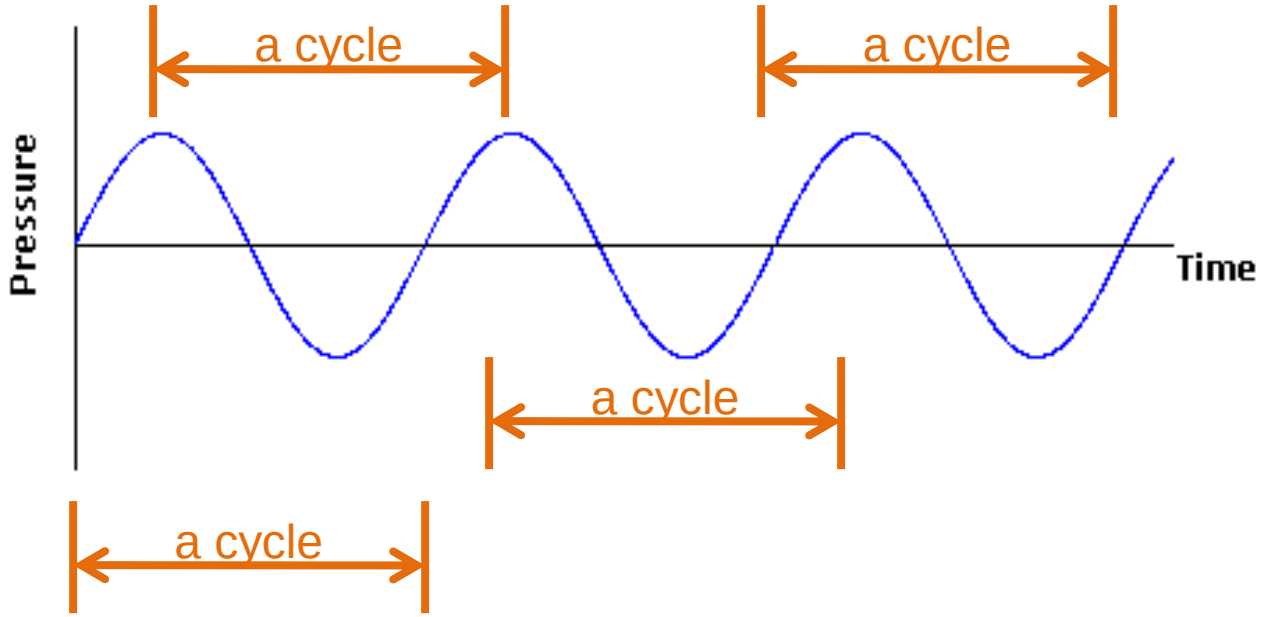
- NOT to interpret the waveform as a representation of the sound wave in space
  - i.e. air molecules are not going up and down



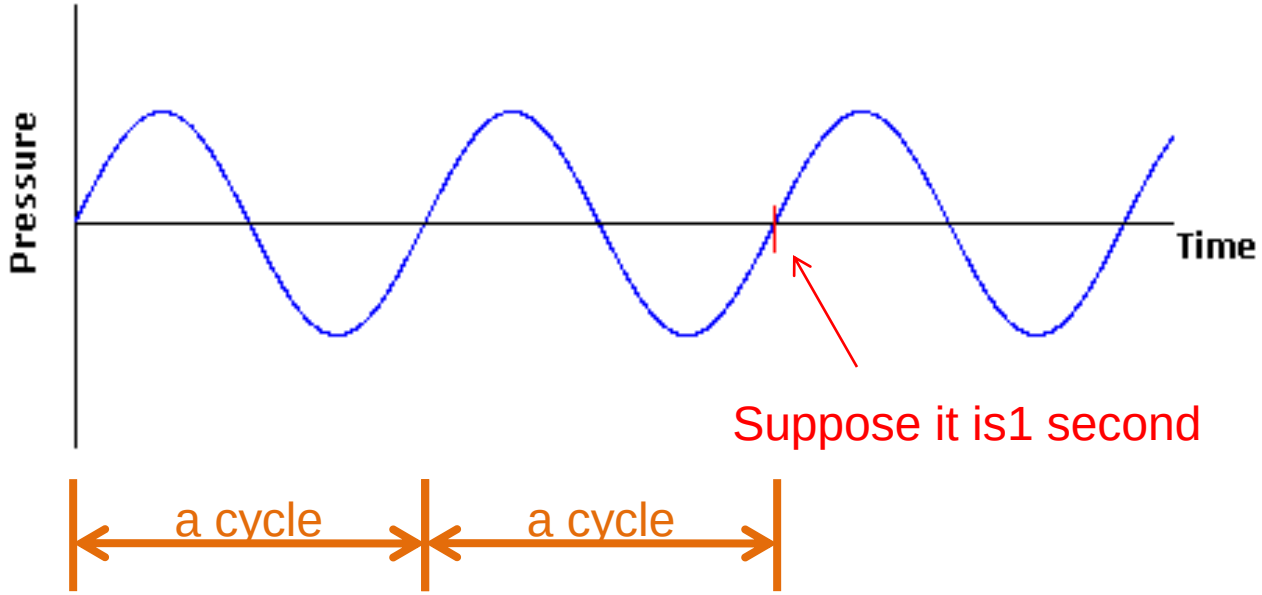
# Frequency of Sound Wave

- Refers to the number of complete back-and-forth cycles of vibrational motion of the medium particles per unit of time
- Unit for frequency: Hz (Hertz)
- 1 Hz = 1 cycle/second

# A Cycle

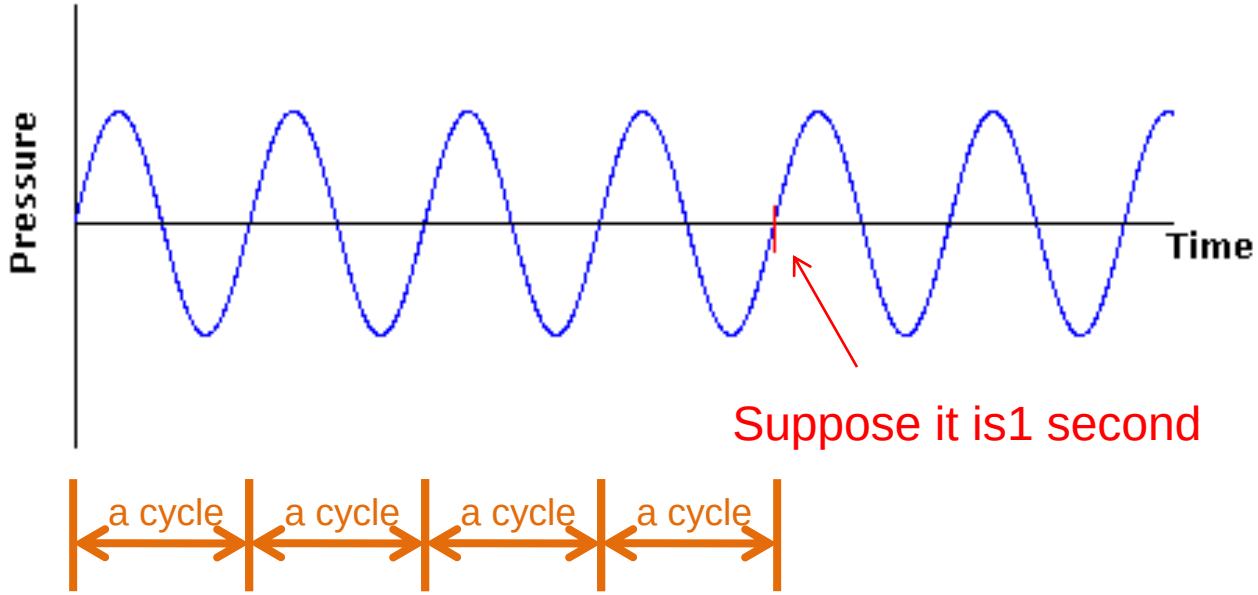


# Frequency




Frequency = 2 Hz (i.e., 2 cycles/second)

# Frequency




Frequency = 4 Hz (i.e., 4 cycles/second)
Higher frequency than the previous waveform.

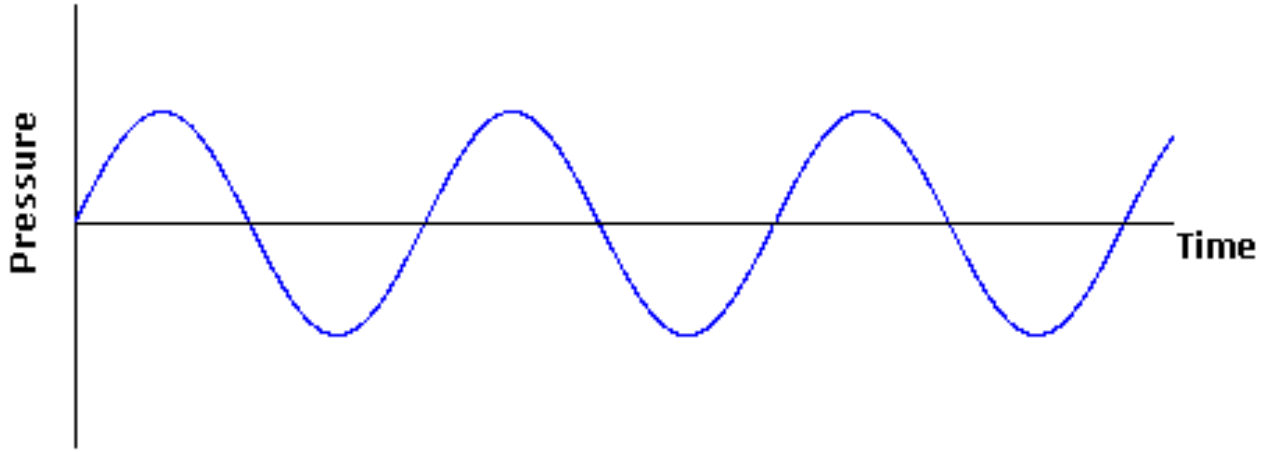
# Pitch of Sound

- Sound frequency
- Higher frequency: higher pitch
- human ear can hear sound ranging from 20 Hz to 20,000 Hz

# Sound Intensity vs. Loudness

- Sound intensity:
  - an objective measurement
  - can be measured with auditory devices
  - in *decibels* (dB)
    - A dB measurement is actually the ratio b/w a chosen reference point on a logarithmic scale and the level that is actually experienced.
- Loudness:
  - a subjective perception
  - measured by human listeners
  - human ears have different sensitivity to different sound frequency
  - in general, higher sound intensity means louder sound

# A Simple Sine Wave Waveform



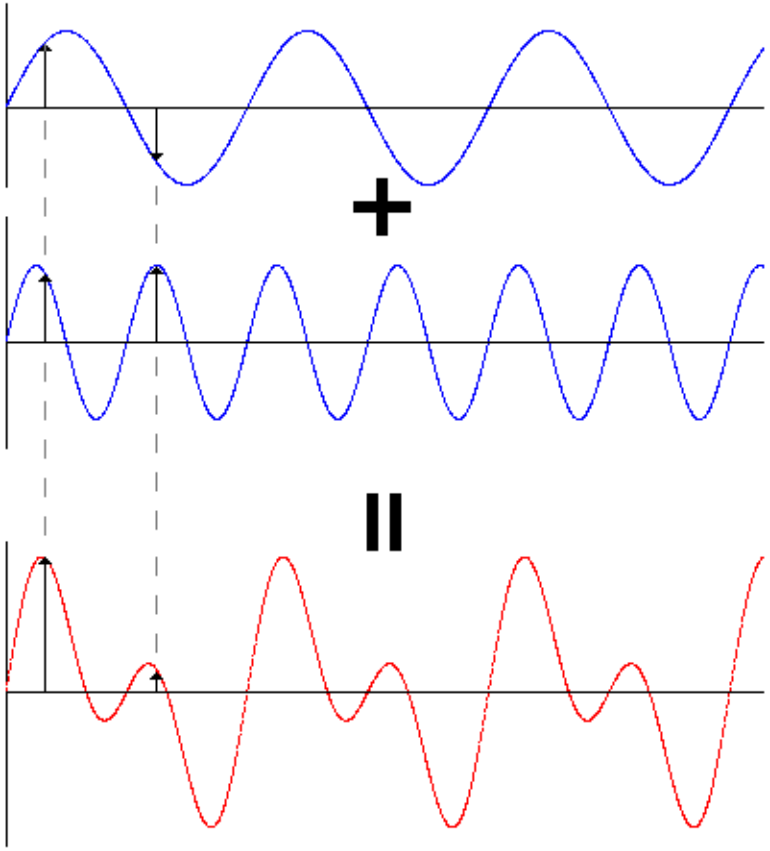

A sinlge sine wave waveform
A single tone

# Adding Sound Waves

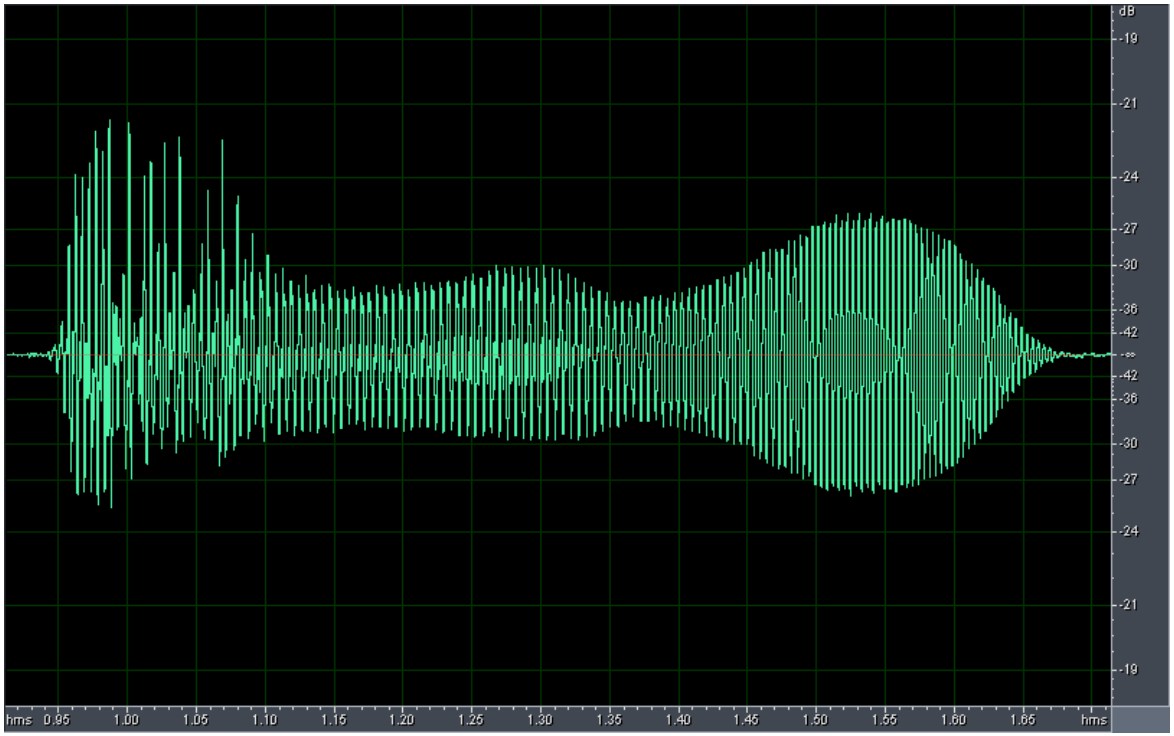
A sinlge sine wave waveform
A single tone

A second sinlge sine wave waveform
A second single tone

A more complex waveform
A more complex sound

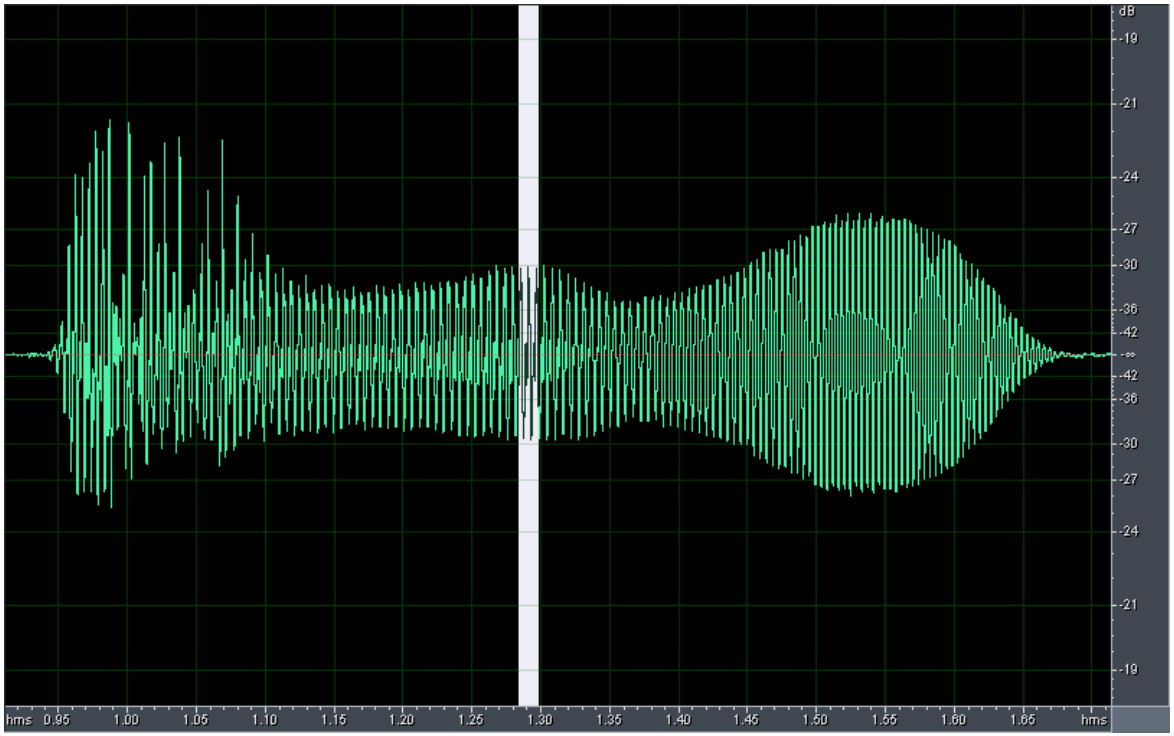


# Waveform Example



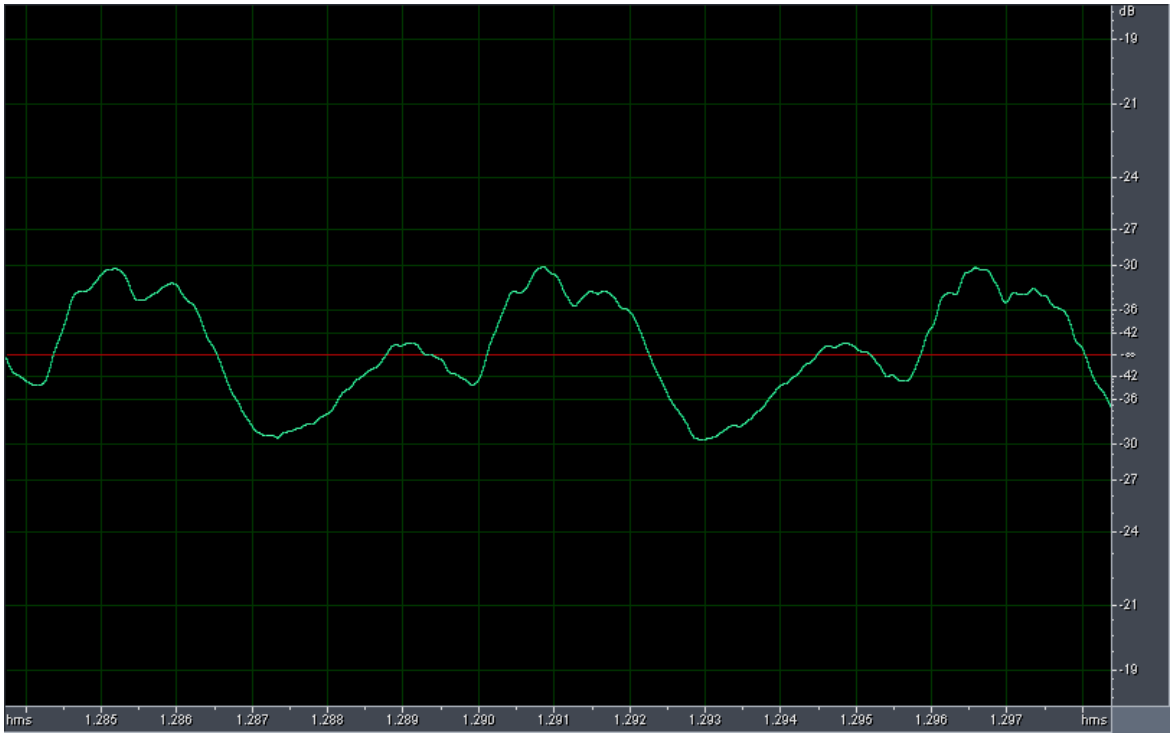

A waveform of the spoken word "one"

# Waveform Example




Let's zoom in to take a closer look

# Waveform Example

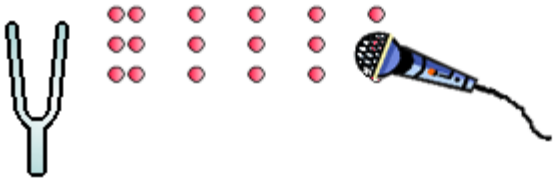



A closer look

# Review Question

Which of the following gives a correct illustration of the sound wave propagation?

A. 

B. 

C. Both are correct; it depends on the type of sound.

# Review Question

A sound with higher ___ is perceived to have a higher pitch.

A. decibels
B. frequency
C. fidelity
D. sampling rate
E. bit depth

# Review Question

The unit used for measuing ___ is Hertz (Hz).

A. amplitude
B. frequency
C. dynamic range
D. sampling rate
E. bit depth

# Review Question

A waveform is a graphical representation of the ___ fluctuations of a sound wave.

A. pressure-time
B. space-time
C. pressure-space

# Review Question

The horizontal axis of a waveform is ___.

A. pressure
B. distance
C. time

# Review Question

The vertical axis of a waveform is ___.

A. pressure
B. distance
C. time

# Review Question

True/False: Zero decibel is when there is absence of sound or no sound wave.

# Review Question

The ___ of a sound relates to the sound intensity or loudness.

A. amplitude
B. frequency
C. sampling rate
D. bit depth
E. dynamic range

Fundamentals of Digital Audio

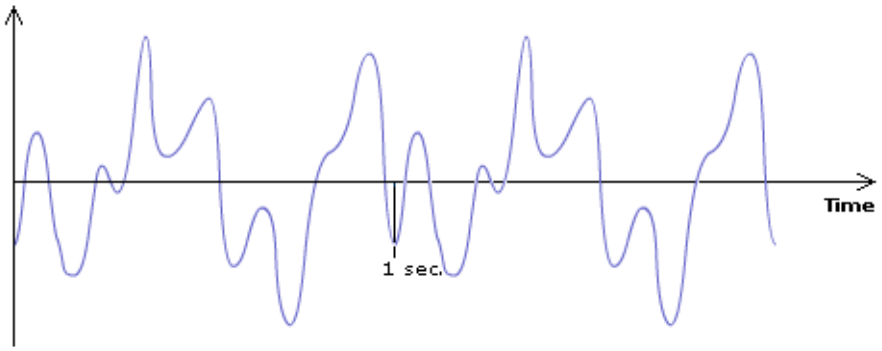
# DIGITIZING SOUND

# In this lecture, you will find answers to these questions:

- What does digitizing audio mean?
- How is sound sampled and quantized in the digitization process?
- How are sampling rate, bit depth, and resolution related to sampling and quantizing?
- How do the choices of the sampling rate and bit depth affect the sound fidelity and details?

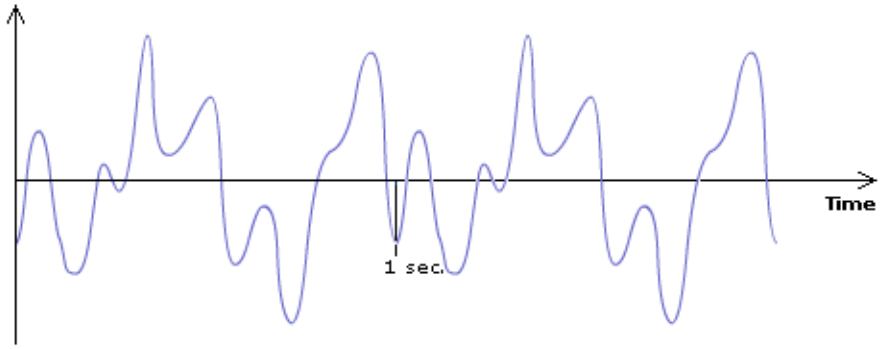
# Suppose we want to digitize the sound wave:

- When the sound is needed to be used in any computer application, we need to converts the air vibrations of sound into an electrical signal, which is called digital signal- a stream of 0's and 1's. The process of conversing analog signals to digital signals is called digitizing.



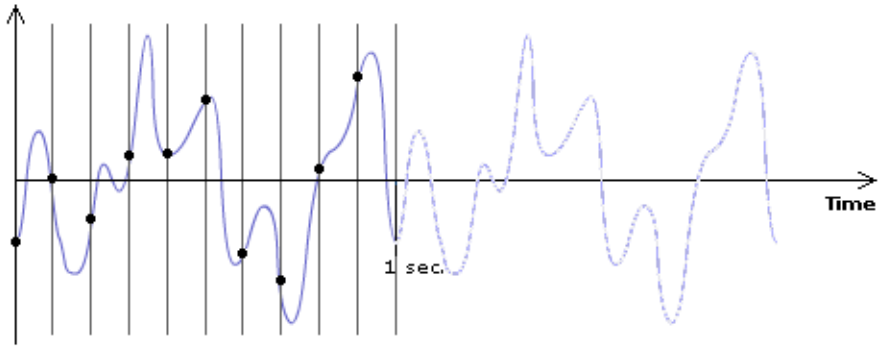
# Step 1. Sampling

The sound wave is sampled at a specific rate into discrete samples of amplitude values.



# Step 1. Sampling

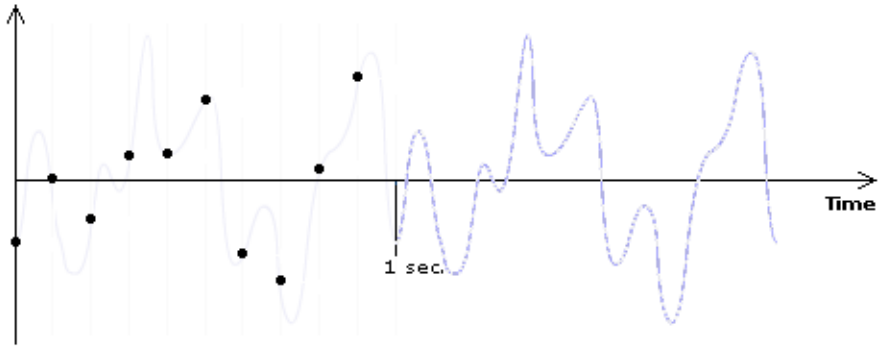
The sound wave is sampled at a specific rate into discrete samples of amplitude values.




Suppose we sample the waveform 10 times a second, i.e., sampleing rate = 10 Hz.

# Step 1. Sampling

The sound wave is sampled at a specific rate into discrete samples of amplitude values.

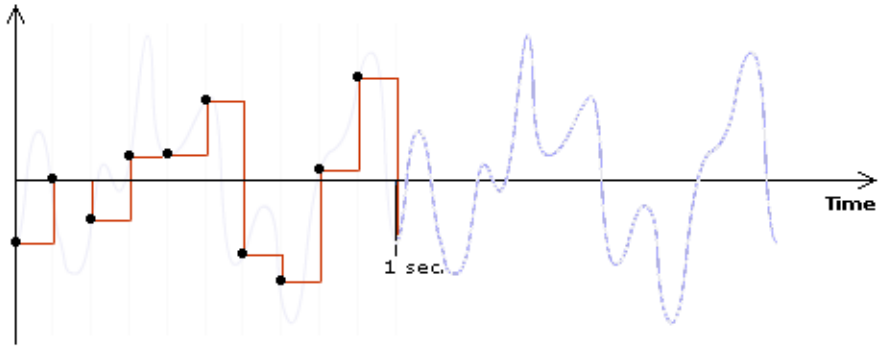



Suppose we sample the waveform 10 times a second, i.e., sampleing rate = 10 Hz.

We get 10 samples per second.

# Step 1. Sampling

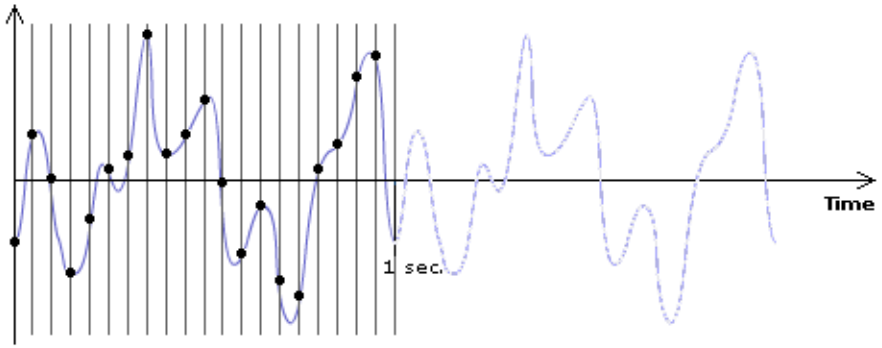
The sound wave is sampled at a specific rate into discrete samples of amplitude values.




Reconstructing the waveform using the discrete sample points.

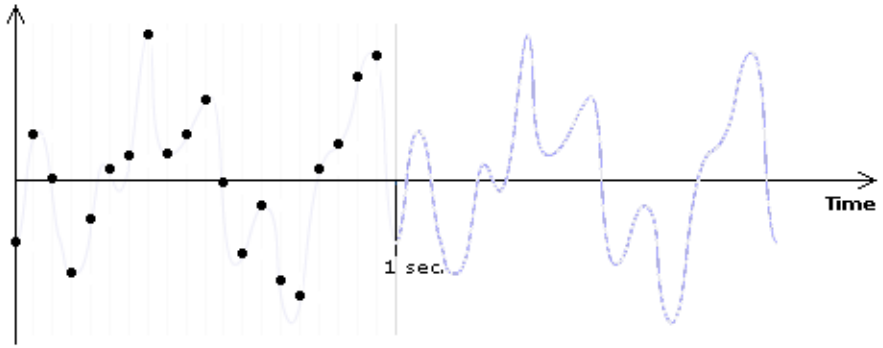
# Step 1. Sampling

What if we sample 20 times a second, i.e., sampling rate = 20 Hz?



# Step 1. Sampling

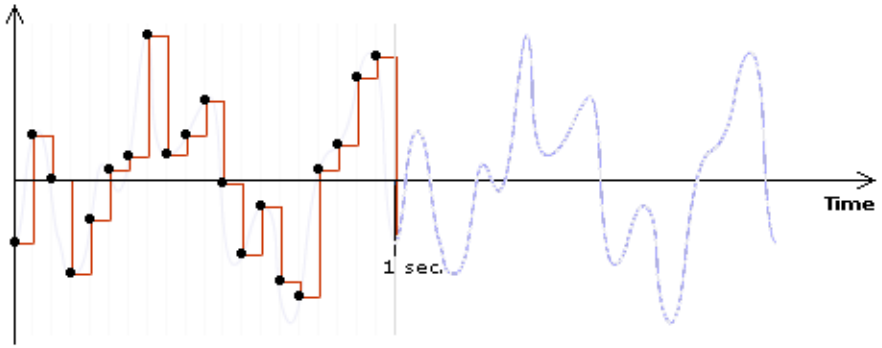
What if we sample 20 times a second, i.e., sampling rate = 20 Hz?




We get 20 samples per second.

# Step 1. Sampling

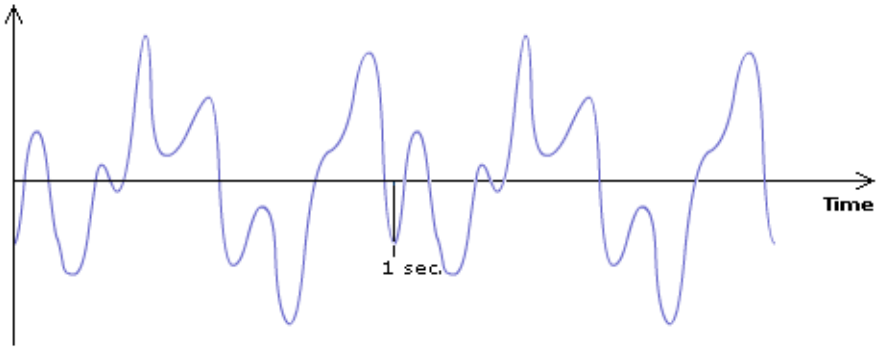
What if we sample 20 times a second, i.e., sampling rate = 20 Hz?



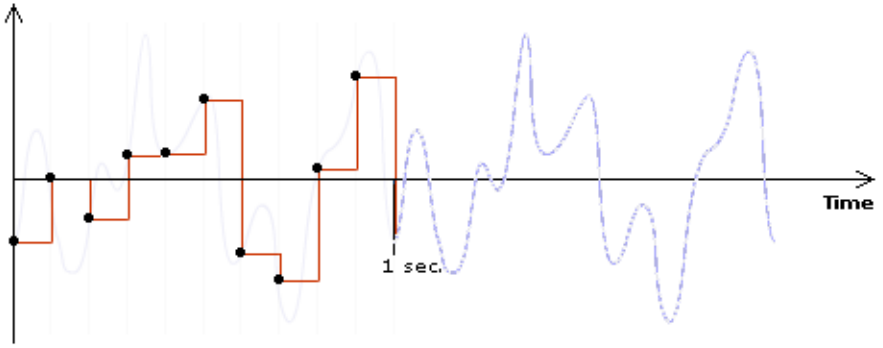

Reconstructing the waveform using the discrete sample points.

# Effects of Sampling Rate

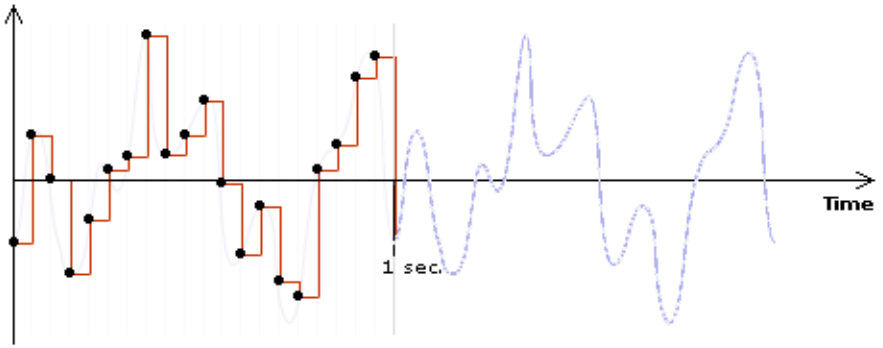
original waveform




sampling rate = 10 Hz




sampling rate = 20 Hz



# Effects of Sampling Rate

Higher sampling rate:

- The reconstructed wave looks closer to the original wave

- More sample points, and thus larger file size

# Sampling Rate Examples

- 11,025 Hz AM Radio Quality/Speech
- 22,050 Hz Near FM Radio Quality (high-end multimedia)
- 44,100 Hz CD Quality
- 48,000 Hz DAT (digital audio tape) Quality
- 96,000 Hz DVD-Audio Quality
- 192,000 Hz DVD-Audio Quality

# Sampling Rate vs. Sound Frequency

- Both uses the unit Hz

BUT:

- sampling rate ≠ sound frequency
- Sample rate: a setting in the digitization process
- Sound frequency: NOT a setting in the digitization process
- Sound frequency: the pitch characteristic of sound
- Higher sampling rate: NOT the pitch characteristic of sound

# Step 2. Quantization

- Each of the discrete samples of amplitude values obtained from the sampling step are mapped and rounded to the nearest value on a scale of discrete levels.

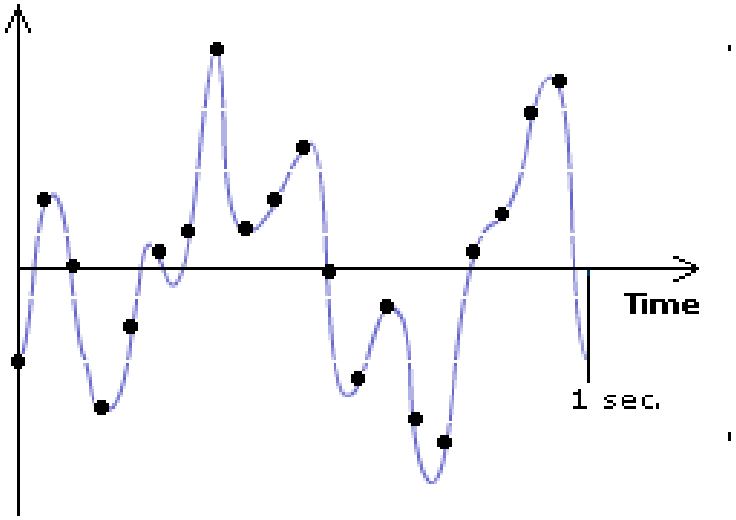
- The number of levels in the scale is expressed in *bit depth*--the power of 2.

- An 8-bit audio allows $2^8 = 256$ possible levels in the scale

- CD-quality audio is 16-bit (i.e., $2^{16} = 65{,}536$ possible levels)

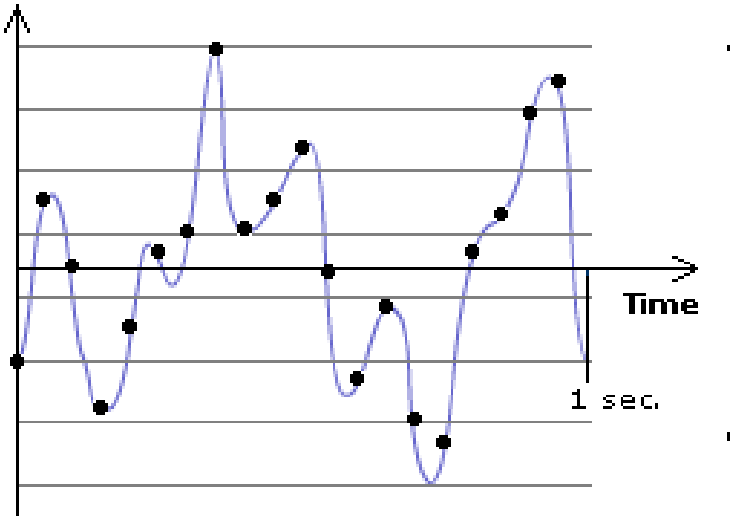
# Step 2. Quantization

Suppose we are quantizing the samples using 3 bits (i.e. $2^3 = 8$ levels).



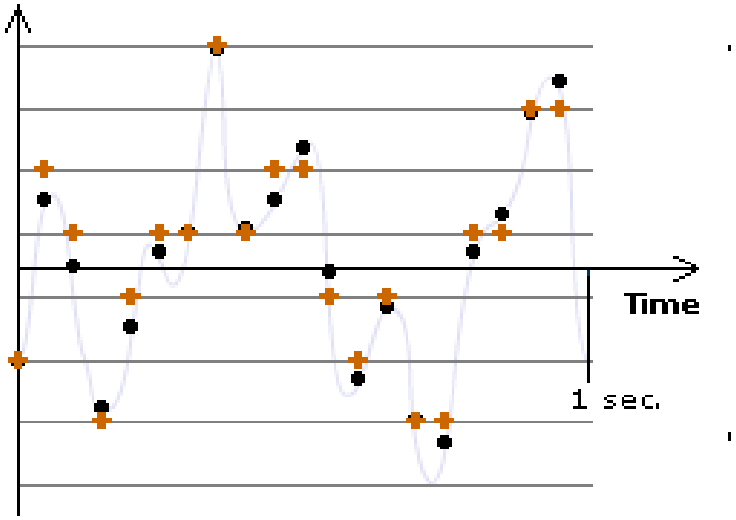
# Step 2. Quantization

Suppose we are quantizing the samples using 3 bits (i.e. $2^3 = 8$ levels).



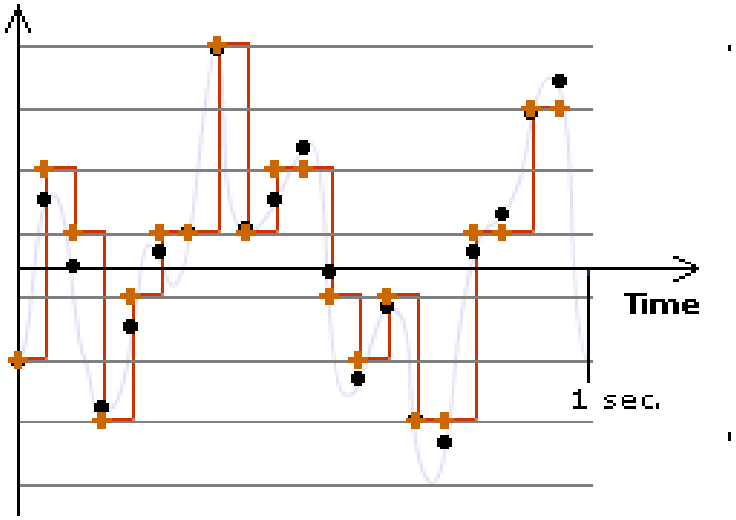
# Step 2. Quantization

Now, round each sample to the nearest level.



# Step 2. Quantization

Now, reconstruct the waveform using the quantized samples.



# Effects of Quantization

- Data with different original amplitudes may be quantized onto the same level
  $\Rightarrow$ loss of subtle differences of samples

- With lower bit depth, samples with larger differences may also be quantized onto the same level.

# Bit Depth

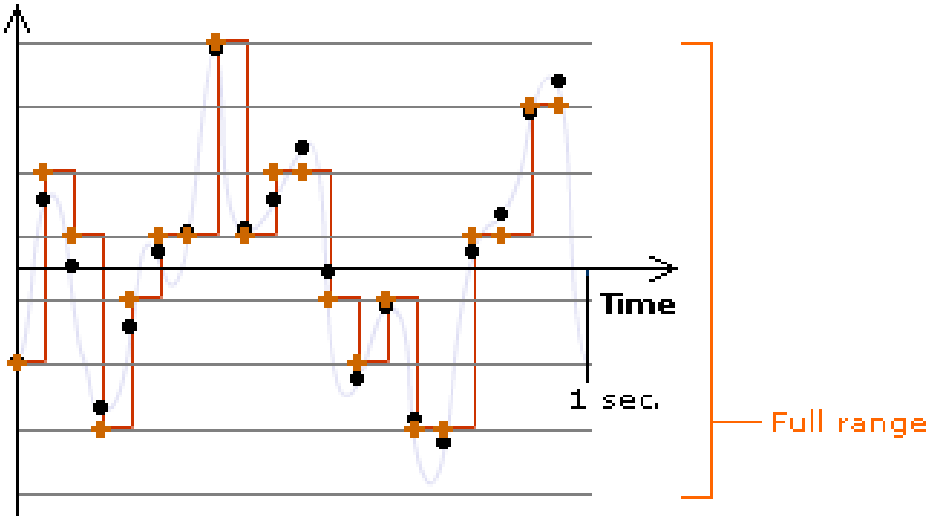
- Bit depth of a digital audio is also referred to as *resolution*.

- For digital audio, higher resolution means higher bit depth.

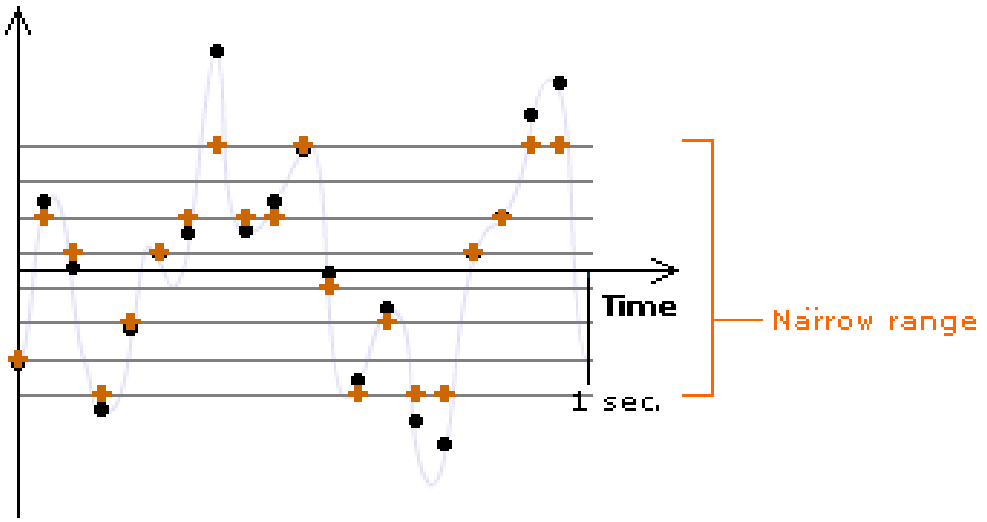
# Dynamic Range

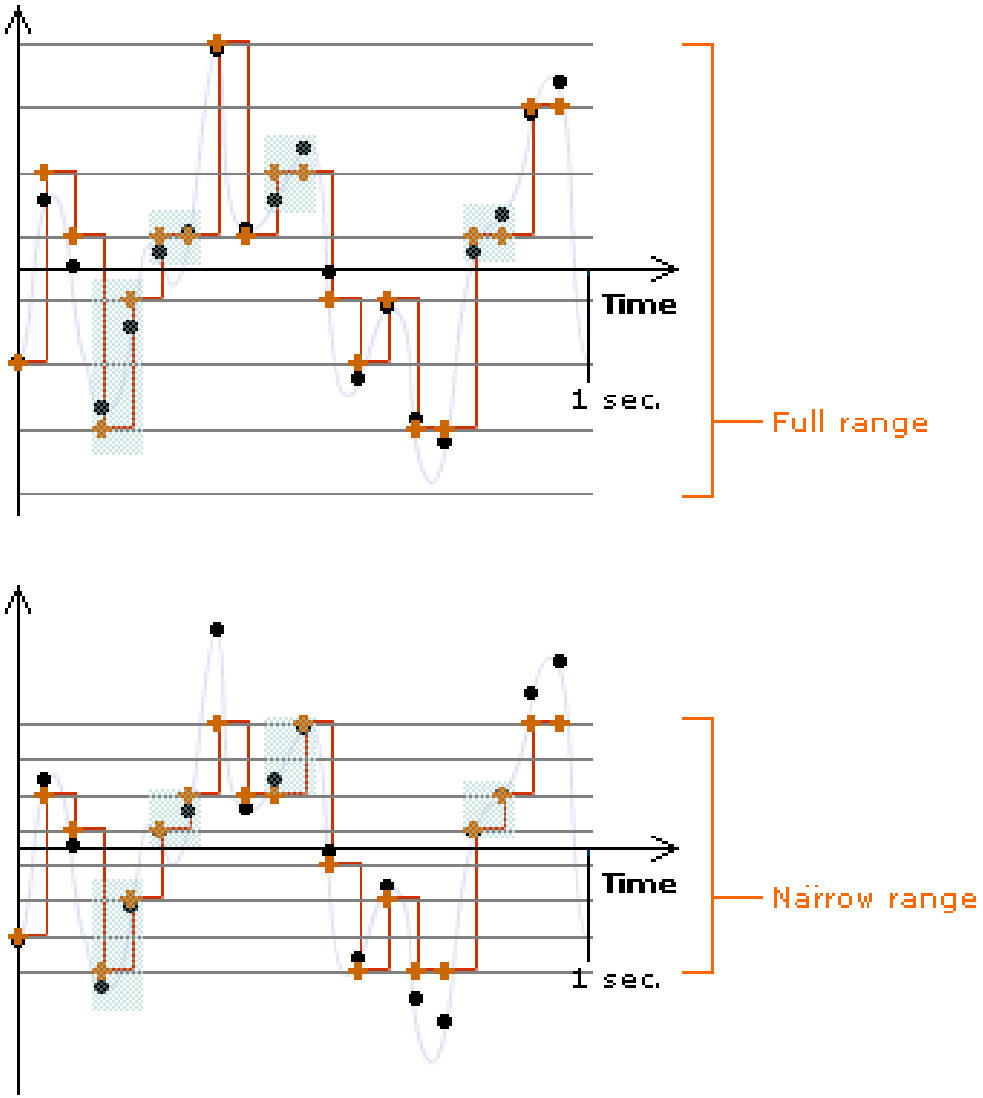
- The range of the scale, from the lowest to highest possible quantization values

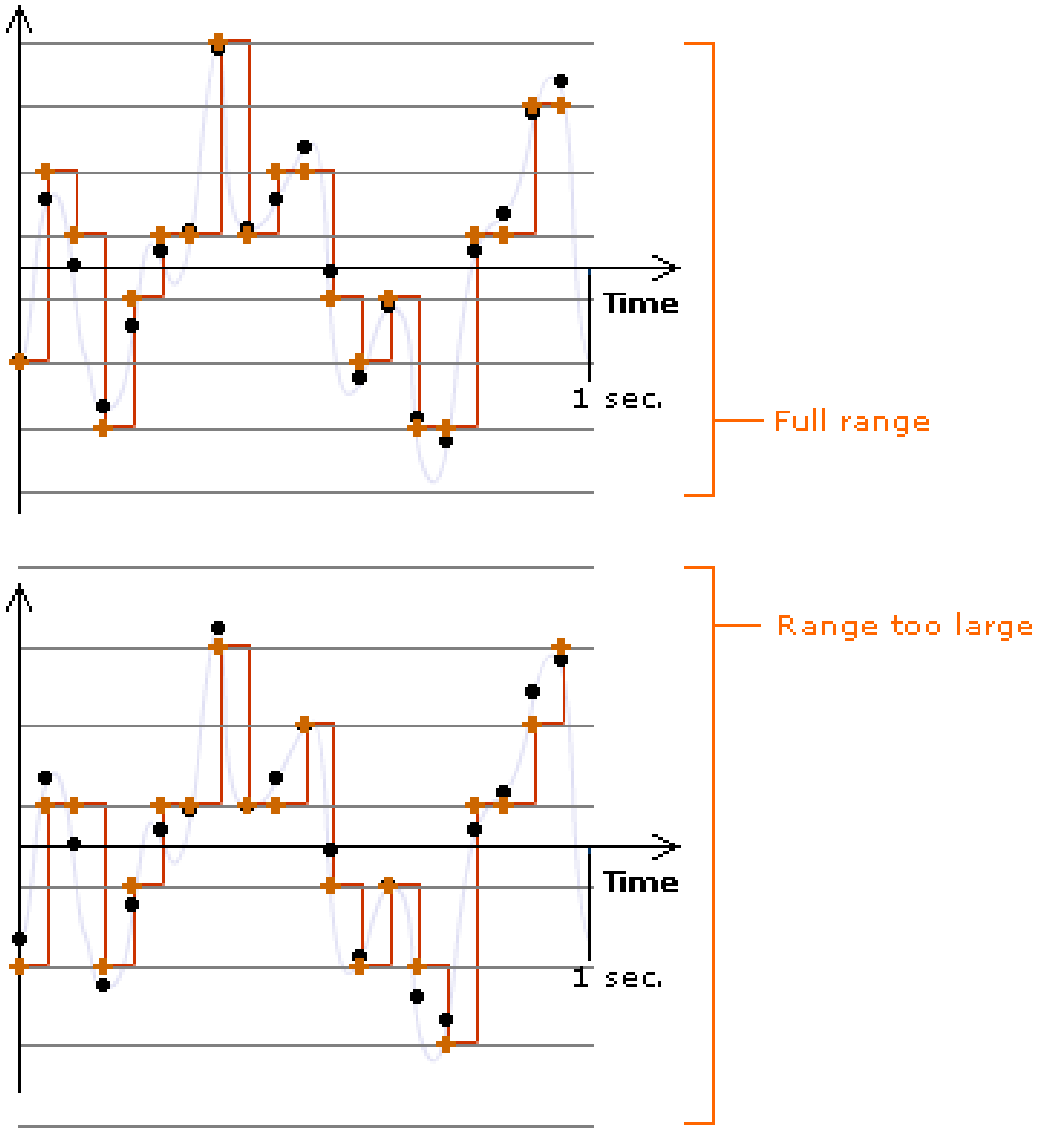
- In the previous example:



# Smaller than the Full Range




Time
1 sec.
Full range
Time
1 sec.
Narrow range


Time
1 sec.
Full range
Time
1 sec.
Range too large

# Review Question

The ___ of a digitized sound affects the accuracy of the sampled amplitudes being stored.

A. amplitude
B. frequency
C. sampling rate
D. bit depth
E. dynamic range

# Review Question

In digital audio, the number of sample points taken per second is called ___.

A. amplitude
B. frequency
C. sampling rate
D. bit depth
E. dynamic range

# Review Question

In digital audio, higher resolution means higher ___.

A. amplitude
B. frequency
C. sampling rate
D. bit depth
E. dynamic range

# Review Question

The unit used for measuring ___ is Hertz (Hz).

A. amplitude
B. frequency
C. sampling rate
D. bit depth
E. dynamic range

# Review Question

How many levels of amplitude values does a 8-bit sound allow?

# Review Question

How many levels of amplitude values does a 16-bit sound allow?

# Review Question

Generally, audio CD music sampling rate is ___ and bit depth is ___.

Fundamentals of Digital Audio

# FILE SIZE, FILE COMPRESSION, AND FILE TYPES OF DIGITAL AUDIO

# In this lecture, you will learn:

- The effects of sampling rate, bit depth, and number of channels on the audio file size
- The general strategies to reduce audio file size
- How to choose sampling rate and bit depth
- How to choose an audio file type

# Choices of Sampling Rate and Bit Depth

Higher sampling rate and bit depth:

- deliver better fidelity of a digitized file
- result in a larger file size (undesirable)

Let's estimate the file size of a 1-minute CD-quality audio file

# 1-minute CD Qualtiy Audio

- Sampling rate = 44100 Hz
  (i.e., 44,100 samples/second)
- Bit depth = 16
  (i.e., 16 bits/sample)
- Stereo
  (i.e., 2 channels: left and right channels)

# File Size of 1-min CD-quality Audio

- 1 minute = 60 seconds
- Total number of samples
  = 60 seconds $\times$ 44,100 samples/second
  = 2,646,000 samples
- Total number of bits required for these many samples
  = 2,646,000 samples $\times$ 16 bits/sample
  = 42,336,000 bits
  This is for one channel.
- Total bits for two channels
  = 42,336,000 bits/channel $\times$ 2 channels
  = 84,672,000 bits

# File Size of 1-min CD-quality Audio

84,672,000 bits
= 84,672,000 bits / (8 bits/byte)
= 10,584,000 bytes
= 10,584,000 bytes / (1024 bytes/KB)
≅ 10336 KB
= 10336 KB / (1024 KB/MB)
≅ <u>10 MB</u>

# Estimate Network Transfer Time

Suppose you are using 1.5Mbps (mega bits per second) broadband to download this 1-minute audio.

The time is no less than

84,672,000 bits / (1.5 Mbps)

= 84,672,000 bits / (1,500,000 bits/seconds)

≅ 56 seconds

# File Size of 1-hour CD-quality Audio

$\cong$ 10 MB/minute $\times$ 60 minutes/hour

= 600 MB/hour

# General Strategies to Reduce Digital Media File Size

- Reduce sampling rate
- Reduce bit depth
- Apply compression

- For digital audio, these can also be options:
  - reducing the number of channels
  - shorten the length of the audio

# Reduce Sampling Rate

- Sacrifices the fidelity of the digitized audio
- Need to weigh the quality against the file size
- Need to consider:
  - human perception of the audio
    (e.g., How perceptibe is the audio with lower sampling rate?)
  - how the audio is used
    - music: may need higher sampling rate
    - short sound clips such as explosion and looping ambient background noise: may work well with lower sampling rate

# Sampling Rate Examples

- 11,025 Hz AM Radio Quality/Speech
- 22,050 Hz Near FM Radio Quality (high-end multimedia)
- 44,100 Hz CD Quality
- 48,000 Hz DAT (digital audio tape) Quality
- 96,000 Hz DVD-Audio Quality
- 192,000 Hz DVD-Audio Quality

# Estimate Thresholds of Sampling Rate Based on Human Hearing

Let's consider these two factors:

1. Human hearing range
2. A rule called Nyquist's theorem

# Human Hearing Range

- Human hearing range: 20 Hz to 20,000 Hz
- Most sensitive to 2,000 Hz to 5,000 Hz

# *Nyquist Theorem*

We must sample at least 2 points in each sound wave cycle to be able to reconstruct the sound wave satisfactorily.

$\Rightarrow$Sampling rate of the audio $\geq$ twice of the audio frequency (called a *Nyquist rate*)

$\Rightarrow$Sampling rate of the audio is higher for audio with higher pitch

# Choosing Sampling Rate

Given the human hearing range (20 Hz to 20,000 Hz) and Nyquist Theorem, why do you think the sampling rate (44,100 Hz) for the CD-quality audio is reasonable?

# Choosing Sampling Rate

If we consider human ear's most sensitive range of frequency (2,000 Hz to 5,000 Hz), then what is the lowest sampling rate may be used that still satisfies the Nyquist Theorem?

A. 11,025 Hz AM Radio Quality/Speech
B. 22,050 Hz Near FM Radio Quality (high-end multimedia)
C. 44,100 Hz CD Quality
D. 48,000 Hz DAT (digital audio tape) Quality
E. 96,000 Hz DVD-Audio Quality
F. 192,000 Hz DVD-Audio Quality

# Effect of Sampling Rate on File Size

File size = duration × sampling rate × bit depth × number of channels

- File size is reduced in the same proportion as the reduction of the sampling rate
- Example: Reducing the sampling rate from 44,100 Hz to 22,050 Hz will reduce the file size by half.

# Effect of Bit Depth on File Size

File size = duration × sampling rate × bit depth × number of channels

- File size is reduced in the same proportion as the reduction of the bit depth
- Example: Reducing the bit depth from 16-bit to 8-bit will reduce the file size by half.

# Most Common Choices of Bit Depth

- 8-bit
  - usually sufficient for speech
  - in general, too low for music
- 16-bit
  - minimal bit depth for music
- 24-bit
- 32-bit

# Audio File Compression

- Lossless
- Lossy
  - gets rid of some data, but human perception is taken into consideration so that the data removed causes the least noticeable distortion
  - e.g. MP3 (good compression rate while preserving the perceivably high quality of the audio)

# Effect of Number of Channels on File Size

File size = duration × sampling rate × bit depth × number of channels

- File size is reduced in the same proportion as the reduction of the number of channels
- Example: Reducing the number of channels from 2 (stereo) to 1 (mono) will reduce the file size by half.

# Midi Audio

- MIDI is a shorthand representation of music stored in numeric form.
- It is not digitized sound.
- A sequencer software and sound synthesizer is required in order to create MIDI scores.
- MIDI is device dependent.
- Length of a MIDI file can be changed without affecting the pitch of the music or degrading audio quality.
- Working with MIDI requires knowledge of music theory.

# MIDI Versus Digital Audio

- MIDI is device dependent while digitized audio is device independent.
- MIDI files are much smaller than digitized audio.
- MIDI files sound better than digital audio files when played on a high-quality MIDI device.
- With MIDI, it is difficult to playback spoken dialog, while digitized audio can do so with ease.
- MIDI does not have consistent playback quality while digital audio provides consistent playback quality.
- One requires knowledge of music theory in order to run MIDI, while digital audio does not have this requirement.

# Common Audio File Types

| File Type | Acronym For | Originally Created By | File Info & Compression | Platforms |
|---|---|---|---|---|
| .wav | | IBM<br>Microsoft | • Compressed or uncompressed<br>• One of the HTML5 audio formats | • Windows<br>• Plays in Web browsers that support the .wav format of HTML5 audio (Firefox, Safari, Chrome, and Opera) |
| .mp3 | MPEG audio layer 3 | Moving Pictures Experts Group | • Good compression rate with perceivably high quality sound<br>• One of the HTML5 audio formats | • Cross-platform<br>•Plays in Web browsers that support the .wav format of HTML5 audio (Safari and IE) |
| .m4a | MPEG-4 format without the video data | Moving Pictures Experts Group | •AAC compression; same compression as the MPEG-4 H.264 without the video data<br>• One of the HTML5 audio formats | Plays in Web browsers that support the AAC format of HTML5 audio (Safari, IE, and Chrome) |

# Common Audio File Types

| File Type | Acronym For | Originally Created By | File Info & Compression | Platforms |
|---|---|---|---|---|
| .ogg or .oga | | Xiph.Org Foundation | • Usually referred to as Ogg Vorbis format<br>• One of the HTML5 audio formats | Plays in Web browsers that support the Ogg Vorbisformat of HTML5 audio (Firefox, Chrome, and Opera) |
| .mov | QuickTime movie | Apple | • Not just for video<br>• supports audio track and a MIDI track<br>• a variety of sound compressors<br>• files can be streamed<br>• "Fast Start" technology | Cross-platform; requires QuickTime player |

# Common Audio File Types

| File Type | Acronym For | Originally Created By | File Info & Compression | Platforms |
|---|---|---|---|---|
| .aiff | Audio Interchange File Format | Apple | compressed, uncompressed | Mac, Windows |
| .au<br>.snd | | Sun | compressed | Sun, Unix, Linux |
| .ra<br>.rm | Real Audio | Real Systems | compressed; can be streamed with Real Server | Cross-platform; requires Real player |
| .wma | Window Media Audio | Microsoft | | |

# Choosing an Audio File Type

Determined by the intended use

- File size limitation
- Intended audience
- Whether as a source file

# File Size Limitations

- Is your audio used on the Web?
  - file types that offer high compression
  - streaming audio file types

# Intended Audience

- What is the equipment that your audience will use to listen to your audio?

- If they are listening on computers, what are their operating systems?
  - cross-platform vs. single platform

# Whether as a Source File

If you are keeping the file for future editing, choose a file type:

- uncompressed
- allows lossless compression

# Review Question

True/False: 8-bit is generally considered to be adequate for recording music.

# Review Question

True/False: MP3 is a good file format to keep as a source file for further editing.

# Review Question

Which of the following file extensions indicate audio files?
BMP
WAV
JPEG
AIFF
MP3
GIF
JPG
PSD
TIFF
WMF

# Review Question

According to Nyquist's theorem, we must sample at least ___ points in each sound wave cycle to be able to reconstruct the sound wave satisfactorily.

In other words, the sampling rate of the audio must be at least ___ of the audio frequency.

# Review Question

The reduction of a digital audio file size can be achieved by ___.

A. reducing the sampling rate
B. reducing the pitch of the audio
C. reducing the bit depth
D. reducing the amplitude of the audio
E. applying file compression techniques

# Review Question

Higher ___ will result in larger file size.

A. amplitude
B. frequency
C. sampling rate
D. bit depth
E. dynamic range

# Review Question

Reducing the bit depth from 16-bit to 8-bit will ___.

A. have no effect on the file size
B. decrease the file size by half
C. decrease the file size to 1/8-th
D. decrease the file size to 1/16-th

# Review Question

Reducing the sampling rate from 44,100 Hz to 22,050 Hz will ___.

A. have no effect on the file size
B. decrease the file size by half
C. decrease the file size to about 1/22-th
D. decrease the file size to about 1/44-th

# Review Question

Reducing the number of channels from 2 (stereo) to 1 (mono) will ___.

A. have no effect on the file size
B. decrease the file size by half
C. decrease the file size to about 1/5-th
D. decrease the file size to about 1/10-th

# Adding Sound to Multimedia Project

- File formats compatible with multimedia authoring software being used along with delivery mediums, must be determined.
- Sound playback capabilities offered by end user's system must be studied.
- The type of sound, whether background music, special sound effects, or spoken dialog, must be decided.
- Digital audio or MIDI data should be selected on the basis of the location and time of use.

# Reference:


- Wong, Y.L. (2013), Digital Media Primer, Pearson Education, Chapter 5
- Vaughan, T (2011), Multimedia: Making it work, McGraw-Hill, USA, 8th Ed, Chapter 5.
- https://www.pacdv.com/sounds/