Audio

By Dr. Heba Hamdy

Assistant Professor, Multimedia Dep.

Digital Audio

- **Audio sources:**
  - Speech.
  - Sounds of instruments, Music.
  - Sounds of all other kinds (the sound of wind, train and ocean).
- Audio needs new methods for coding and processing.

Audio processing is a key task in multimedia systems

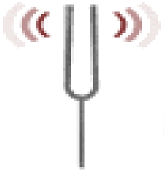
  - Audio coding (MPEG audio, mp3, AAC and others)
  - Authoring and representation (composition)
  - Analysis and searching (retrieval and database)
  - 3D sound, etc.

We will focus on basic audio processing, MPEG audio and related topics.

Sound Issues

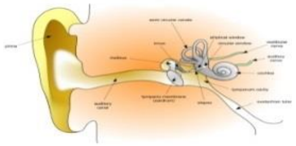
Sound includes the spoken word, voices, music and even noise.

It is a complex relationship involving:



- a vibrating object (sound source)

- a transmission medium (usually air)



- a receiver (ear) and;
- a preceptor (brain).

Sound Creation

Sound

- consists of pressure waves moving through a medium, such as air.
- It is heard by the ear (also called audio)
- it is analog nature and its waves are known as a continuous waveform

Volume: the higher the wave the louder the sound

Something vibrates in the air

Waves of pressure

Ear drums will translate these changes in wave Forms as sound

Audio

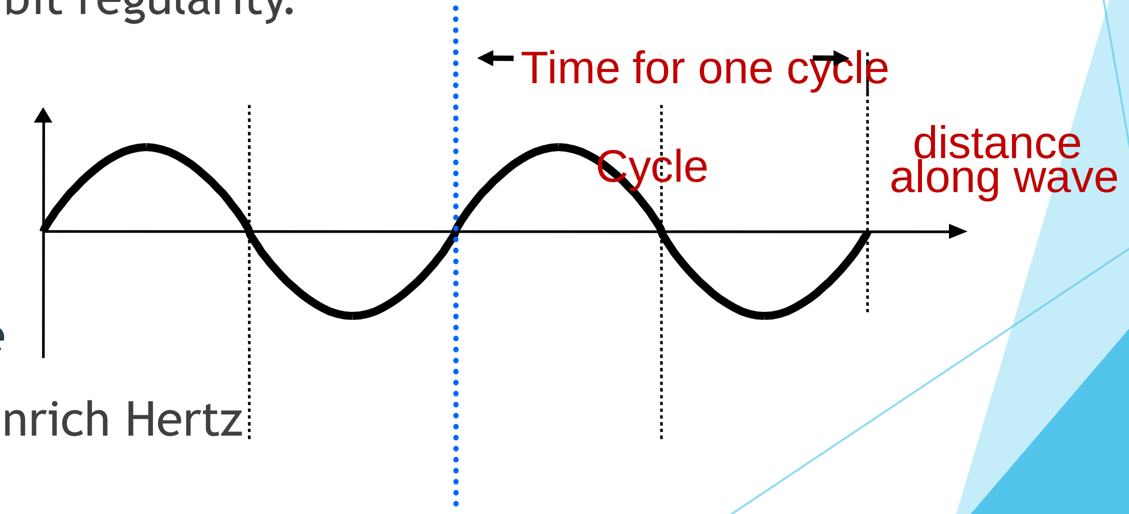
Types of Audio in Multimedia Applications:

1. **Music**-set the mood of the presentation, enhance the emotion ,illustrate points
2. **Sound effects**-to make specific points, e.g., squeaky doors, explosions, wind,...
3. **Narration**-most direct message, often effective

The most reliable audio formats in the multimedia world are **WAV** and **MP3**. There are many programs that also use **WMA, OGG, AAC, AIFF** and others

Physics component of Sound

- **Wave:** The compressions and refractions of air in a longitudinal waveform.
- **Waveform:** the measurement of the air particles **speed** and the **distance** that they travel.
- **Periodic waveform :** A waveform that repeats itself at regular intervals.
- **Noise :** Waveforms that do not exhibit regularity.



The unit of regularity is called a cycle

This is known as Hertz (or Hz) after Heinrich Hertz

- One cycle = 1 Hz
- Sometimes written as kHz or kiloHertz (1 kHz = 1000 Hz)

The characteristics of sound waves

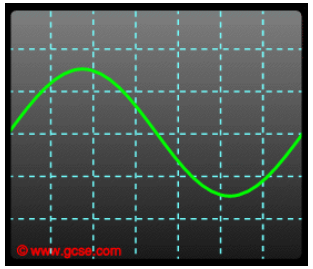
Sound Wave is described in terms of two characteristics:

- **Frequency** (or pitch)
- **Amplitude** (or loudness)

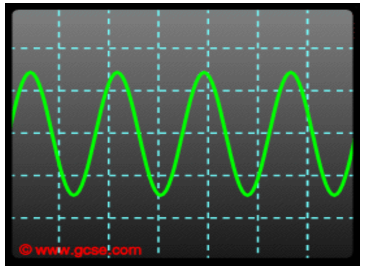
Frequency

Frequency: how many cycles occur in one second.

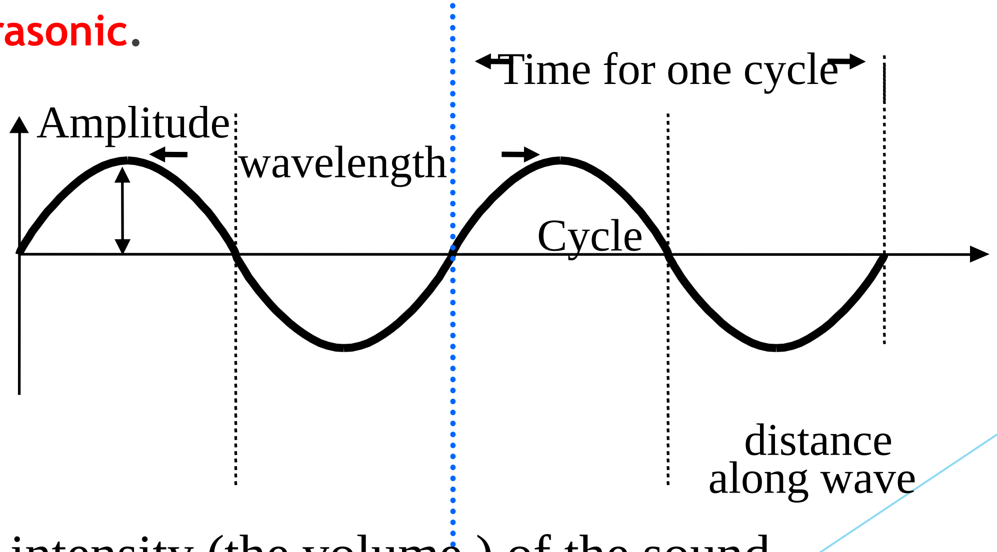
- Measured in Hertz that corresponds to the pitch of a sound.
- Optimally, people can hear from **20 Hz to 20,000 Hz (20 kHz)**
  - **Sounds below 20 Hz are infrasonic**
  - **sounds above 20 kHz are ultrasonic.**



Low pitch



High pitch

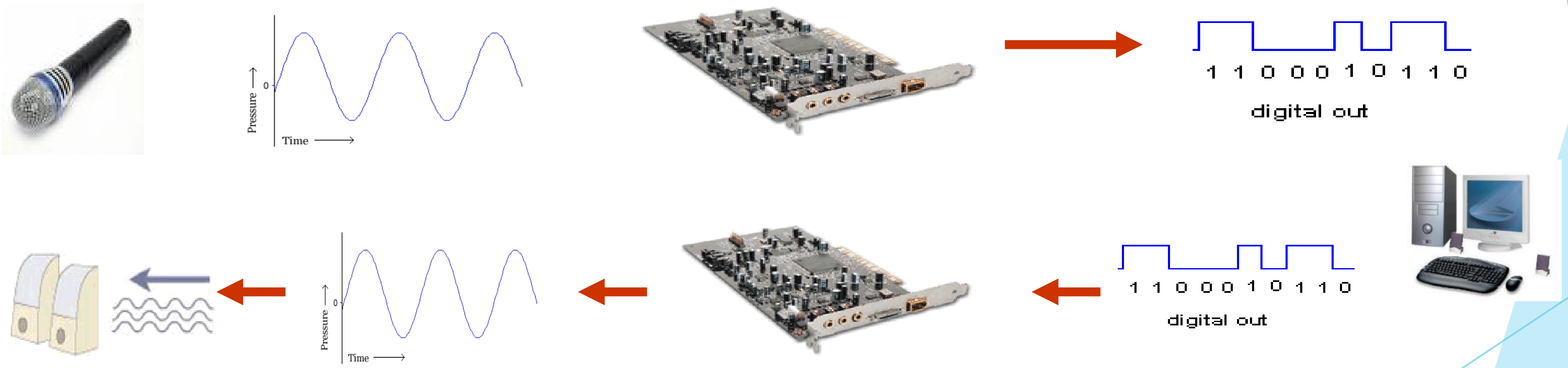


Amplitude: Measures the relative intensity (the volume) of the sound,
It measures in decibels (dB) unit

How is Sound Recorded?

- Computers have a sound card which converts the sound wave from a microphone as analog to digital form (Sampling operation)

Input path



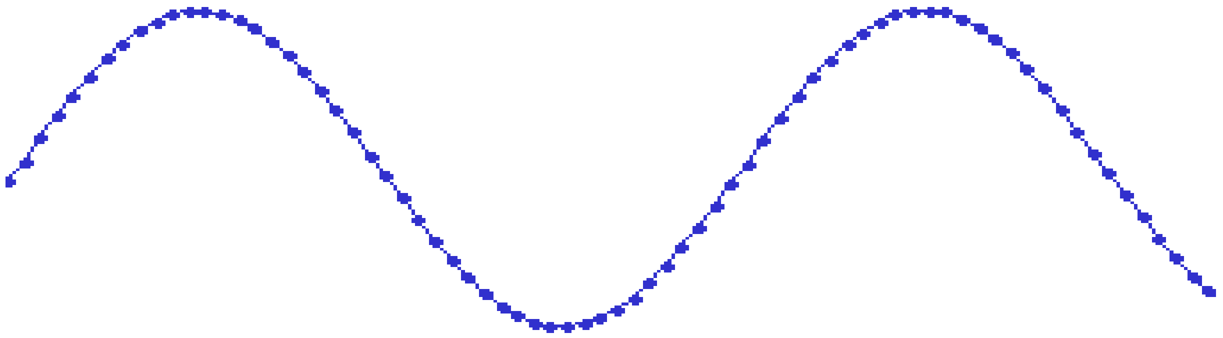
Output path

Sound card has:

- (ADC) **Analog-to-Digital Converter** for recording
- (DAC) **Digital-to-Analog Converter** for playing audio.

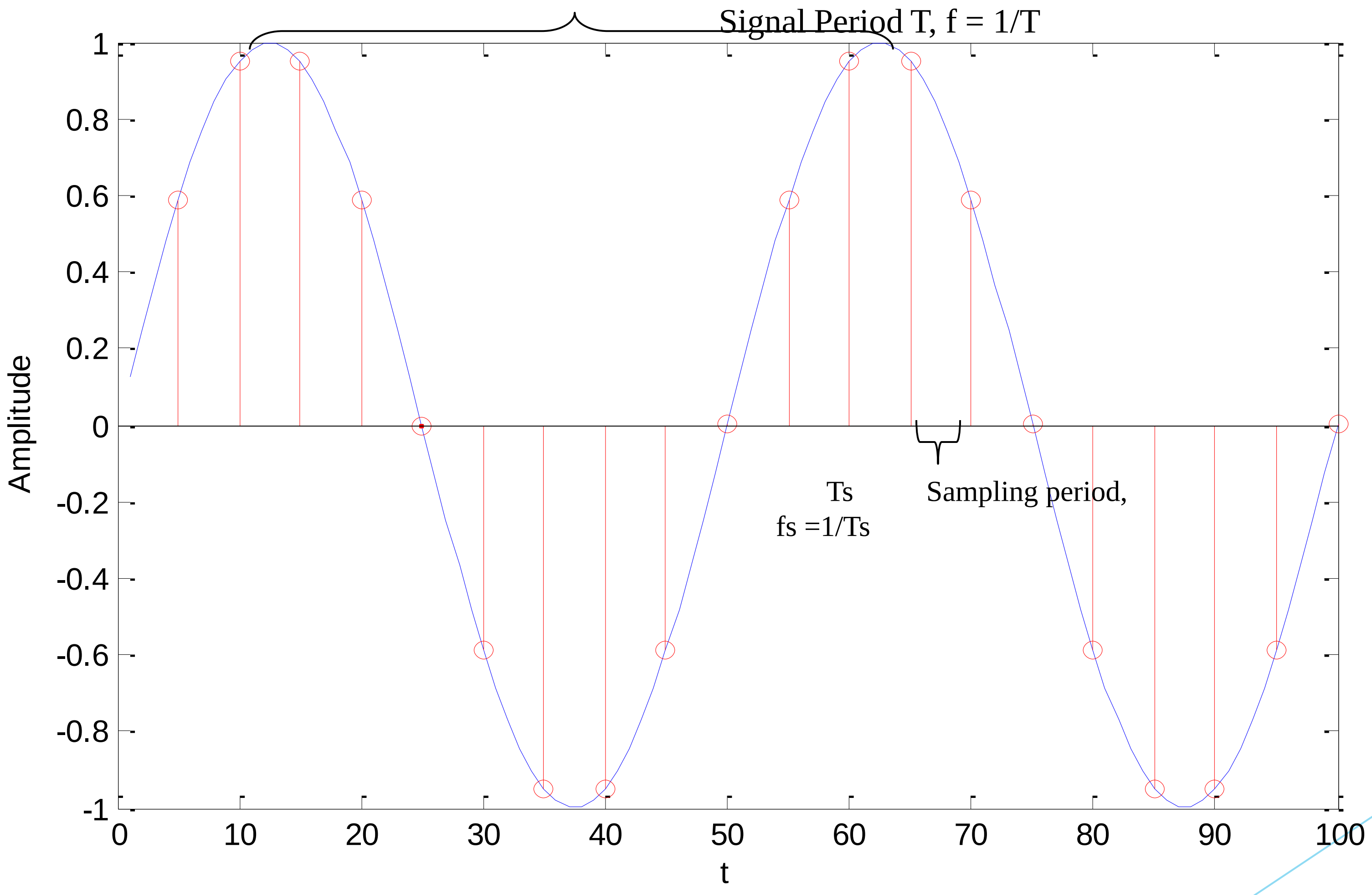
Sampling operation

- **Sampled audio**
  - Captured by sampling an analogue waveform at a set rate



- Each dot in the figure represents one audio ***sample***
- To convert **analog sounds to digital sounds**
  → thousands of **samples** are taken of the sound waves and recorded as bits.

Sampling for an Audio Signal



QUALITY OF DIGITAL RECORDING

DEPENDENT ON

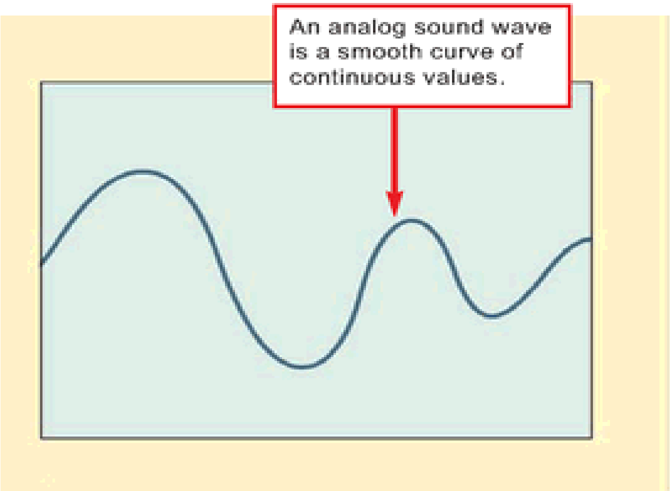
1. **Sample Rate**
2. **Sample Size**
3. **Channels**
4. **Codecs**

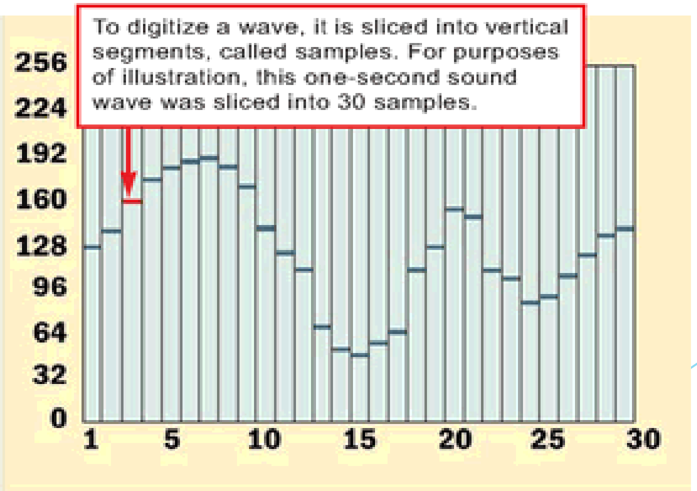
(1)SAMPLE RATE

Sample Rate is number of samples taken in a second

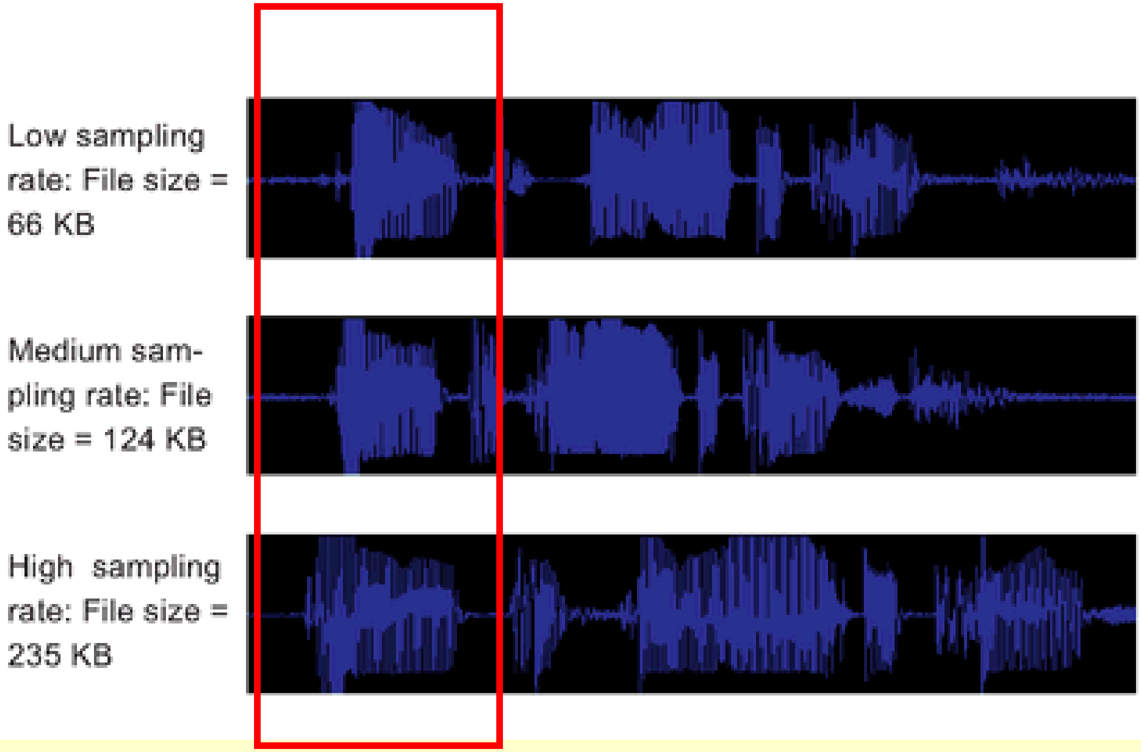
E.g. 1000 samples per second = 1 KHz (1000Hz)

- **The higher the sampling rate**, the more the measurements are taken (**better quality**).
- **The lower** the sampling rate, the lesser the measurements are taken (**low quality**).





Multimedia SAMPLE RATE



There are three sampling frequencies often used in multimedia are:

8 KHz **VOICE ONLY (telephone quality)**

11.025 KHz **AM Quality**

22 KHz **FM Quality**

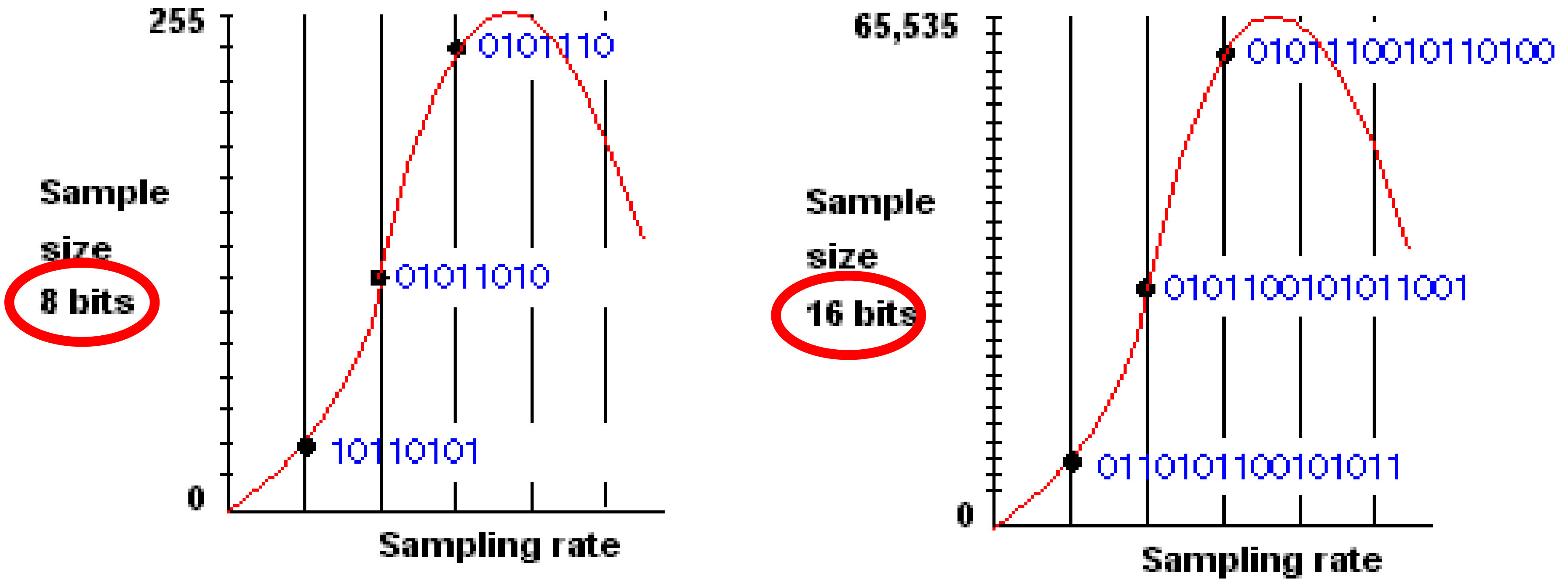
44.1 KHz **CD**

(2) SAMPLE SIZE -

SAMPLE SIZE (bit depth)

Each measurement taken is represented by a value . How many bits do we use to represent that value?

→The higher the sample size – better sound ⟶bigger file

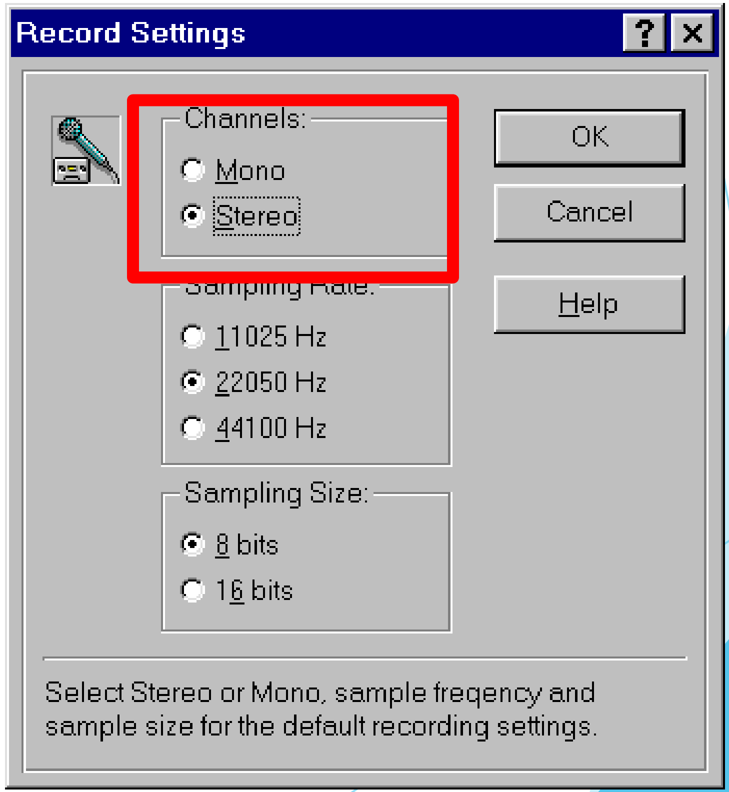


(3) CHANNELS -

CHANNELS : Mono vs. Stereo

Mono signal

- one channel(stream) of data reproduced equally on both speaker channels



Stereo signal

- consists of two channels (streams)of data
- double the size of mono file with higher quality

(4) CODECS -

CODECS: Software programs that greatly reduce the audio file size

Codecs for audio use lossy compression

- **Removes redundant and less-significant data**
- **Each time compression applied, quality of level of the file diminishes**
- **File formats for compression use as : (.mp3, .ra)**

Common File Formats that are uncompressed:

- **wav (very common, 44KHz, 16bit)**
- **aiff**
- **cdda (Red Book) → Standard for CDs, 44KHz, 16 bit per sample, 2 channels.**

The size of a file is a constant calculation based on:

- ✓**Sample rate**
- ✓**Sample size (bit Depth)**
- ✓**number of channels (mono or stereo),**
- ✓**codec**

Comparison of file sizes for an
uncompressed 16 bit, 44.1 kHz WAV music file:

Size in Bytes	**1 minute**	**2 minutes**	**3 minutes**
Mono	**5,168 (5.2 MB)**	**10,336 (10.3 MB)**	**15,504 (15.5 MB)**
Stereo	**10,336 (10.3 MB)**	**20,672 (20.6 MB)**	**31,007 (31 MB)**

How big can audio get?

- An example of uncompressed sound with CD quality for 1 minute of audio:
  1 minute of recording → **60 seconds**

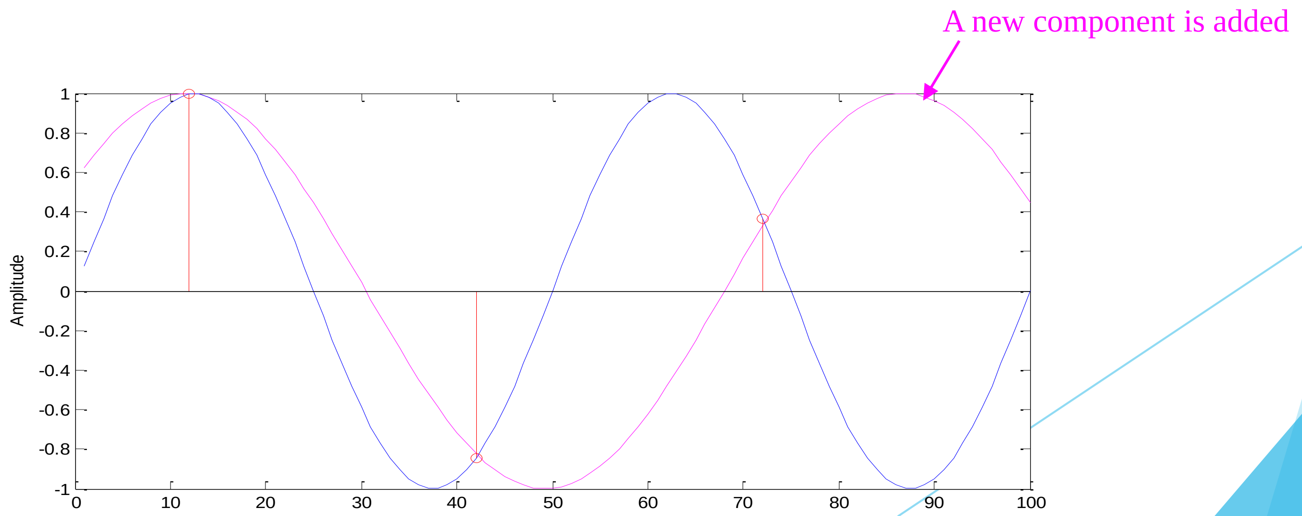
- 60 * 44,100 samples/second → **2,646,000 samples**
- 2,646,000 samples * 16bits per sample → 42,336,000 bits
- 42,336,000 bits * 2 (stereo, 2 channels) → 84,672,000 bits
- 84,672,000 bits / (8bits per byte) → 10,884,100 → **About 10 MB (Megabytes)!!!**

Sampling problem

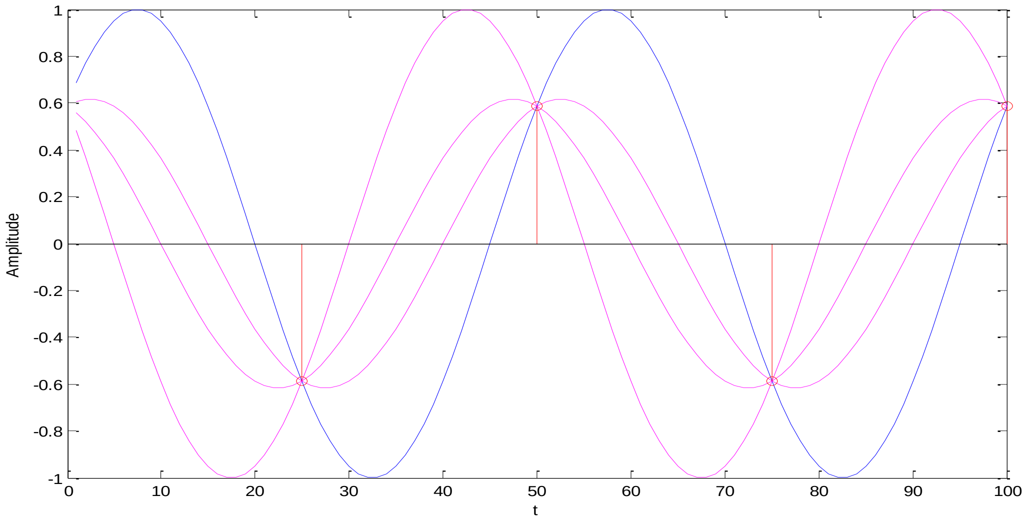
Aliasing: In sound and image generation, aliasing is a Distortion caused due to the generation of a false (alias) frequency along with the correct one when doing a low rate of frequency (Data) sampling .

- For images, this produces a jagged edge, or stair-step effect.
- For sound, it produces a buzz

Aliasing is Cured or subdued by antialiasing techniques such as filtering or super-sampling.



Number of aliasing signals



There are infinite number of possible sin waves going through the sampling points

Frequency Decomposition

Any signal can be represented as the summation of sin waves (possibly infinite number of them).

- We can use “Fourier Transform” to compute these frequency components.
- We can now extend our analysis to any signals.
- If we have a signal has frequency components {f1 < f2 < f3 … < fn}

SO:

- **what is the minimum sampling frequency we should use?**

Nyquist Theorem

- The Nyquist Theorem, also known as the sampling theorem, is a principle that engineers follow in the digitization of analog signals. For analog-to-digital conversion (ADC) to result in a faithful reproduction of the signal, slices, called samples, of the analog waveform must be taken frequently. The number of samples per second is called the sampling rate or sampling frequency.

- Any analog signal consists of components at various frequencies. The simplest case is the sine wave, in which all the signal energy is concentrated at one frequency. In practice, analog signals usually have complex waveforms, with components at many frequencies. The highest frequency component in an analog signal determines the bandwidth of that signal. The higher the frequency, the greater the bandwidth, if all other factors are held constant.

Nyquist Theorem

- Suppose the highest frequency component, in hertz, for a given analog signal is fmax. According to the Nyquist Theorem, the sampling rate must be at least 2fmax, or twice the highest analog frequency component. The sampling in an analog-to-digital converter is actuated by a pulse generator (clock). If the sampling rate is less than 2fmax, some of the highest frequency components in the analog input signal will not be correctly represented in the digitized output. When such a digital signal is converted back to analog form by a digital-to-analog converter, false frequency components appear that were not in the original analog signal. This undesirable condition is a form of distortion called aliasing.

Nyquist Theorem

- **For lossless digitization, the sampling rate should be at least twice the maximum frequency response.**
- **In mathematical terms:**
  - The necessary condition of reconstructing a continuous signal from the sampling version is that the sampling frequency
    - $f_s > 2*f_{max}$

f_s :is the sampling frequency

f_{max} :is the maximum frequency in the signal (the highest frequency component in the signal).

- **If a signal's frequency components are restricted in**
  - [f1, f2], we need fs >2 (f2-f1).

Features of Audio Files

There are two most important parameters you will be thinking about when working with digital audio:

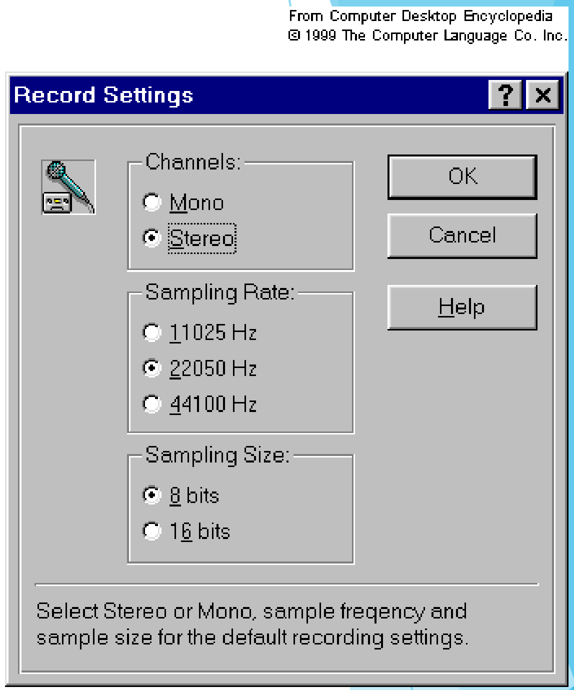
- **Sound quality:**

will be your major concern if you want to broadcast your programme on FM or short wave.

- **Audio file size.**

creating audio for exchange over the Net (uploading/downloading),or to include it on a Web Site, file size will be one of your greatest concerns

Sound-Editing Software



Sound editing programs characteristics

- record sounds in different sample sizes and resolutions
- save sounds in different file formats
- In addition to:
  - Edits: unwanted noise, pauses, trimming
  - special effects : Fade-ins, fade-outs, background music.
  - Several files: can be mixed or spliced together
  - Save the file: in a format intended for the audience application
  - Most sound files are saved in: .mp3 , .wav.

Sound on the Web

Using audio on the Web?

Audio files are quite large

File Size and File format is important parameters

File Size

- Select small sized files
- The smaller the file, the faster the webpage will download, and the faster it will start up playing the sound file

File Formats

Consider:

- Platform Support
- Browser Support
- Compressed or Uncompressed quality
- Compatibility of Application Use the file

Audio File Formats

Audio Format	File Extension	Advantages	Disadvantages
Advanced Audio Compression	.aac	•Good sound quality •Used on iTunes	•Copy protected •Limited to approved devices
Audio Interchange Format	.aif /.aiff	•Excellent sound quality •Supported without a plug-in	•Uncompressed so large files
MP3	.mp3	•Good sound quality even though compressed •Can be streamed over the Web	•Requires standalone player or browser plug-in
Real Audio	.ra, .rx	•High Compression •Very small files •Can be streamed over the web	•Sound quality not great •Requires a player or plug-in
Wave	.wav	•Good sound quality •Supported without a plug-in	•Uncompressed, very large files
Windows Media Audio	.wma	•Good sound quality even though compressed •Used on music download sites	•Files can be copy protected •Requires Windows Media Player 9 or higher

Assignment

- IT Students [team group, 2 students]:

Select on Audio Edit Program, Make a video demo and upload on Facebook Group.

- IS Students [team group, 2 students]:

Select on Image Edit Program, Make a video demo and upload on Facebook Group.

- MM Students [Individual]:

Select on Image Edit Program, Make a video demo and upload on Facebook Group.