UNIT III
AUDIO MEDIA

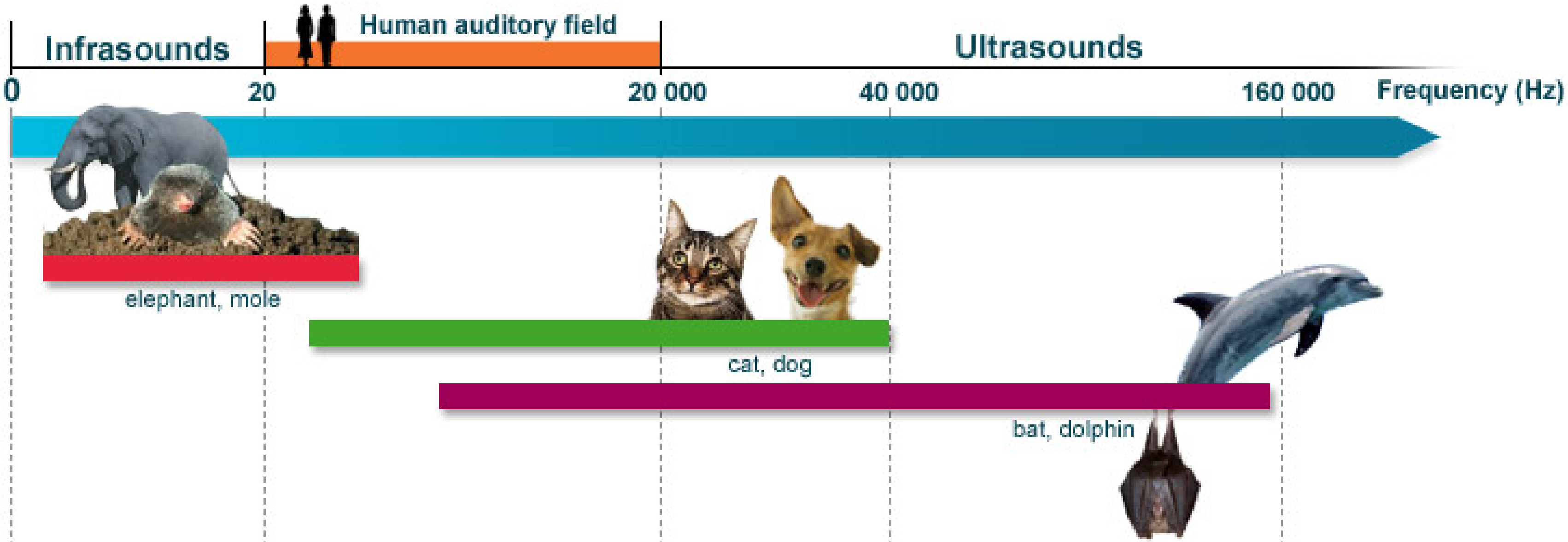
What is ‘Audio’?

- Audio means ‘of sound’ or ‘of the reproduction of sound’.
- It also refers to the range of sound frequencies which can be heard by humans.
- Specifically, it refers to the range of frequencies detectable by the human ear — approximately 20Hz to 20kHz.
- **20Hz** is the lowest-pitched sound we can hear, **20kHz** is the highest pitch we can hear.
- In physics, sound is a form of energy known as *acoustical energy*.

Sound Vs Audio

<u>Classes of Sound:</u>

- **Voice**
  - Defined as talking.
- **Music**
- **Sound Effect:**
  - Voice or Music; but often created by natural events like thunderclap, wind and door slamming.



- **Audio work** involves the production, recording, manipulation and reproduction of sound waves. To understand audio you must have a grasp of two things:
- **Sound Waves:** What they are, how they are produced and how we hear them.
- **Sound Equipment:** What the different components are, what they do, how to choose the correct equipment and use it properly.

The Field of Audio Work

- The field of audio is vast, with many areas of specialty.
- Hobbyists use audio for all sorts of things, and audio professionals can be found in a huge range of vocations.
- Some common areas of audio work include:

Studio Sound Engineer
Live Sound Engineer
Musician
Music Producer
DJ
Radio technician

Film/Television Sound Recordist
Field Sound Engineer
Audio Editor
Post-Production Audio Creator

- In addition, many other professions require a level of audio proficiency.
- For example, video camera operators should know enough about audio to be able to record good quality sound with their pictures.

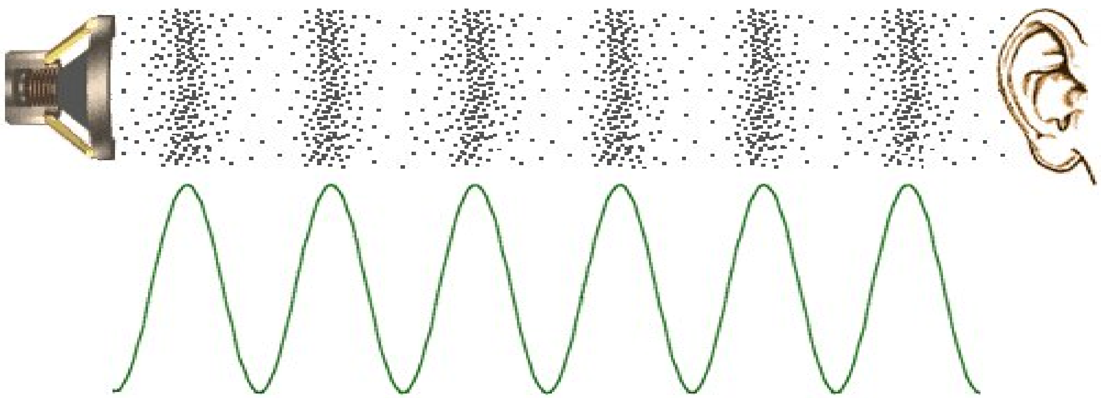
How Sound Waves Work

- Sound waves exist as variations of pressure in a medium such as air. They are created by the vibration of an object, which causes the air surrounding it to vibrate. The vibrating air then causes the human eardrum to vibrate, which the brain interprets as sound.

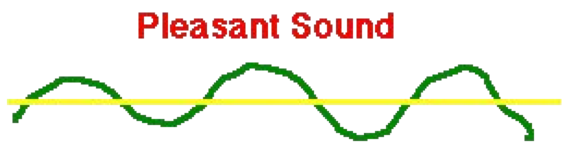
- When an object vibrates, it causes movement in surrounding air molecules.
- These molecules bump into the molecules close to them, causing them to vibrate as well.
- This makes them bump into more nearby air molecules.
- This "chain reaction" movement, called sound waves, keeps going until the molecules run out of energy.

How do We Hear?

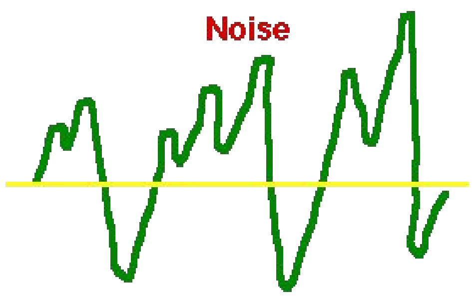
- Sound waves are variations of pressure in a medium such as air.
- Sound created by the vibration of an object, which causes the air surrounding it (medium) to vibrate.
- Vibrating air causes the human eardrum to vibrate, which the brain interprets as sound.



- Vibrations in the air create waves of pressure that are perceived as sound.
- Sound comprises the spoken word, voices, music and even noise.
- Sound waves vary in sound pressure level (amplitude) and in frequency or pitch.
- ‘Acoustics’ is the branch of physics that studies sound.
- Sound pressure levels (loudness or volume) are measured in decibels (dB).
- Sound waves travel through air in much the same way as water waves travel through water.
- In fact, since water waves are easy to see and understand, they are often used as an analogy to illustrate how sound waves behave.

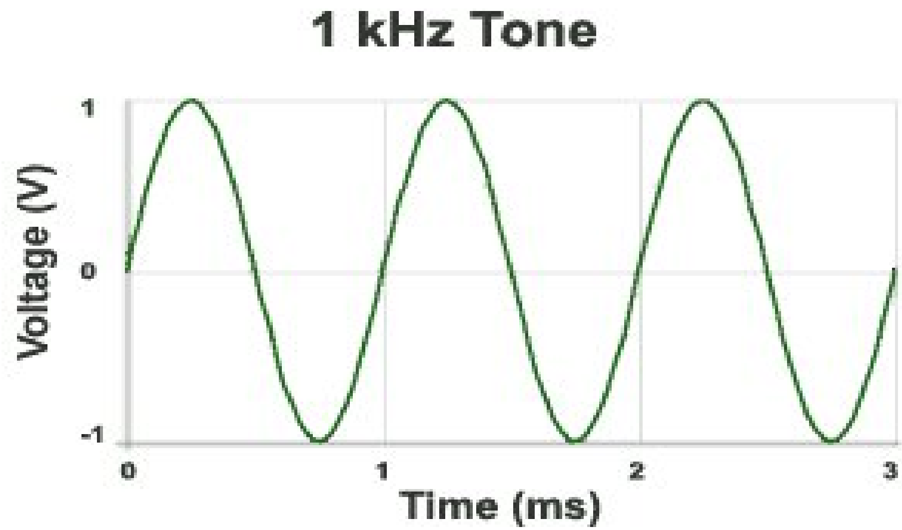


- A pleasant sound has a regular wave pattern. The pattern is repeated over and over.



- But the waves of noise are irregular. They do not have a repeated pattern.

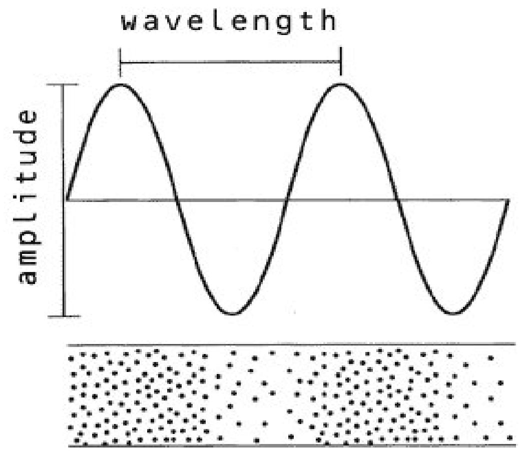
- Sound waves can also be shown in a standard ***x Vs. y* graph**, as shown below.
- This allows us to visualize and work with waves from a mathematical point of view. The resulting curves are known as the '**waveform**' (i.e. the form of the wave.)



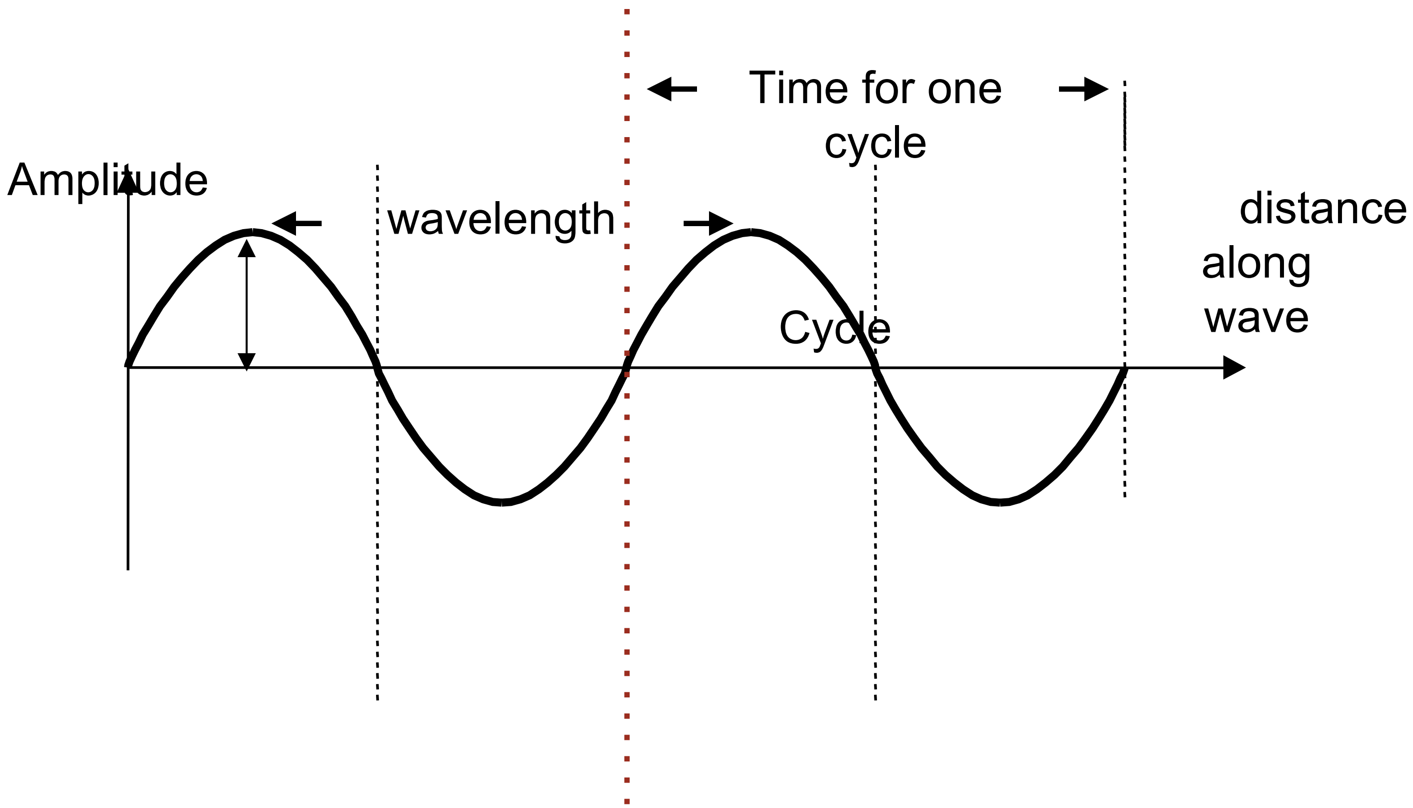
Characteristic of Sound Waves

- Sound is described in terms of two characteristics:
  - **Frequency** (or pitch)
  - **Amplitude** (or loudness)

- Wavelength
- Period
- speed

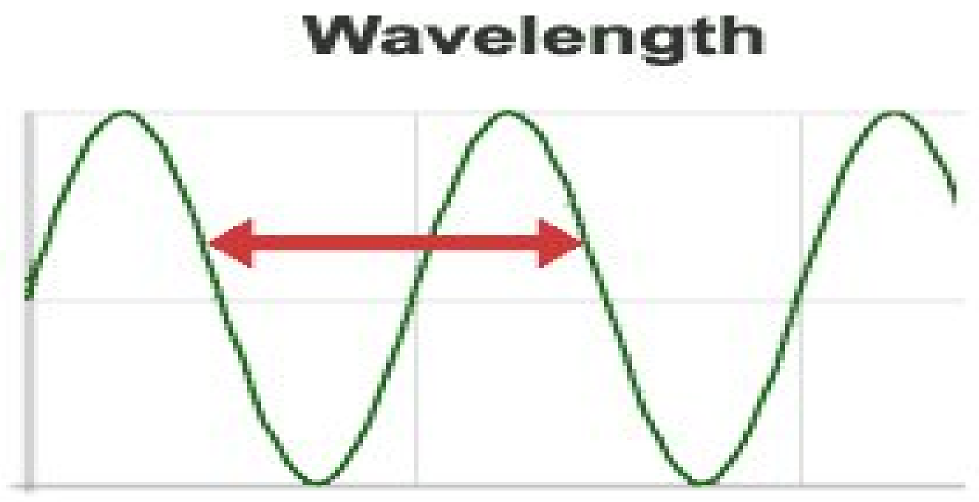


Characteristic of Sound Waves



Sound Wave Properties

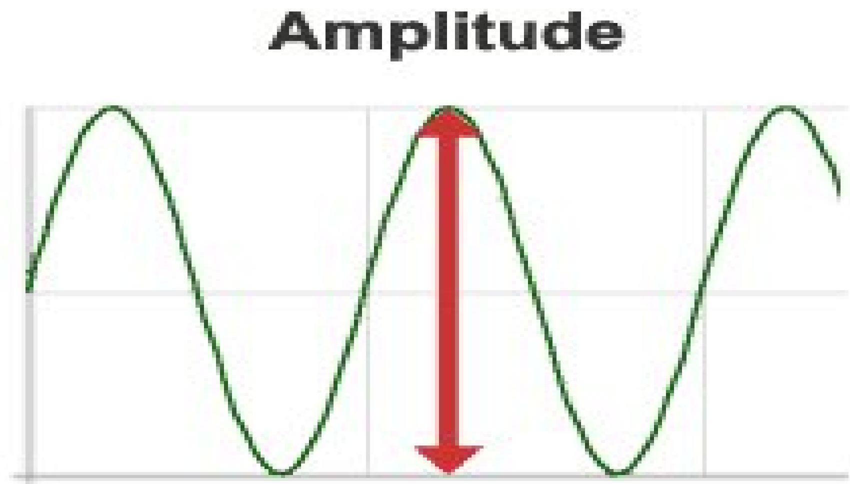
- **Wavelength:** The distance between any point on a wave and the equivalent point on the next phase. Literally, the length of the wave.



Sound Wave Properties

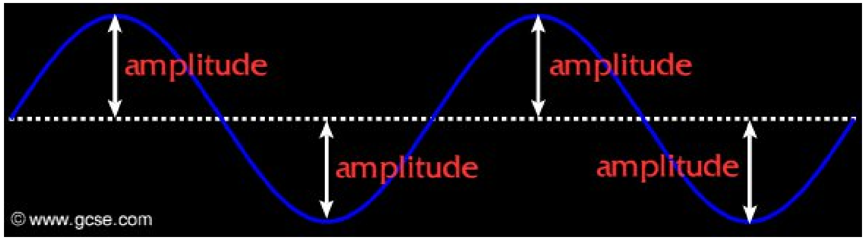
Amplitude: The strength or power of a wave signal. The "height" of a wave when viewed as a graph.

Higher amplitudes are interpreted as a higher volume, hence the name "amplifier" for a device that increases amplitude.

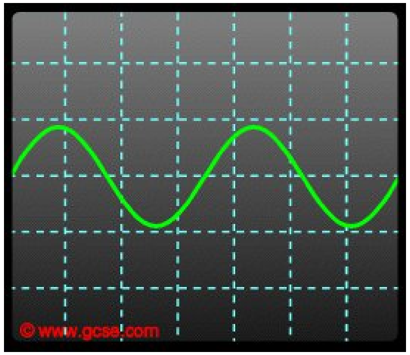


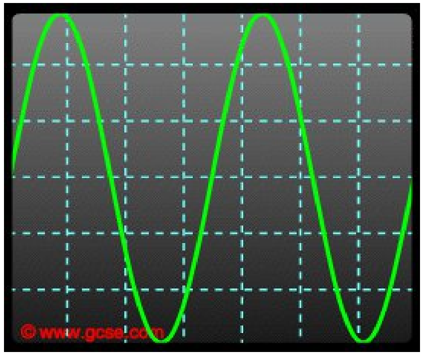
Amplitude

- ***Amplitude*** is the *maximum displacement* of a wave from an equilibrium position.
  - The louder a sound, the more energy it has. This means loud sounds have a large amplitude.



Quiet





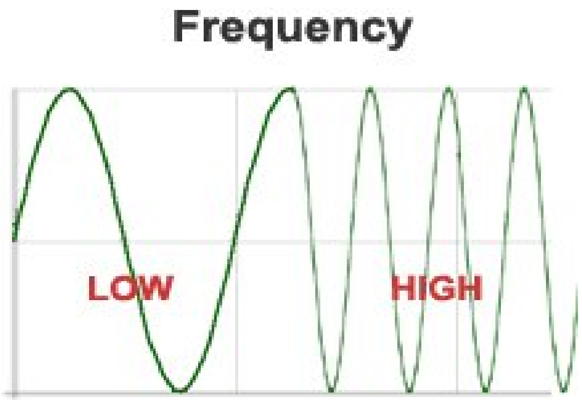
Loud

Low amplitude   High Amplitude

- The amplitude relates to how loud a sound is.

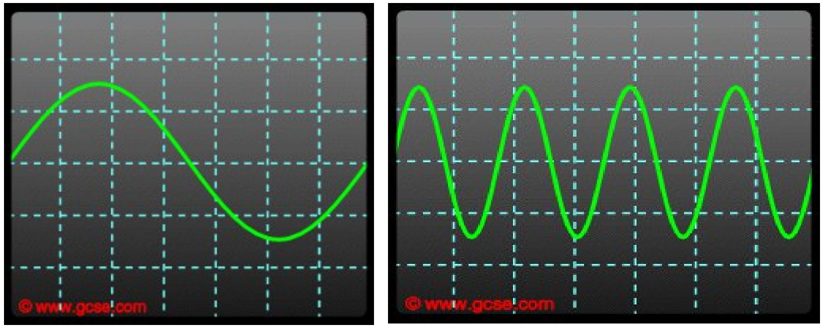
Properties of Sound

- Frequency: Number of times the wavelength occurs in one second.
- Measured in Hertz (Hz), or cycles per second.
- The faster the sound source vibrates, the higher the frequency, the higher the pitch
- Example: singing in a high-pitched voice forces the vocal chords to vibrate quickly.



Frequency

- ***<u>Frequency</u>*** is a measure of how many cycles occur in one second. This is measured in *Hertz* (abbreviation Hz) and directly corresponds to the ***pitch*** of a sound.
  - The <u>more</u> frequent <u>vibration</u> occurs <u>the higher the pitch</u> of the sound.



Low pitch High pitch

- Optimally, people can hear from **<u>20 Hz to 20,000</u>** Hz (20 kHz)
  - Sounds below 20 Hz are infrasonic
  - sounds above 20 kHz are ultrasonic.

DIGITAL AUDIO

Digital audio

- It is the actual representation of sound stored in the form of digital numbers (Samples)
- It represents the loudness of the sound at a slice of time
- It sounds the same every time played.

Digital Representation

- Analog audio signals are typically represented as waveforms
  - **Simple**
  - **Complex.**
- A **simple sinusoidal wave** corresponds to a pure tone at a single frequency, or pitch.
- The amplitude of the wave gives the strength of the sinusoid at that time.

Digital Representation

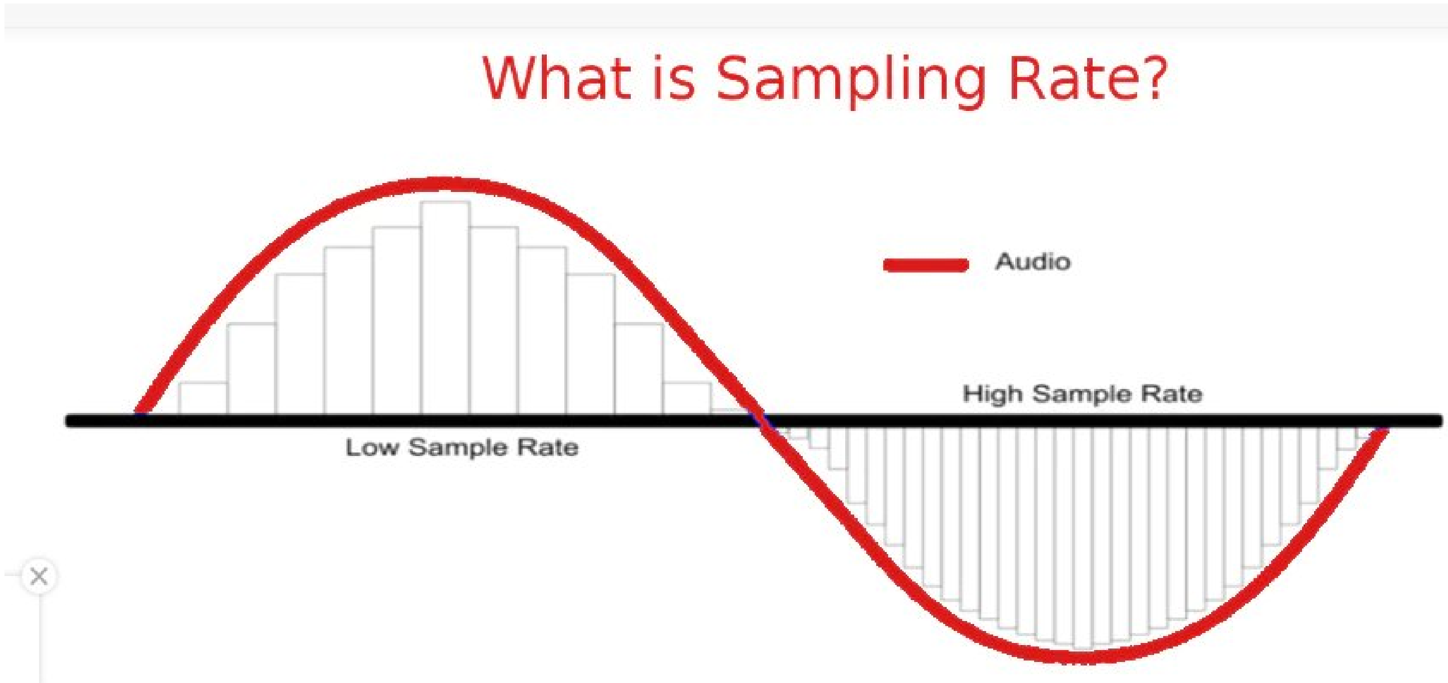
- A **complex wave** consists of multiple frequencies or sinusoidal waves combined together.
- Most audio signals that are of interest to us, such as voice, music, and so on, are composed of multiple frequencies, where the amplitude of the complex signal is the joint combination of the amplitudes of the individual frequencies.
- This understanding, which is normally referred to as the **frequency domain representation** of the analog signal.

Digital Representation

- Digitizing an analog audio signal requires
  - **Sampling and**
  - **Quantization.**
- The process of conversion to digital sound is known as **pulse code modulation (PCM).**
- The analog sound is sensed at evenly spaced time intervals, producing digital audio samples.

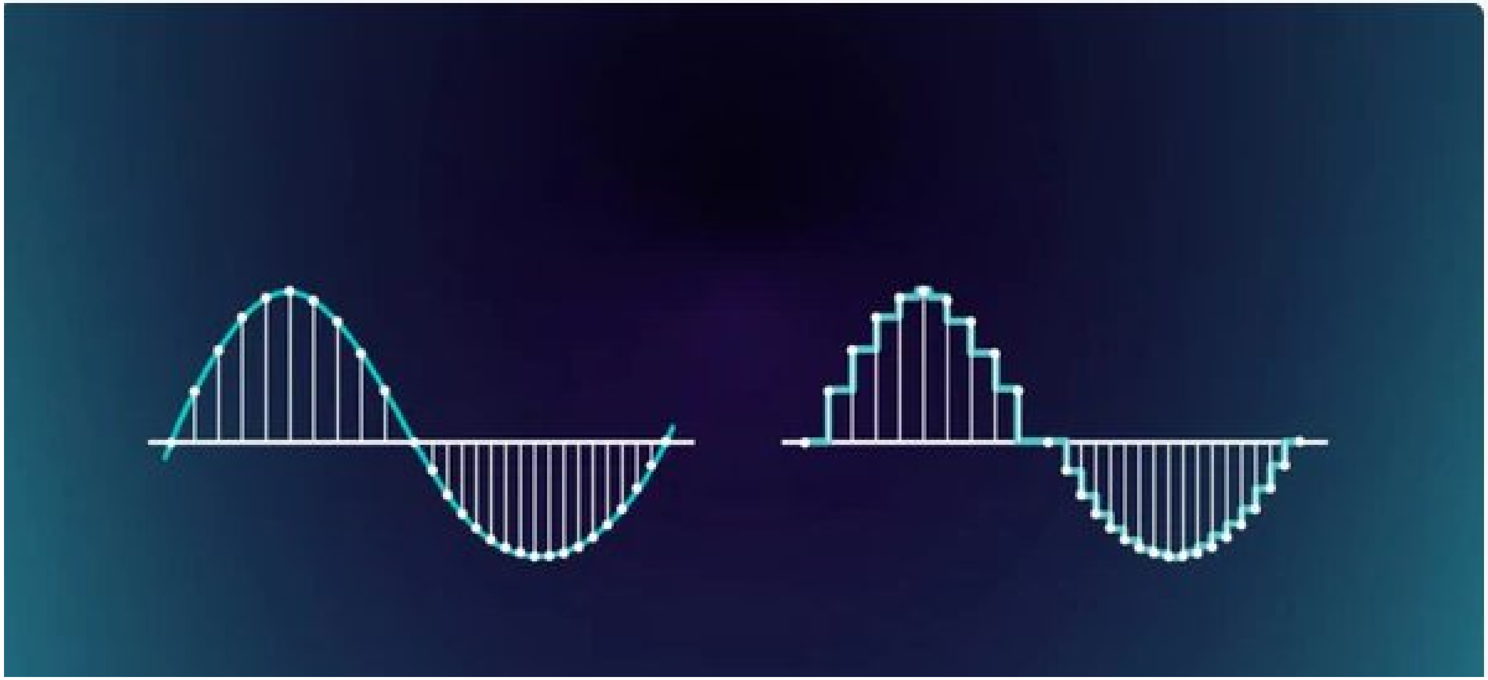
Digital Representation

- The number of samples per time unit or per second (**sampling rate**) must be specified during the digitization process.
- Given a sample, the signal amplitude at that position is encoded on a fixed number of bits. All samples are represented by the same number of quantization bits.
- The sampling rate and the quantization bits per sample are the main properties of the PCM signal and need to be carefully chosen so that it is possible to reconstruct the analog equivalent.



What is Sampling Rate?
Audio
High Sample Rate
Low Sample Rate

- *Sampling Rate*



Digital Representation

- The digital audio signal is finally rendered by converting it to the analog domain and so the choice of sampling rate and sample size need to be chosen appropriately in order to faithfully re-create the original sound.

Digital Representation

- In addition to sampling rate and quantization, another characteristic commonly used to describe audio signals is the number of channels, which may be one (mono), two (stereo), or multichannel (surround sound). Also of growing interest in research and industry is the notion of spatial audio. Mono and stereo sound technology are the most commonly used.

Commonly Used Audio Formats

File suffix or logo	Filename	File type	Features
.wav	WAV	Uncompressed PCM coded	Default standard for audio on PCs. WAV files are coded in PCM format.
.au	G.711 μ-law, or ITU μ-law	Uncompressed audio	Universal support for telephone. Packs each 16-bit sample into 8 bits, by using logarithmic table to encode with a 13-bit dynamic range. Encoding and decoding is very fast.
GSM 06.10	Global System for Mobile Communication	Lossy Compressed mobile audio	International standard for cellular telephone technology. Uses linear predictive coding to substantially compress the data. Compression/decompression is slow. Freely available and, thus, widely used
.mp3	MPEG1 Layer3	Compressed audio file format	Uses psychoacoustics for compression Very good bandwidth savings and, hence, used for streaming and Internet downloads.

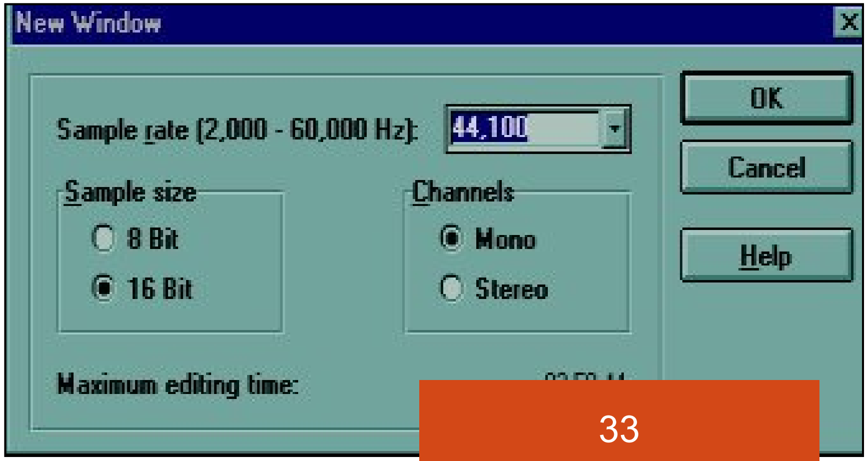
.ra	Real Audio	Compressed format	Proprietary to Real Audio. Capable of streaming and downloading. Comparable quality to mp3 at high data rates but not so at low data rates
AAC	Advanced Audio Codec MPEG4	Compressed format	Superior quality to .mp3.
.mid	MIDI—Musical Instrument Digital Interface	Descriptive format	MIDI is a language of communication among musical instruments. Description achieved by frequencies, decays, transients, and event lists. Sound has to be synthesized by the instrument.
DOLBY DIGITAL	Dolby Digital (formerly called	Compressed 5.1 surround sound	De facto standard of home entertainment (Dolby AC-3) Distributed with DVD, HDTV systems. Provides five discrete channels—center, left, right, surround left, and surround right—plus an additional six for LFE.
DIGITAL dts SURROUND	DTS Surround Sound	Compressed 5.1 surround sound	Alternate to Dolby Digital. Distributed with DVDs, but not HDTV. Has higher data rate compared with Dolby Digital.
LUCASFILM THX ULTRA	THX Surround Sound	Compressed 5.1 surround sound	Designed for movie theaters (THX Ultra) as well home theaters (THX Select). Has become the select brand for surround sound today.
LUCASFILM THX ULTRA	THX Surround Sound Extended	Compressed 6.1 or 7.1 surround sound	Jointly developed by Lucasfilm, THX and Dolby Laboratories. Also known as Dolby Digital ES. Has a surround back channel, placed behind audience achieving 360^0 of sound.

Characteristics of digital sound

- Three main characteristics :
  - Frequency
    - defines the number of samples per second (or per other unit) taken from a continuous signal to make a discrete signal.
    - For time-domain signals, it can be measured in hertz (Hz).
  - Sound resolution / Amplitude measurement
    - Number of bits used to represent a sample.
  - Channel
    - Mono or stereo

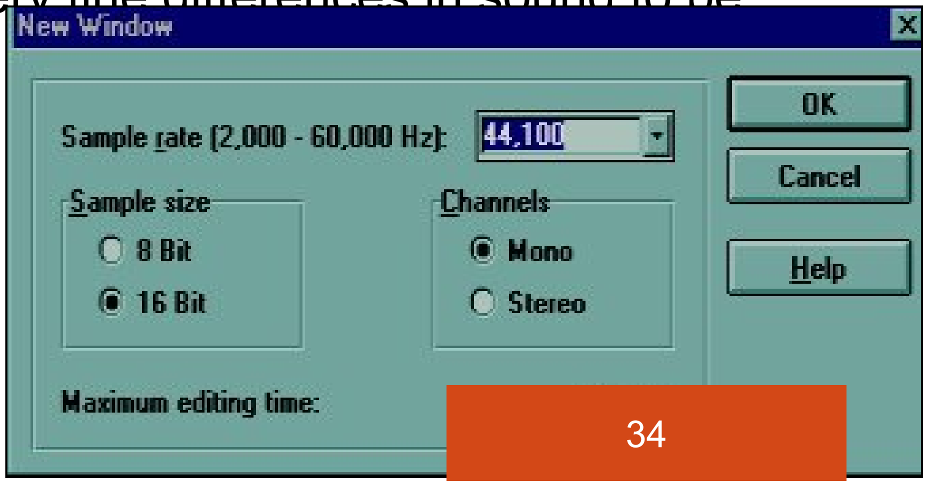
Frequency

- A **higher frequency** sampling rate means more samples; **better quality**.
- The more the samples there are, the **more storage space** will be needed.
- **Higher Frequency -> higher quality -> higher storage space**
- Sound cards are able to record sound at different sampling rates.
- Depending on the user's choice sound can be recorded at 11.025 kHz, 22.05 kHz and 44.1 kHz which is CD quality.



Sound resolution / Amplitude Measurement

- Based on 8bits (1 byte).
- 8 bits for 256 levels & 16 bits for 65536 levels.
- The number of bit-sampling too, will affect the size of the file.
- The higher the bit-sampling, the larger the size of the file.
- Usually, for narrations, 8-bit sampling is quite sufficient.
- If you want high quality sound, 16-bit will be a preferred choice.
- Higher sound resolution allows very fine differences in sound to be recorded.



Sound channel

- Whether you want mono or stereo sound will **affect** the size of the file.

- Mono means sound will be playing from one channel whereas stereo means two channels.

- Therefore, stereo sound will require larger storage space than mono sound.

Reasons to use digital audio

- It is consistent: the digital media will sound as good at the end as it did in the beginning when it was created.
- A wider selection of application software and system support for digital audio is available.
- The preparation and programming required for creating digital audio do not demand knowledge of music theory; working with MIDI data usually does require familiarity with musical scores, keyboards, and notation as well as audio production.

When to use digital audio

- You don’t have control over playback hardware.
- You have the computing recourses and bandwidth to handle digital files.
- You need spoken dialog.

Benefits of using digital audio

- Sound can be permanently stored in inexpensive CD.
- Consistent sound quality without noise or distortion.
- Duplicate will sound exactly the same as the master copy.
- Digital sound can be played at any point of the sound track. (random access)
- It can also be integrated with other media.
- Can be edited without loss in quality.

Principles of Audio Recording

Audio quality is essential to create a pleasant experience, and by extension adequate engagement from listeners. Depending on the gravity of the audio defects, post processing can be used to clean the audio file, but it is best to avoid these during the recording stage.

There are a few things to consider when creating a high quality audio recordings:

- **Noise**
- **Reverberation**
- **Audio Clipping**
- **Compression**

Principles of Audio Recording

Noise

- Noise refers to the unwanted signal that gets recorded to a medium.
- Noise can take different forms, the most common of which is background noise.
- Background noise is usually caused by recording in a noisy or untreated environment.
- Recordings in the ABP Microstudio should not be impacted by background noise, but for improvised recording setups, closing doors and windows and choosing a quiet time might help reduce noise.

Principles of Audio Recording

Reverberation

- Reverberation refers to noise bouncing back off hard surfaces and mixing back into the original signal.
- Reverberation can be mitigated by avoiding recording in large empty spaces.

Principles of Audio Recording

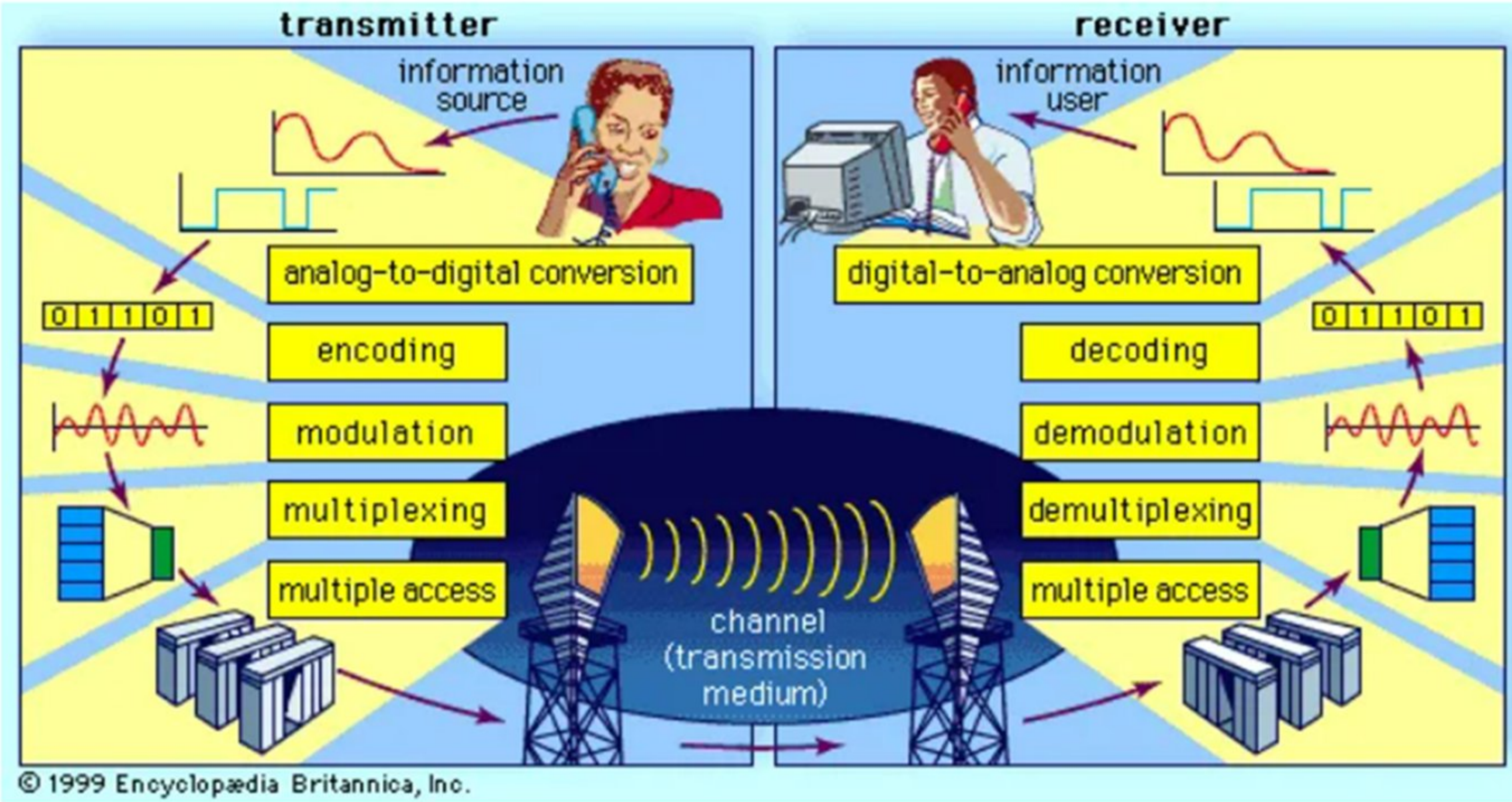
Audio Clipping

- When sounds are too low, they don't get picked up by the microphone, but this also happens on the other end , sounds that are too loud also don't get picked up, and the phenomenon called audio clipping occurs.
- Audio clipping is identifiable by a pop or a crackling sound, and is unfortunately extremely difficult to fix in post production.
- Pop filters can also be fitted to microphones to reduce the risk of clipping.

Principles of Audio Recording

- **Compression**
- Compression is the processing of making an audio file smaller removing some of the information contained in it.
- Uncompressed files usually contain high levels of details which are often not discernable to the human ear, but which can be necessary during post-processing.

ANALOG TO DIGITAL CONVERSION


transmitter
receiver
information source
information user
analog-to-digital conversion
digital-to-analog conversion
0 1 1 0 1
0 1 1 0 1
encoding
decoding
modulation
demodulation
multiplexing
demultiplexing
multiple access
multiple access
channel (transmission medium)
© 1999 Encyclopædia Britannica, Inc.

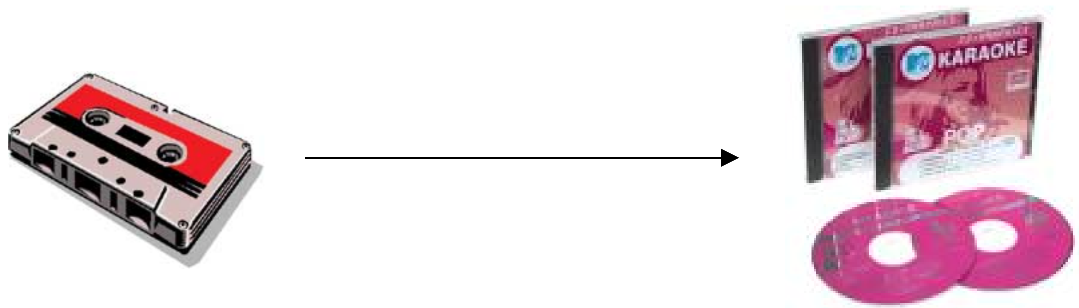
Analogue to Digital Audio

Analogue audio

- The name for an electronic signal that carries its information of sound as continuous fluctuating voltage value.
- Stored in non digital tape or audio tape recording of sound.

<u>Digitizing</u>

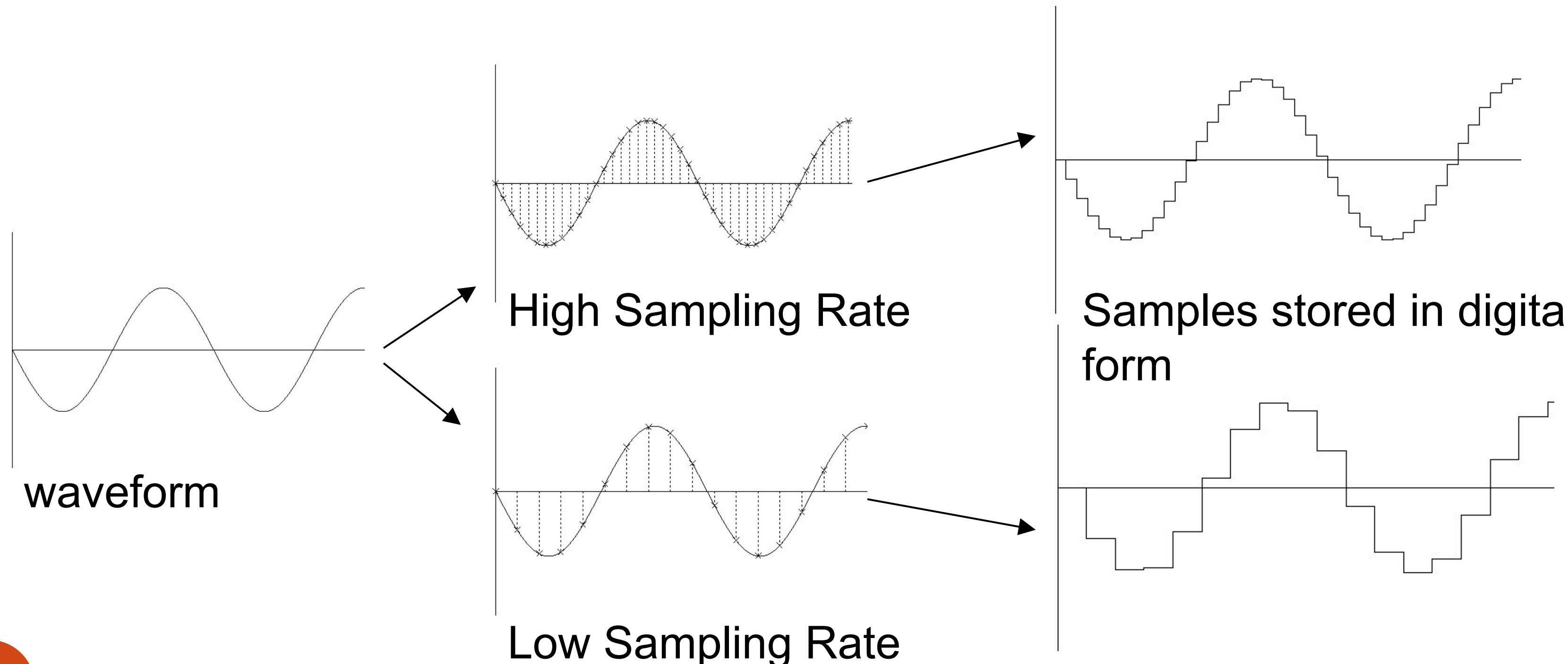
- the process of converting an analog signal to a digital one.



- A sound is recorded by making a <u>**measurement of the amplitude**</u> of the sound at regular intervals which are defined by the <u>"sampling rate“</u> (frequent of sample point taken).
- The process of taking the measurement is called <u>"sampling"</u> and each measurement is called a "sample point".

Digital Audio

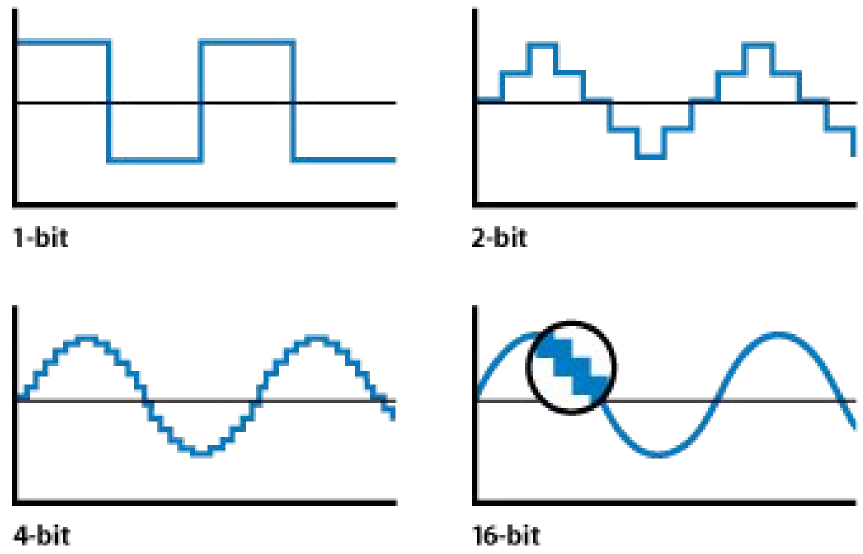
- Digital audio - data are stored in the **<u>form of samples point</u>**.
- Samples represent the amplitude (or loudness) of sound at a discrete point in time.
- Quality of digital recording depends on the ***<u>sampling rate</u>***, <u>the number of samples point taken per second (Hz)</u>.



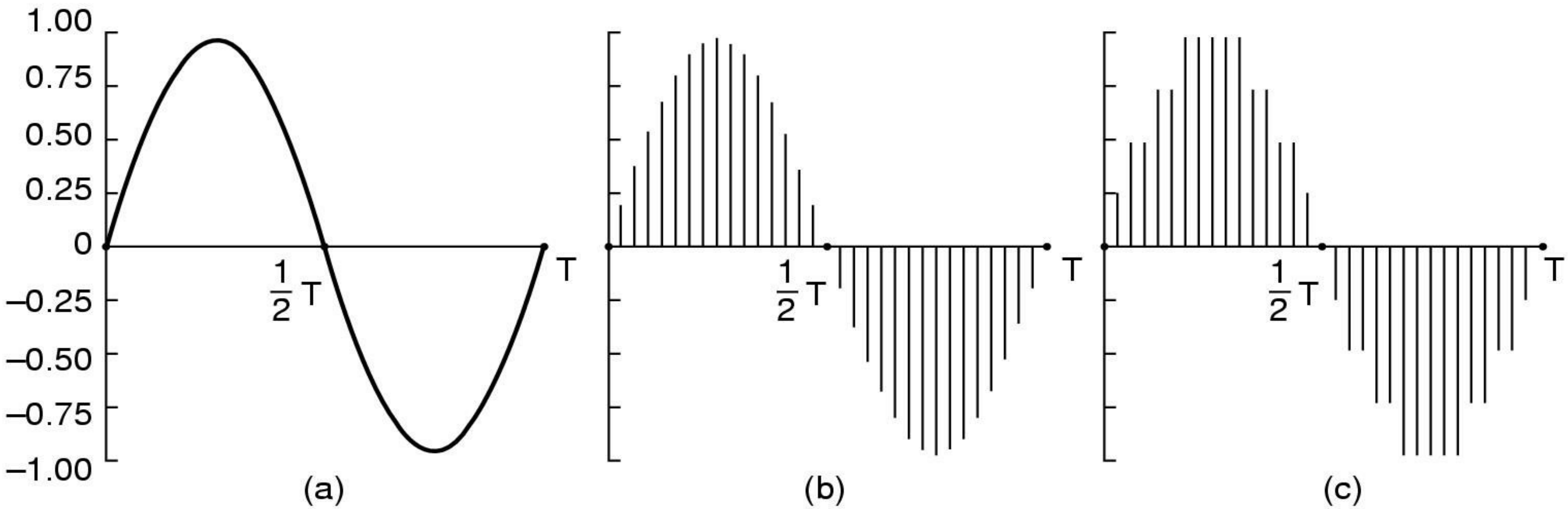
- To represent sound digitally, we must convert this varying voltage into a series of numbers representing its amplitude.

- This process is known as analog-to-digital conversion.
- Audio data consisting of such numbers is said to be in ***pulse code modulation*** format, abbreviated PCM.

- Speech is analog in nature and it is converted to digital form by an analog-to-digital converter (ADC).
- A transducer converts pressure to voltage levels.
- Convert analog signal into a digital stream by discrete sampling
- Discretization both in time and amplitude (quantization)



Audio Encoding Concepts

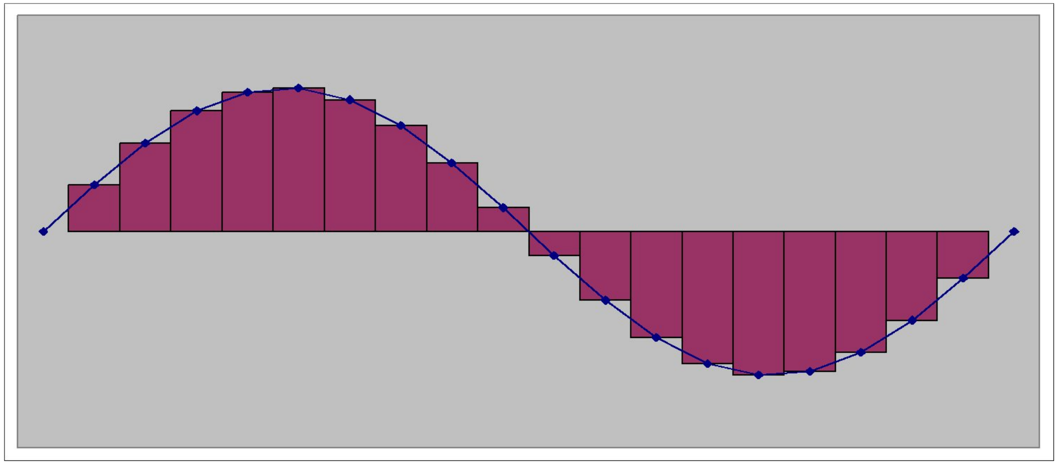


Audio Waves Converted to Digital

- electrical voltage input
- **sample voltage levels** at intervals to get a vector of values: (0, 0.2, 0.5, 1.1, 1.5, 2.3, 2.5, 3.1, 3.0, 2.4,...)
- A computer measures the amplitude of the waveform at regular time intervals to produce a series of numbers (**samples**).
- The ADC process is governed by various factors such as sample rate and quantization: binary number as output

Sample rate and Quantization

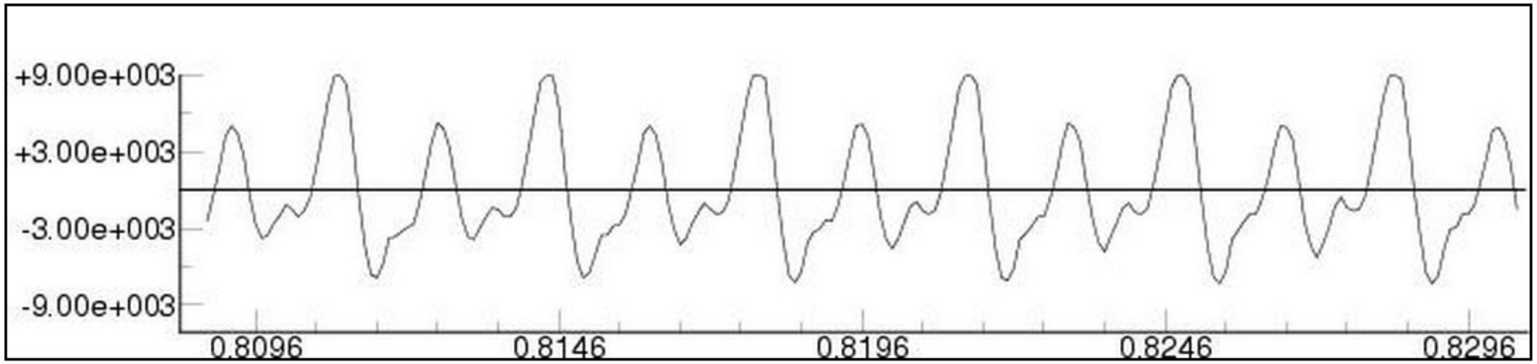
- The ADC captures a snapshot of the electric voltage on an audio line and represents it as a digital number that can be sent to a computer.
- By capturing the voltage thousands of times per second, you can get a very good approximation to the original audio signal.



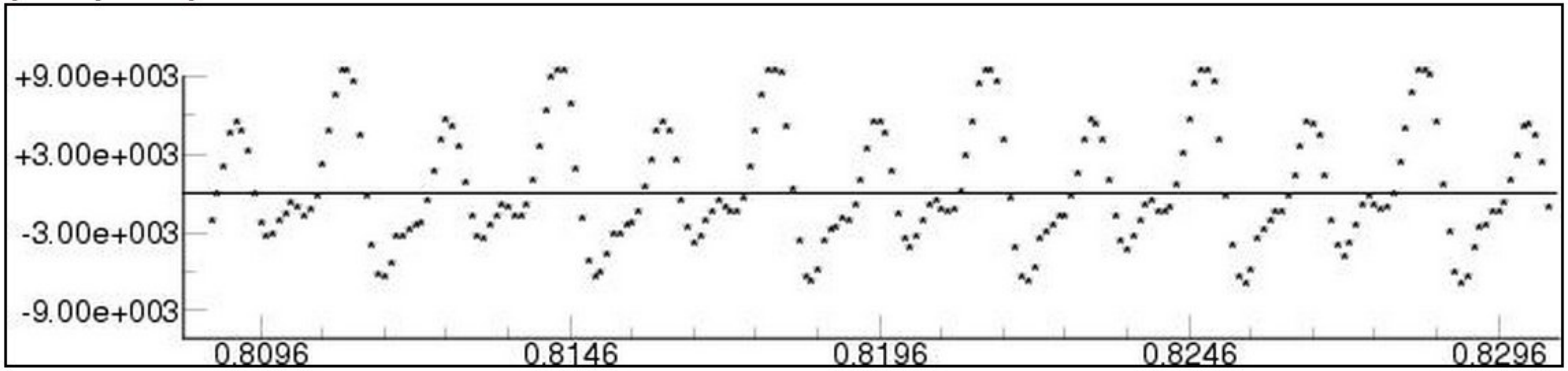
- A digital system such as a computer cannot directly represent a continuous signal.
- Instead, it must measure the signal at a finite set of discrete times. This is known as ***sampling*.**
- Further more, it must make use of a finite number of discrete amplitude levels.
- This is known as ***quantization***. The number of levels used is known as the ***resolution*.** The resolution is usually expressed in *bits*, that is, as the base-2 logarithm of the actual number. A system with a resolution of 8 bits makes use of 2^8 = 256 levels.
- A system with 16 bit resolution makes use of 2^16 = 65,536 levels.

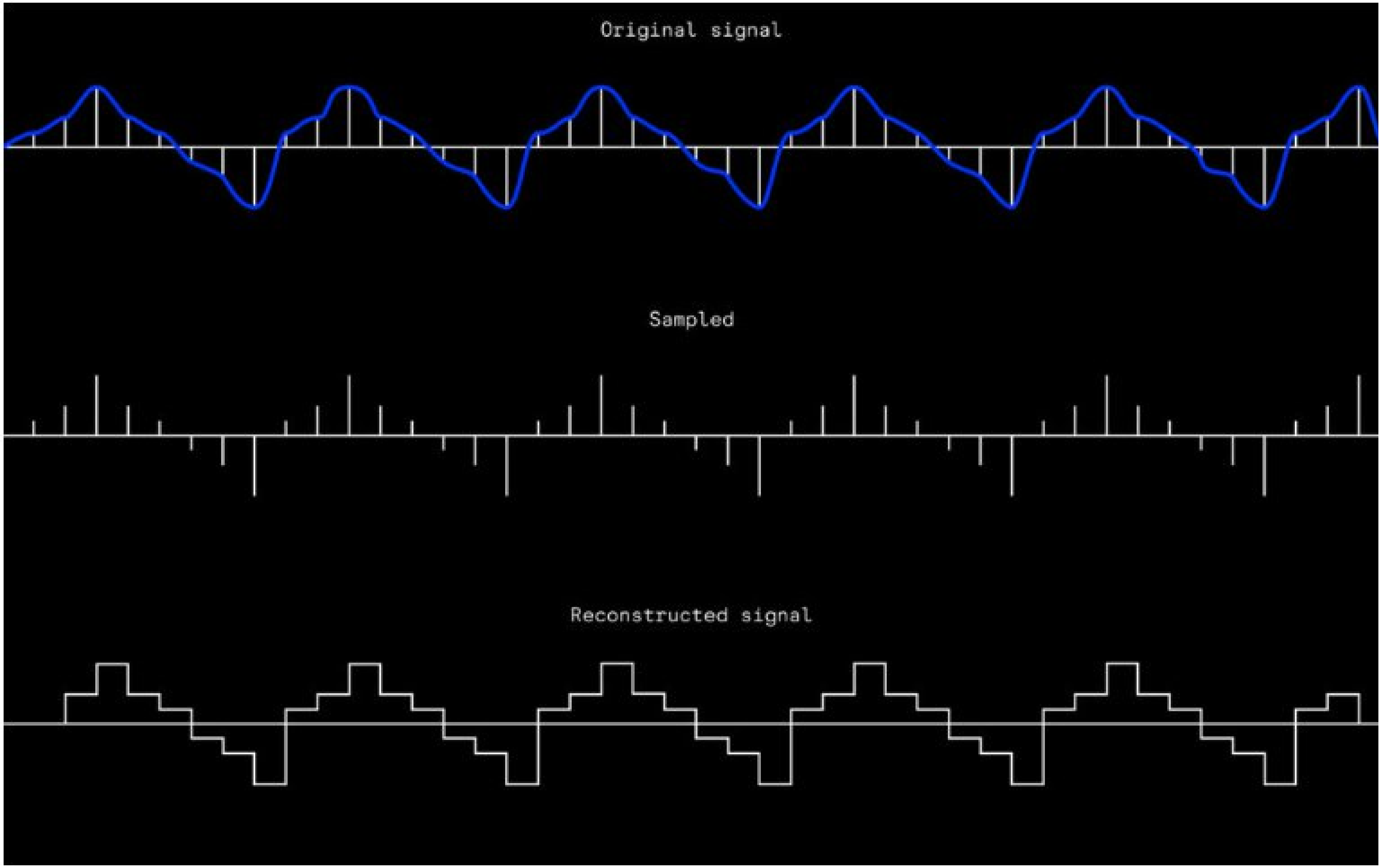
- The sampling rate and resolution determine the quality of the digital representation of the sound. “
- CD-quality" sound has a resolution of 16 bits and a sampling rate of 44,100 samples per second.

Here is a continuous waveform:



And here is a sampled and quantized representation of it:

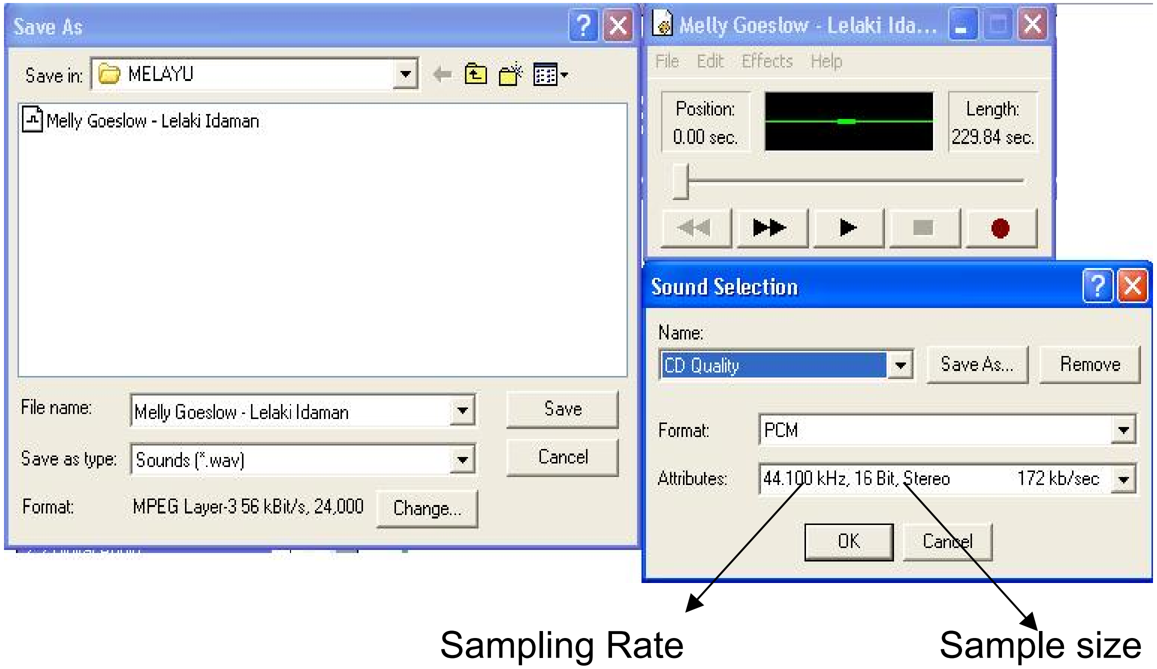




Original signal
Sampled
Reconstructed signal

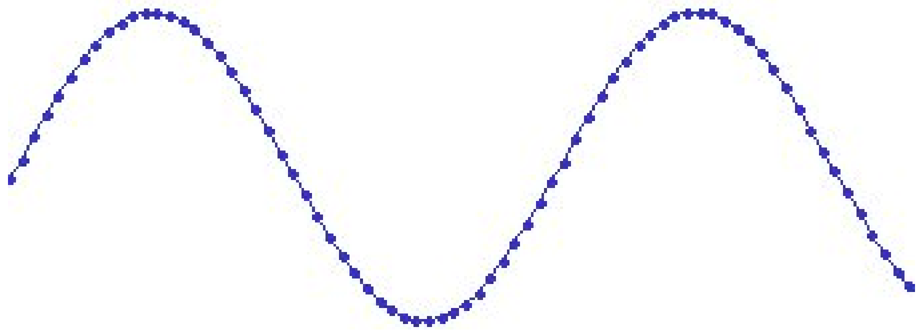
Digital audio sampling

- Quality factors for digital audio file :
  - Sampling Rate
  - Sample Size (resolution) the number of bits used to record the value of a sample in a digitized signal.



- Other than that, it also depends on:
  - The quality of original audio source.
  - The quality of capture device & supporting hardware.
  - The characteristics used for capture.
  - The capability of the playback environment

- **Sampling:** is converting the sound wave to numbers using the magnitude of the wave. That is from analog to digital. How?
- Divide the horizontal axis (the time dimension) into discrete uniform pieces.
- Sample rate (frequency): is how often the samples are taken (fraction of a second)
- Sample size: the amount of information stored about each sample 8 or 16 bits .

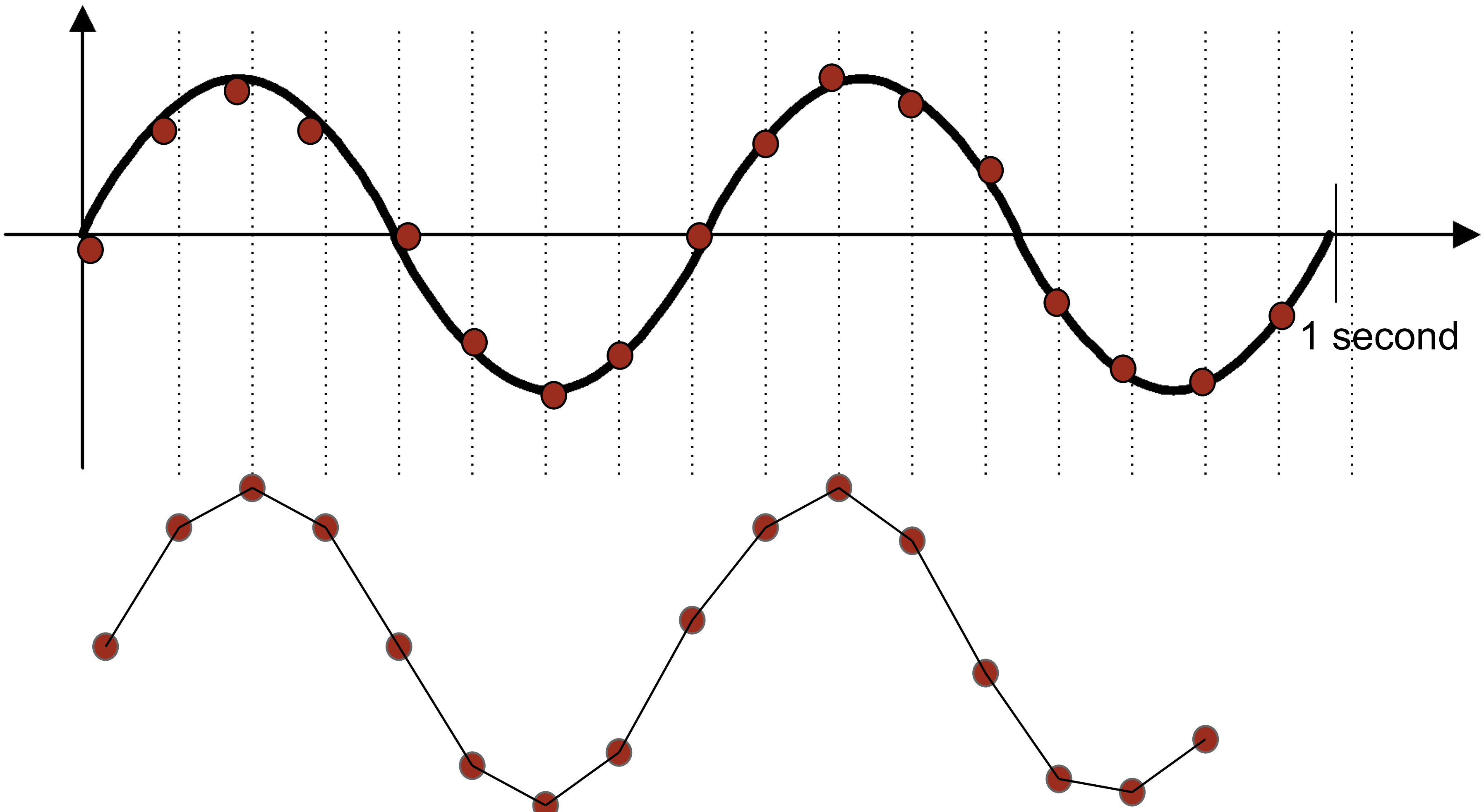


Each dot in the figure above represents one audio *sample*.

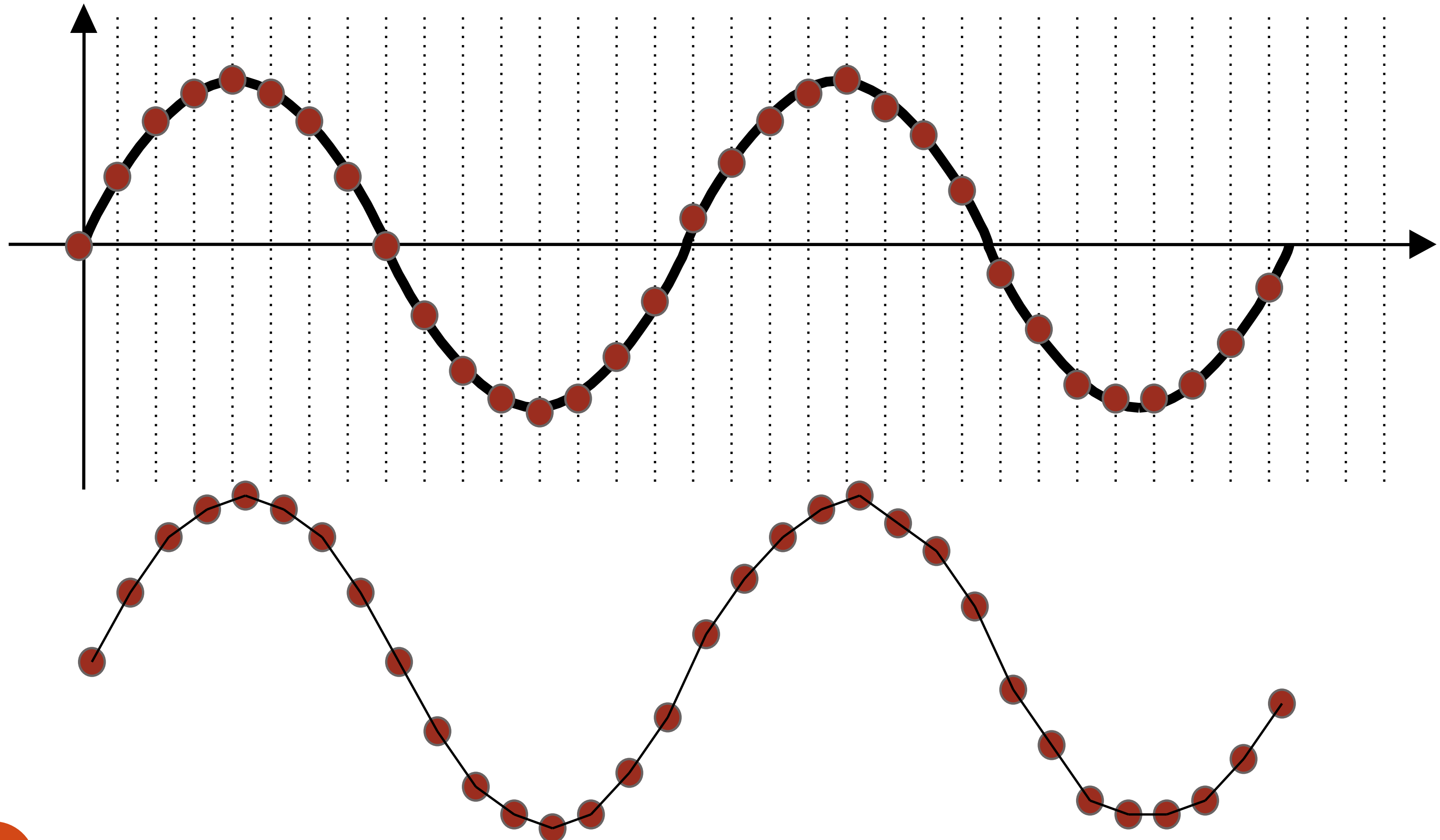
- **Sample rate:** The rate at which the samples are captured or played back, measured in Hertz (Hz), or samples per second. An audio CD has a sample rate of 44,100 Hz, often written as 44 KHz for short. This is also the default sample rate that Audacity uses.

- **Sample format or sample size**: Essentially this is the number of digits in the digital representation of each sample. Think of the sample rate as the horizontal precision of the digital waveform, and the sample format as the vertical precision. An audio CD has a precision of 16 bits, which corresponds to about 5 decimal digits.

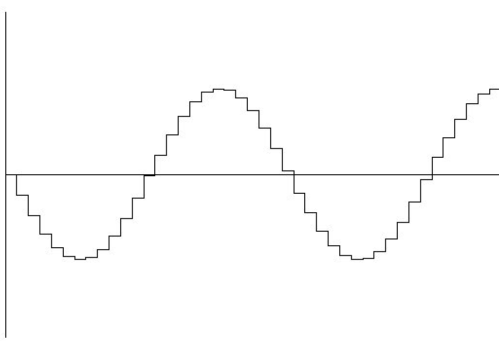
Digital Sampling



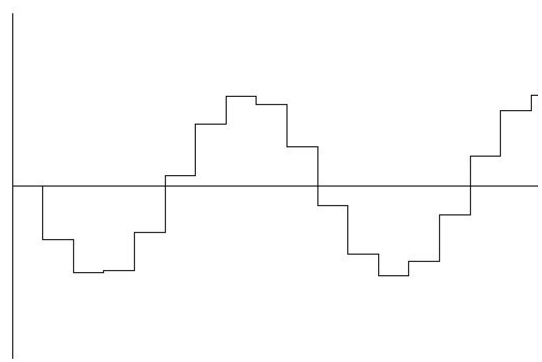
Digital Sampling



- There are three sampling frequencies most often used in multimedia are 44.1 kHz, 22.05 kHz and 11.025 kHz.
  - The **<u>higher the sampling rate</u>**, the **<u>more the measurements</u>** are taken (better quality).
  - The **lower the sampling rate**, the **lesser the measurements** are taken (low quality).
- The number of bits used to describe the amplitude of sound wave when sampled, determines the sample size.



High Sampling Rate



Low Sampling Rate

Sampling Rate

- The sampling rate is the number of times per second that the amplitude of the signal is measured and so has dimensions of samples per second.
- The higher the sampling rate, the more accurately the sampled signal will represent the original signal.

Audio Encoding

- The best-known technique for voice digitization is **Pulse-Code Modulation (PCM)**.
- Voice 4000 Hz
  - What is the PCM sampling rate?
- PCM provides analog samples which must be converted to digital representation. Each of these analog samples must be assigned a binary code. Each sample is approximated by being **quantized** as explained next.

- ***Quantization (sample precision):*** the resolution of a sample value.
- Divide the vertical axis (signal strength) into pieces. Sometimes, a non-linear function is applied.
  - Samples are typically stored as raw numbers (linear PCM format) or as logarithms (u-law or A-law)
  - Quantization depends on the number of bits used measuring the height of the waveform
  - 8 bit quantization divides the vertical axis into 256 levels.
  - 16-bit CD quality quantization results in over 65536 values

Resolution

- The error produced by quantizing a signal is known as *quantization noise.*
- The quality of a quantized signal may be measured by computing the signal-to-noise ratio (SNR), where the noise in question is the quantization noise.
- Each bit of resolution adds approximately 6 decibels to the signal-to-noise ratio.

Digital audio sampling

- How many Samples to take?
  - **11.025 KHz** — Speech (Telephone 8KHz)
  - **22.05 KHz** — Low Grade Audio (WWW Audio, AM Radio)
  - **44.1 KHz** — CD Quality

Digital Audio

- Audio resolution determines the accuracy with which sound can be digitized.
- Size of a monophonic digital recording = sampling rate x (bit resolution/8) x 1.
- Size of stereo recording = sampling rate x duration of recording in seconds x (bit resolution/8) x 2.

What is bit rate?

- It describes the number of bits being produced per second.
- The ratio of the number of bits that are transferred between devices in a specified amount of time, typically one second.
- In other words, it measures how much data is transmitted in a given amount of time.
- Bit rate is of critical importance when it comes to storing a digital signal, or transmitting it across networks, which might have high, low, or even varying bandwidths.
- It is commonly measured in bits per second (bps), kilobits per second (Kbps), or megabits per sec.(Mbps).

$$\text{Bit rate} = \frac{\text{Bits}}{\text{Second}} = \left(\frac{\text{Samples produced}}{\text{Second}}\right) \times \left(\frac{\text{Bits}}{\text{Sample}}\right)$$

$$= \text{Sampling rate x Quantization bits per sample}$$

- Ideally, the bit rate should be just right to capture or convey the necessary information with minimal perceptual distortion, while also minimizing storage requirements. Typical bit rates produced for a few widely used signals are shown in Figure

Signal	**Sampling rate**	**Quantization**	**Bit rate**
Speech	8 KHz	8 bits per sample	64 Kbps
Audio CD	44.1 KHz	16 bits per sample	706 Kbps (mono) 1.4 Mbps (stereo)
Teleconferencing	16 KHz	16 bits per sample	256 Kbps
AM Radio	11 KHz	8 bits per sample	88 Kbps
FM Radio	22 KHz	16 bits per sample	352 Kbps (mono) 704 Kbps (stereo)
NTSC TV image frame	Width – 486 Height – 720	16 bits per sample	5.6 Mbits per frame
HDTV (1080i)	Width – 1920 Height – 1080	12 bits per pixel on average	24.88 Mbits per frame

- They describe the quality of an audio or video file.
- For Ex: MP3 audio file compressed at 192 Kbps has greater dynamic range and sounds slightly more clear than the same audio file compressed at 128 Kbps.
- It is because more bits are used to represent the audio data for each second of playback.
- **Compression:** A method of reducing the size of a digital file. (To be discussed later on in detail)

Raw Data Rate

- Sampling frequency= *f* (Hz)
- Each sample represented by *R* bits
- Raw data rate (bit rate):

- T = f x R (bits per second, or bps)

Digital Audio Signals

- Frequency band of sound: human hearing frequency range: 20Hz-20 KHz.
- Sampling rate > 40 KHz (Actual sampling rate of CD-Audio = 44.1 KHz)
- Bit rate for CD quality audio signal (44.1 KHz, Quantization:16 bits, 2 channels):

- T = 44100 x 16 x 2 (bits per second, or bps)
- CD quality stereo sound → 10.6 MB / min

Calculate audio data size

- The formula to calculate audio data size:
  - C = number of channels (mono = 1 , stereo = 2)
  - S = sampling rate in Hz (cycles per second)
  - T = Time (seconds)
  - B = bytes (1 for 8 bits, 2 for 16 bits)

File Size = C * S * T * B

Exercise: Calculate audio data size

Calculate a 30 seconds 16-bit, 44.1 kHz stereo music

- **Step 1**
  - 44,100 x 2 bytes (or 16-bits) = 88,200 bytes
- **Step 2**
  - 88,200 x 2 (for stereo) = 176,400 bytes
- **Step 3**
  - 176,400 x 30 seconds = 5,292,000 bytes

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)!!!**

Analog to Digital Conversion

1. The first step in converting an analog sound wave into digital audio data is using a type of analog device called a **transducer**.
2. Microphones are transducers because they help change one type of energy into another through a special component called a diaphragm.
3. The diaphragm will begin vibrating in response to incoming sound waves, which creates an alternating electrical current in the microphone

Analog to Digital Conversion

4. This AC voltage is its own analog signal, but one that can be passed through an **ADC**, or an **analog-to-digital converter**.

5. An ADC takes an analog signal that's continuous in both time and amplitude, and translates it into a series of numbers that most accurately represent its amplitude (how loud it is) at very specific points in time

6. It does this by taking a snapshot of the analog signal every so often – this is actually called **sampling,** and the amount of snapshots being taken is determined by **sample rate**.

Analog to Digital Conversion

7. As each snapshot is taken based on the sample rate, the amplitude represented by that incoming electrical current needs to be measured and stored digitally through the process of **quantization**

8. The **bit resolution** (or **bit depth**) determines the range of values that can be used to represent amplitude in digital terms. Higher bit resolution results in less **quantization error**, which occurs when analog amplitudes and their digital representations don't quite match up (this is why audio recorded at 8-bit resolutions can sound quite noisy).

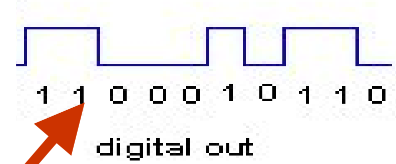
Analog to Digital Conversion

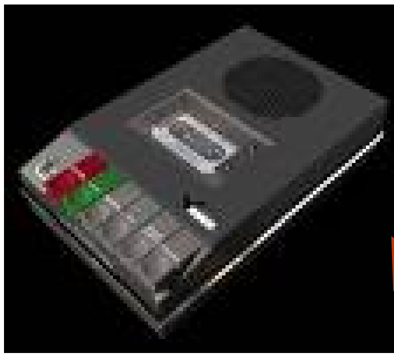
9. Once all of this sampled amplitude data is encoded as a sequence of binary values, the process is complete!

10. We now have a digital representation of an analog signal, which can be copied and transmitted an infinite number of times without any loss in quality.

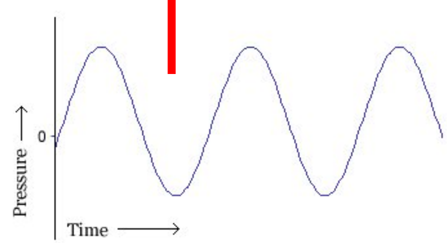
11. This data is commonly stored as an uncompressed audio file like .wav or .aiff

How is Sound Recorded?





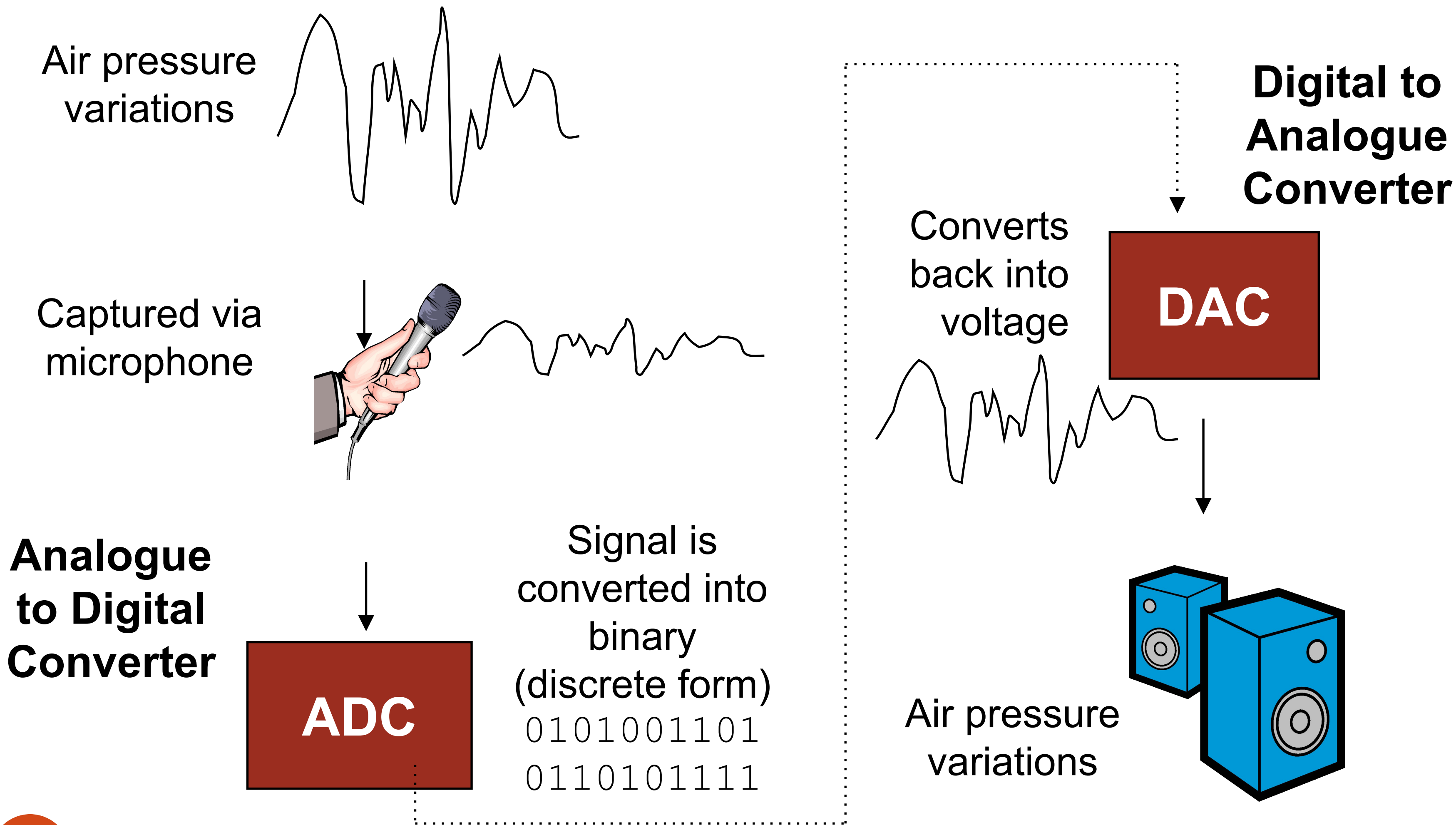
(electrical signal)
analog



- **Microphone** translates movement into electrical signals (analog). Then **tape recorder** translates the waveform from an electrical signal on a wire, to a magnetic signal on a tape (analog)

- **Analog-to-Digital Converter (ADC).** The ADC captures a snapshot of the electric voltage on an audio line and represents it as a digital number that can be sent to a computer.

Capture & Playback of Digital Audio



Recording Audio Files

Recording Audio Files on the pc

Uses either:

i. Microphone

- connect microphone to the microphone port and record using sound recorder

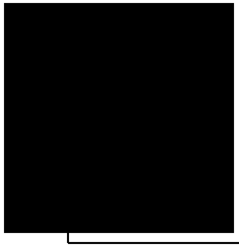
Recording Audio Files

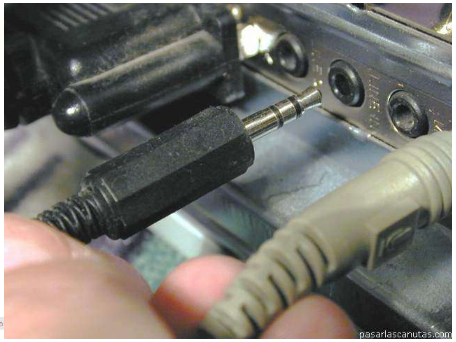
ii. CD-ROM Drive

- Move music files from CD to hard drive or;
- Play the cd and then record using the sound recorder.

iii. Line-in

- pressing play on the audio source, which is connected to the computer’s audio line-in socket. Record using the sound recorder.





Line in port on the pc

Audio cable

- Once a recording had been completed, it almost always needs to be edited.
- Basic sound editing operations include trimming, splicing and assembly, volume adjustments and working on multiple tracks.
- Additional available sound editing operations include format conversion, resampling or downsampling, fade-ins and fade-outs, equalization, time stretching, digital signal processing, and reversing sounds.

How is Sound Recorded and Played?

A sound card is typically equipped to accept input from a microphone and send output to speakers or headphones. For processing waveform files, a sound card contains a special type of circuitry called a **digital signal processor**, which performs three important tasks. It transforms digital bits into analog waves when you play back a waveform audio file. It transforms analog waves into digital bits when you make a sound recording. It also handles compression and decompression, if necessary.

- In computers, an audio card contains a special built-in processor and memory for processing audio files and sending them to *speakers* in the computer.
- An audio file is a record of captured sound that can be played back. Sound is a sequence of naturally analog signals that are converted to digital signals by the audio card, using a microchip called an analog-to-digital converter (ADC).
- When sound is played, the digital signals are sent to the speakers where they are converted back to analog signals that generate varied sound.

Memory Required for 1 Minute of Digital Audio

File Type	44.1 KHz	22.05 KHz	11.025 KHz
16 Bit Stereo	10.1 Mb	5.05 Mb	2.52 Mb
16 Bit Mono	5.05 Mb	2.52 Mb	1.26 Mb
8 Bit Mono	2.52 Mb	1.26 Mb	630 Kb

Preparing digital audio files

- Two important aspects have to be taken in considerations while preparing sounds:

1. Balancing the sound quality with the available hardware resources.
   - Remember the relation between sampling frequency and sound quality?
2. Setting proper recording levels to get clean recording. If Recording level too high, it will introduce noise. Conversely if it is too low it is useless.

QUALITY OF DIGITAL RECORDING

- **DEPENDENT ON**

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

QUALITY OF DIGITAL RECORDING

Quality

Sample rate (kHz)	Resolution (bits)	Stereo/Mono	File size (30 sec)	Comment
11	8	Mono	330K	Marginal quality
22	8	Mono	660K	AM quality
22	16	Stereo	2.6MB	FM quality
44	8	Mono	1.3MB	none
44	16	Mono	2.6MB	none
44	16	Stereo	5.3MB	CD quality

Audio File format

- An audio file format is a file format for storing digital audio data on a computer system.
- It can be a raw bit stream, but it is usually a container format or an audio data format with defined storage layer.

What are codec?

- It comes from the word "**CO**DER-**DEC**ODER".
- A **codec** is a device or computer program capable of encoding and/or decoding a digital data or signal.
- It codes/decodes according to the audio file format specifications.
- For Ex: WAV audio file format is usually coded in the OCM format and Macintosh AIFF audio files.

AudioFormat/Codec

- It is important to distinguish between a file format and an audio codec.
- A codec performs the encoding and decoding of the raw audio data while the data itself is stored in a file with a specific audio file format.
- In other words, Codec contains both an ADC and DAC running off the same clock.

3 Categories of Audio Files

There are 3 categories in which certain Audio files belong to:-

- **Uncompressed**: Audio files that are not compressed and are capable of having a large file size. Ex) .Wav
- **Lossless:** Audio files that are compressed but doesn't lose any quality to the file. Ex) .WMA
- **Lossy:** Audio files that lose some quality when being compressed. Ex) .Mp3

More audio file formats:

- **Uncompressed formats**

  wav & aiff

- **Lossless compression**

  FLAC, lossless Windows Media Audio (WMA). ALAC

- **Lossy compression**

  MP3, Windows Media Audio (WMA) and AAC.

Uncompressed formats

- They are often referred to as **Pulse-code modulation** (PCM formats)
- As the name suggests these formats are not compressed nor do use compression.
- It means all the data is available, at the risk of large file sizes.
- For Ex:
  - Wav
  - AIFF

Compression

- The need to compress media data is motivated by both storage requirements and transmission requirements.
- Audio data takes up a lot of space, at least in comparison with text. A single second of compact disc audio takes up about as much space as 15,000 words of ASCII text, that is, 60 pages of a typical book. Here is a chart showing the amount of space occupied by different durations of monaural sound at different sampling rates. A 10GB disk, for example, will hold only about 31 hours of audio at the CD-rate.

	1 second	1 minute	1 hour
44,100 samples/second 16 bit	188.2KB	5.3MB	317.5MB
22,050 samples/second 16 bit	44.1KB	2.6MB	158.8MB
16,000 samples/second 16 bit	32.0KB	1.9MB	115.2MB

Data Compression

- Data compression requires the identification and extraction of source redundancy.
- In other words, data compression seeks to reduce the number of bits used to store or transmit information.
- Compression techniques are of two basic types: *lossless* and *lossy*. A *lossless* compression technique is one that yields a compressed signal from which the original signal can be reconstructed perfectly. No information is lost as a result of the compression.
- A *lossy* compression technique is one that discards information. The original signal cannot be reconstructed perfectly from a signal compressed by a lossy method.

Codecs

- A program or hardware device that compresses and decompresses data is known as a *codec*, short for "**co**mpressor - **dec**ompressor".
- Codecs are compression technologies and have two components, an **encoder** to compress the files, and a **decoder** to decompress.
- The term is a blend of the words *coder* and *decoder*, as well as *compression* and *decompression*.

Codecs

- codecs compress -- or shrink -- media files such as video, audio and still images in order to save device space and to efficiently send those files over a network such as the internet.
- Codecs are made up of an encoder and decoder. The encoder compresses a media file, and the decoder decompresses the file.

Codecs

- Codecs are invisible to the end user and come built into the software or hardware of a device. For example, Windows Media Player, which comes pre-installed with every edition of Windows, provides a limited set of codecs that play media files.
- Codecs serve a major purpose as, without them, media files would take up much more storage space

Lossless Compression

- Lossless compression techniques achieve compression by removing the redundancy in the signal.
- This is normally achieved by assigning new codes to the symbols based on the frequency of occurrence of the symbols in the message.
  - More frequent symbols are assigned shorter codes and vice versa.
  - The lossless compression algorithms make use of the probabilistic models to compute efficient codes for the symbols.
- Lossless compression techniques are not widely used because the amount of compression that they produce is relatively small.

Lossless Audio Formats

- A lossless compressed format stores data in less space by eliminating unnecessary data.
- Uncompressed audio formats encode both sound and silence with the same number of bits per unit of time.
- In a lossless compressed format, the music would occupy a smaller portion of the file and the silence would take up almost no space at all
- Lossless compression formats enable the original uncompressed data to be recreated exactly.

Lossless Compression

- The degree of compression obtained depends on the content of the file. With speech, lossless compression reduces the size of the file at best to about 25% of its original size, at worst to about 50%. Quiet classical music compresses almost as well as speech, while "noisy" modern music, tends to compress poorly, often to about 75% of its original size.
- At present, the main users of lossless compression appear to be fans of recordings of live concerts. FLAC ("free lossless audio codec)
- In areas such as phonetics research, the use of lossless compression is desirable. Those generating audio data should consider using one of the lossless techniques if they are going to compress at all.

Lossless Compression

- Lossless compression can recover the exact original data after compression.
- They provide a compression ratio of about 2:1.
- It is used mainly for compressing database records, spreadsheets or word processing files, where exact replication of the original is essential.
- Examples: Run Length Encoding (RLE), Lempel Ziv Welch (LZW), Huffman Coding.

What is Lempel–Ziv–Welch (LZW) Algorithm ?

- LZW compression works by reading a sequence of symbols, grouping the symbols into strings, and converting the strings into codes.
- Because the codes take up less space than the strings they replace, we get compression

Lossless compression

- It applies compression to an uncompressed audio file, but it doesn’t lose information or degrade the quality of the digital audio file.
- It produce exactly the same digital signal as the original audio source.
- For Ex:
  - FLAC
  - WMA
  - ALAC

Lossy compression

- It will result in some loss of data as the compression eliminates unnecessary information.
- It basically it removes what it sees as irrelevant information.
- Lossy compression typically achieves far greater compression but somewhat reduced quality than lossless compression by simplifying the complexities of the data.
- A variety of techniques are used, mainly by exploiting psychoacoustics, to remove data with minimal reduction in the quality of reproduction

- It has a small file size but poor sound quality.

Lossy compression

- There are numerous lossy compression techniques, most of which are now rarely encountered.
- Two lossy compression techniques are of some importance: minidisc and mp3.
- Minidisc compression is important because minidisc recorders have been used for the collection of linguistic data. MP3 compression is important because a great deal of audio is distributed in this form.
- For Ex:
  - MP3
  - WMA
  - AAC.

Lossy Compressed Audio Formats

- Most formats offer a range of degrees of compression, generally measured in bit rate.
- The lower the rate, the smaller the file and the more significant the quality loss.
- Examples… MP3, Vorbis (.ogg ,.oga), Musepack (.mpc, .mp+, .mpp),AAC(.m4a, .m4b, .m4p, .m4v, .m4r, .3gp, .mp4, .aac), ATRAC (.aa3,.oma,.at3),Windows Media Audio Lossy (WMA lossy)(.wma)

Compression Ratio

- Compression ratio

$$\frac{\text{original data size}}{\text{compressed data size}} : 1$$

Different File Formats and its packaging

Common File Types

- Here we describe the most common audio file types. The discussion here will also give the reader a good general idea of the organization of audio files.
  - Raw Sound Files
  - AU/SND Files
  - WAVE Files
  - AIFF Files
  - MP3 Files
  - OGG Files
  - RAM Files

Raw Sound Files

- Sound files that consist of nothing but PCM audio data are called *raw* sound files.
- Some audio i/o devices, especially older devices intended for research rather than the commercial market, produce such files. They are no longer commonly seen.
- Since raw sound files have no header in which to store information, it is necessary to know their sampling rate, resolution, signedness, and number of channels.

Raw Sound Files

- Filename suffixes are sometimes used to convey the resolution and signedness.
- For example, the suffix *.sb* is likely to indicate that samples consist of one byte, that is, have a resolution of 8 bits, and are signed.
- The suffix *.uw* in this system indicates that each sample is represented by a two-byte word, that is, has a resolution of 16 bits, and is unsigned.

Signedness

- The integers used to represent amplitude values may be *signed* or *unsigned*.
- A signed number is one that may be either positive or negative.
- An unsigned number may never be negative.
- Whether the numbers used are signed or unsigned has no effect on the resolution.
- The number of distinct amplitude levels remains the same.

Wave (.wav)

- WAV was developed jointly by Microsoft and IBM.
- They have complex format and are more common on Windows-based systems.
- Commonly used for storing uncompressed CD-quality sound files
- Offer the best quality.
- But these files are very large (10 MB / min)
- Hence they are unsuitable for everyday exchange the internet.

- **Advantage-** WAV format is easy to transform and compress into MP3 or other formats.

Aiff (.aiff or .aif)

- Audio Interchange File Format.
- It similar to .wav and has high quality sound.
- The format was co-developed by Apple Computer in 1988 based on Electronic Arts' Interchange File Format
- Commonly used on Apple Macintosh computer systems.
- **Disadvantage**-not supported by all web browsers

AIFF Files

- AIFF format is widely used on Apple computers and, as a result, in professional audio processing software. Like, RIFF/WAVE, AIFF ("Audio Interchange File Format") is a derivative of the Interchange Format Files format developed by Electronic Arts.
- It is simpler than RIFF/WAVE format in that it is intended only for audio data and supports a smaller range of audio data formats.
- An AIFF file consists of a header followed by one or more "chunks". A minimal AIFF sound file therefore consists of a header and a sound chunk.
- In addition to sound chunks, a variety of other chunks are possible, including markers of positions in the waveform data, and comments.
- The audio data in an AIFF file is always uncompressed PCM. The header contains information about the number of channels, sampling rate, and resolution. The audio data, like all integer data in this format, is stored in big-endian format.

MP3 (.mp3)

- **M**otion **P**icture Experts Group 1, Layer **III** or MPEG Audio layer III
- Designed by the Moving Picture Experts Group.
- MP3 was introduced in 1992
- It is the most successful audio-standard since WAV.
- It compresses an ordinary music-CD to 1/10 of it's original size - thus 12 hours of music could be stored on a recordable CD
- Several bit rates are specified in the MP3: 32, 40, 48, 56, 64, 80, 96, 112, 128, 144, 160, 192, 224, 256 and 320 kbit/s

Windows Media Audio (WMA)

- It is an audio data compression technology developed by Microsoft. The name can be used to refer to its audio file format or its audio codecs. It is a proprietary technology that forms part of the Windows Media framework.

- WMA consists of four distinct codecs.
  - The original WMA codec, known simply as *WMA*, was conceived as a competitor to the popular MP3 and RealAudio codecs.
  - *WMA Pro*, a newer and more advanced codec, supports multichannel and high resolution audio.
  - A lossless codec, *WMA Lossless*, compresses audio data without loss of audio fidelity (the regular WMA format is lossy).
  - *WMA Voice*, targeted at voice content, applies compression using a range of low bit rates.

Container format

- A WMA file is in most circumstances contained in the Advanced Systems Format (ASF), a proprietary Microsoft container format for digital audio or digital video.
- The ASF container format specifies how metadata about the file is to be encoded, similar to the ID3tags used by MP3 files.
- Metadata may include song name, track number, artist name, and also audio normalization values. This container can optionally support digital rights management (DRM) using a combination of elliptic curve cryptography key exchange, DES block cipher, a custom block cipher, RC4 stream cipher and the SHA-1 hashing function.

WMA Container format

- Since 2008 Microsoft has also been using WMA Professional in its Protected Interoperable File Format (PIFF) based on the ISO Base Media File Format and most commonly used for Smooth Streaming, a form of adaptive bit rate streaming over HTTP.
- Related industry standards such as DECE UltraViolet and MPEG-DASH have not standardized WMA as a supported audio codec, deciding in favor of the more industry-prevalent MPEG and Dolby audio codecs.

Different Formats: An Comparison in nutshell

- **WAV** and **AIFF**: Both WAV and AIFF are uncompressed formats, which means they are exact copies of the original source audio.
- The two formats are essentially the same quality; they just store the data a bit differently.
- AIFF is made by Apple, so you may see it a bit more often in Apple products, but WAV is pretty much universal.
- However, since they're uncompressed, they take up a *lot* of unnecessary space. Unless you're editing the audio, you don't need to store the audio in these formats.

Streaming audio

What If…?

- Instead of having to wait until a sound file was fully downloaded, the file could begin playing while it was still downloading?
- Well, that's what **streaming audio** (and video) is all about.

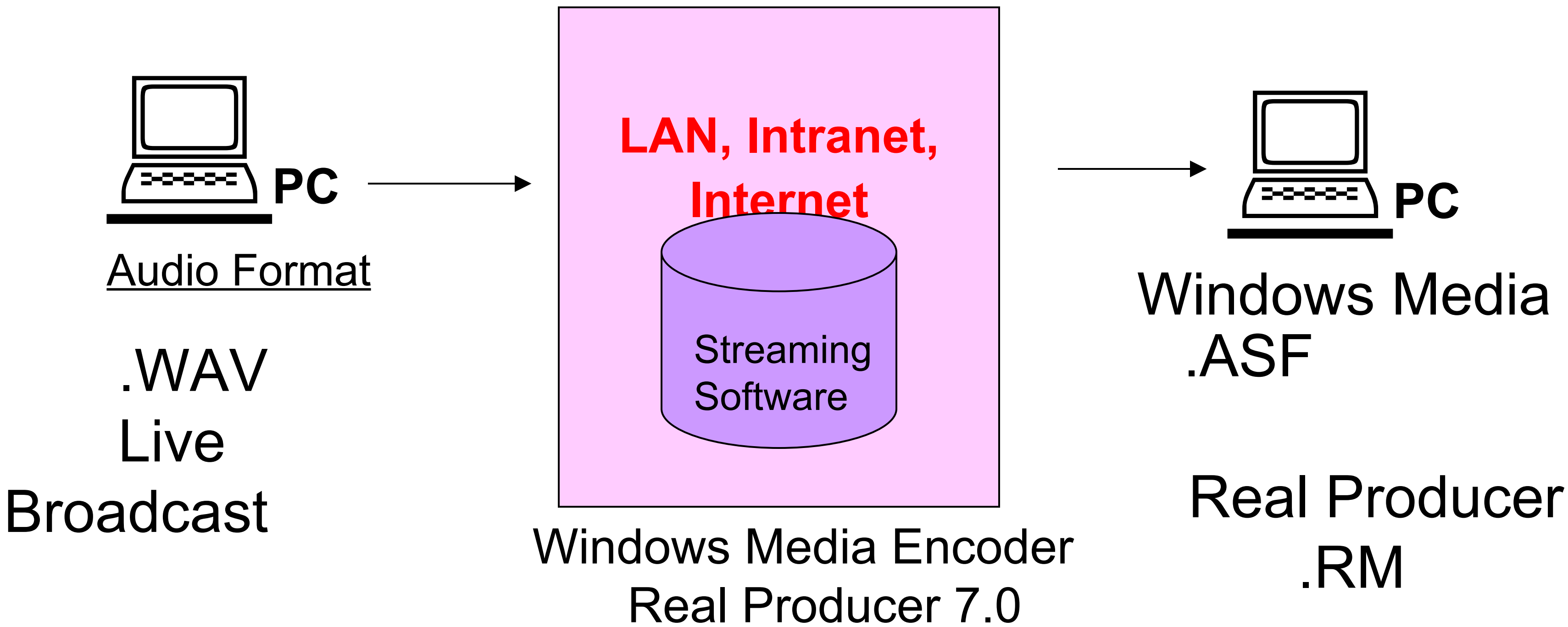
So what is *Streaming* audio?

- It's a method of delivering an audio signal to your computer over the Internet, and differs from the "normal" method of receiving Internet audio in one important way: instead of having to download a ".wav", ".au" or other type of file completely before being able to listen to it, you hear the sound *as it arrives* at your computer, and therefore do not have to wait for a complete download (which would be difficult with a live broadcast anyway!).
- As the data arrives it is buffered for a few seconds and then playback begins. As the audio is playing, more data is constantly arriving (or streaming), and as long as you are receiving a constant stream of data, you should hear constant audio.

How it works…?

- With streaming audio, the sound file is downloaded into a buffer, then sent to the sound card.
- The file starts playing a few seconds after downloading begins, and if the Web doesn't bog down too badly, will keep playing more or less smoothly to the end.

Audio Streaming Concept



.ASF stands for Advanced Systems Format (formerly **Advanced Streaming Format**,
Active Streaming Format)

Audio Streaming Concept cont.

- Streaming audio is a technology which has revolutionized how Internet users define "multimedia."
- Prior to 1995, for a person surfing the web to hear audio, he or she had to download a huge audio file before being able to hear even a few minutes of speech or music.
- The transfer ratio of an audio file is typically 5:1, meaning for every minute of audio, the user has to wait 5 minutes for that one-minute segment to download.
- To hear a 30-minute audio program, the listener would have to wait over 2 hours, a delay which is not only unreasonable, but expensive if Internet access is charged by the hour.
- Even if a person was willing to wait for the amount of time needed to download the file, any technical difficulty during transfer would mean that the entire file would be lost. Also, the amount of available hard drive space would also pose a constraint, since audio files are so large.

- All of these factors inhibited the widespread usage and availability of a true multimedia experience.
- In 1995, Progressive Networks (later renamed Real Networks) introduced a new technology which solved many of the existing problems with audio. Streaming audio enables the listener to hear sound in real time: the user is listening to the sound file as it is being transmitted instead of having to wait for the file to download first.
- Since the sound data is sent in "packets" over the network, each set of data is being heard immediately as it sent.
- The user does not have to store the information on the hard drive or risk more than a waiting a substantial amount of time only to lose the file; even if a packet of information is lost, the rest of the audio file is intact and can still be played back by the listener.

How They Work?

- Streaming audio uses codec (coder/decoder) technology to compress the audio file for distribution across the Internet.
- Once the original audio file has been digitized, it is encoded using a vendor-specific algorithm which compresses the data and allows efficient transmission over the network.
- The audio data is then sent across Internet lines using either UDP, TCP, or IP Multicasting as a protocol to communicate the information.
- As the data is received, the browser uses a helper application to play back the file, buffering the first portion of the data. Buffering allocates a portion of memory on the user's hard drive to store a few packets of video or audio information.

How They Work?

- As each buffer is played, new data is received from the server.
- This way, the player always can get information from the buffer instead of waiting to receive data from the server.
- Streaming technology allows the user to skip forward or back to any portion of the file, since the transmission is "bidirectional": in addition to the server transmitting information to the player, the player can also transmit information to the server, requesting a specific audio packet of data.

Audio Streaming Concept cont.

- Streaming audio technologies relies on:
  - Sound sequences
  - Compression schemes
- Compression schemes (encoding) decreases the audio’s bandwidth requirements:
  - Lowering the audio’s sampling rate
  - Filtering high frequencies

Audio Streaming Advantages

- Real time audio content.
- Low bandwidth media used.
- No waiting for downloading audio file.
- Internet users can enjoy a live online program.

Limitations of Streaming Technology

- Audio and video quality is influenced by a variety of factors. The smoothness of playback can be judged to some degree by the amount of "packet loss."
- Since both audio and video information is streamed over the Internet in small packets, a lost packet means that the set of data contained in that packet needs to be re-sent by the server, causing small gaps in the file being played back to the user.
- The amount of packet and overall quality of sound is determined by three factors: connection speed, the underlying transmission protocol, and the compression algorithm used by the vendor.
- A faulty router between the two may cause the stream to drop, as can inadequate bandwidth from either machine. There are just so many possibilities; it would send a sane person into a screaming fit thinking of them all.

Where can Audio Streaming Applied

- You can include streaming files in a Web page the same way you include any audio file. The catch is that your Web server must have the RealAudio extensions installed, and many hosting services charge an extra fee for this.
- Streaming audio also makes it possible to broadcast live or prerecorded music over the Internet on a continuous basis. Internet "radio stations" can broadcast to a worldwide audience, and therefore can specialize in a particular type of music, as no geographically-limited station possibly could.

Audio Streaming Applications

- Long-distance or automated training
- Seminars
- Concerts
- Speeches
- Music samples
- Online corporate messages
- Hear the news / Radio

Audio Streaming Products

- Window Media Technologies (Microsoft)
- RealSystem G2 (RealNetworks)
- Shockwave Streaming Audio (Macromedia)
- IBM Bamba (IBM)
- Streamworks (Xing Technology)
- Media Player (Netscape)

PART E Audio Effects

Introduction

- **Sound effects** (or **audio effects**) are artificially created or enhanced sounds, or sound processes used to emphasize artistic or other content of films, television shows, live performance, animation, video games, music, or other media.
- In motion picture and television production, a sound effect is a sound recorded and presented to make a specific storytelling or creative point *without* the use of dialogue or music.
- The term often refers to a process applied to a recording, without necessarily referring to the recording itself.

Introduction

- In professional motion picture and television production, dialogue, music, and sound effects recordings are treated as separate elements.
- Dialogue and music recordings are never referred to as sound effects, even though the processes applied to them, such as reverberation or flanging effects, often are called "sound effects".

Sound Effects

- A sound other than speech or music made artificially for use in a play, film, or other broadcast production.

The function of sound effects is three fold:

- Simulating reality
- Creating illusion
- Mood

Simulating reality

In a western barroom fight our hero is hit over the head with a whiskey bottle

- The bottle is fake. It becomes real with the addition of an actual glass bottle crash from the sound editors library.
- In gun battles the weapon actually is actually loaded with blanks and what is called quarter loads which means one-fourth of normal amount of gunpowder contained in a real bullet.
- The actual sound is just slightly louder than a cap pistol until the sound editor has completed work.
- These are but two of the more obvious examples of the sound effect taking a fake bit of theatrics and making it real by adding a real sound.

You see it - you hear it - you must believe it!

Creating Illusion

- Creating illusion was one of the biggest additions to the art of film by sound.

- Eg: 1
- *A man and a woman walk into a café. Several other people are sitting at various table in deep conversation. The main couple sits at a table and a series of close ups for their conversation are presented*

- By adding the sound of the off-scene diners the audience is convinced that they are still in the café. Obviously, the producer does not want to pay a group of extras to sit off camera. The sound editor places them there with his crowd walla for the sound (Walla is an industry term for the sound of people talking without hearing specific words)

Eg 2

A woman is sitting in her living room. The door opens and her husband walks into the room.

- With the addition of a few sound effects, it is possible to inform the audience that he has driven up to the house, parked his car, walked to the door, and used his key to unlock the door.
- None of this was shot. It was an illusion created with effects.

Eg 3.

A safari makes it through the jungle.

- The sound editor cuts a lion roar.
- Not has he placed a lion in the film where none exists but he has also placed the safari in danger.
-

Mood

T*he leading man drives up to a house.*

- As he parks, we hear the sound of a small dog yapping.
- No particular danger is perceived. Inside is probably a child or an old lady.
- Change the small dog yapping to the sound of a vicious Doberman, and the mood is again changed.

Disciplines of Effects

ISOLATED SOUND

- Isolated sounds include the sounds of everyday items like doorbells, car horns and telephone rings. These are the real and specific sounds that you hear such as dogs barking, guns firing, doors slamming and car tires screeching during a car chase.

SPECIALTY EFFECTS

- Specialty effects include sounds that are designed to be used for objects and places that exist only in our imaginations. You might use special effects in cartoon, fantasy and science fiction productions. They would enhance the production by creating a special sound for a space transporter, the magic of a genie's brass lamp, or for gigantic purple people eater cartoon animals.

FOLEY SOUNDS

Jack Donovan Foley

- Foley sounds are synchronized with the visuals in a motion picture or television program.
- This synchronization process is also called audio sound replacement.
- The magic of Foley places footsteps into a film soundtrack – recreates the rattle of a coffee cup being put down on a table – provides the realistic rustle of clothing and the ever popular punch in the face.
- These and many other sounds must be dubbed during post production on the film's soundtrack in order to be heard by the audience.
- They are named after Jack Foley – a pioneer in the sound replacement field.

BACKGROUND AMBIENCES

- Background ambience tracks provide the basic environment for a production. They provide all of the subtle atmosphere that makes the film-goer or radio listener really feel like he is in a courtroom, a church, a train station, a thunderstorm or a jungle.

Discussion on some Audio Effects

Equalization

- *Equalization* means boosting or reducing the levels of various frequencies in a signal. At it's most basic, equalization can mean turning the bass/treble controls up or down. Advanced equalizers have fine controls for specific frequencies.
- Moderate use of equalization (often abbreviated as "EQ") can be used to "fine-tune" the tone quality of a recording; extreme use of equalization, such as heavily cutting a certain frequency can create more unusual effects.
- Common uses for equalization include correct signals which sound unnatural and reducing feedback.

Filtering

- Equalization is a form of filtering. In the general sense, frequency ranges can be emphasized or attenuated using low-pass, high-pass, band-pass or band-stop filters. Band-pass filtering of voice can simulate the effect of a telephone because telephones use band-pass filters.

Compression & Limiting

- *Compression* means reducing the dynamic range of a signal. All signal values above a certain adjustable threshold are reduced in gain relative to lower-level signals. This creates a more even signal level, reducing the level of the loudest parts.

- Level compression is not to be confused with audio data compression, where the amount of data is reduced without affecting the amplitude of the sound it represents.

- *Limiting* is an extreme form of compression. Rather than smoothly reducing the gain of successively higher levels, all signal above the threshold is limited to the same gain. This creates a very hard cut-off point, over which there is no increase in level.

Expansion & Noise Gating

- *Expansion* means increasing the dynamic range of a signal. High level signals maintain the same (or nearly the same) levels, low level signals are reduced (attenuated). This creates a greater range between quiet and loud. Expansion is the opposite of compression.
- *Noise gating* is an extreme form of expansion — signals below a certain point are either heavily attenuated or eliminated completely. This leaves only higher level signals and removes background noise when the signal is not present.

Delay / Echo

- *Delay* is a simple concept — the original audio signal is followed closely by a delayed repeat, just like an echo. The delay time can be as short as a few milliseconds or as long as several seconds. A delay effect can include a single echo or multiple echoes, usually reducing quickly in relative level.
- Delay also forms the basis of other effects such as reverb, chorus, phasing and flanging.
- *echo* - to simulate the effect of reverberation in a large hall or cavern, one or several delayed signals are added to the original signal.
- To be perceived as echo, the delay has to be of order 35 milliseconds or above. Short of actually playing a sound in the desired environment, the effect of echo can be implemented using either digital or analog methods.

Reverb

- *Reverb* is short for *reverberation*, the effect of many sound reflections occurring in a very short space of time. The familiar sound of clapping in an empty hall is a good example of reverb.
- Reverb effects are used to restore the natural ambience to a sound, or to give it more fullness and body.

Reverse echo

- a swelling effect created by reversing an audio signal and recording echo and/or delay whilst the signal runs in reverse.
- When played back forward the last echos are heard before the effected sound creating a rush like swell preceding and during playback.

Chorus

- delayed signal is added to the original signal with a constant delay. The delay has to be short in order not to be perceived as echo, but above 5 ms to be audible. If the delay is too short, it will destructively interfere with the un-delayed signal and create a flanging effect. Often, the delayed signals will be slightly pitch shifted to more realistically convey the effect of multiple voices.

- It is designed to make a signal sound like it was produced by multiple similar sources. For example, if you add the chorus effect to a solo singer's voice, the results sounds like.... a chorus.

Phasing

- *Phasing*, AKA *phase shifting*, *is* another way of creating an unusual sound, a sweeping, whooshing effect often used in music.
- The effect is created by mixing the original signal with another version of itself which has been phase-shifted. This results in various out-of-phase interactions over time which gives the sweeping effect.
- The signal is split, a portion is filtered with an all-pass filter to produce a phase-shift, and then the unfiltered and filtered signals are mixed.
- The phaser effect was originally a simpler implementation of the flanger effect since delays were difficult to implement with analog equipment. Phasers are often used to give a "synthesized" or electronic effect to natural sounds, such as human speech.
- The voice of C-3PO from Star Wars was created by taking the actor's voice and treating it with a phaser.

Flanging

- To create an unusual sound, a delayed signal is added to the original signal with a continuously variable delay (usually smaller than 10 ms).
- This effect is now done electronically using DSP, but originally the effect was created by playing the same recording on two synchronized tape players, and then mixing the signals together.
- As long as the machines were synchronized, the mix would sound more-or-less normal, but if the operator placed his finger on the flange of one of the players (hence "flanger"), that machine would slow down and its signal would fall out-of-phase with its partner, producing a phasing effect.

Overdrive

- effects such as the use of a fuzz box can be used to produce distorted sounds, such as for imitating robotic voices or to simulate distorted radiotelephone traffic (e.g., the radio chatter between starfighter pilots in the science fiction film *Star Wars*).
- The most basic overdrive effect involves *clipping* the signal when its absolute value exceeds a certain threshold.

Pitch shift

- similar to pitch correction, this effect shifts a signal up or down in pitch. For example, a signal may be shifted an octave up or down. This is usually applied to the entire signal, and not to each note separately.
- One application of pitch shifting is pitch correction. Here a musical signal is tuned to the correct pitch using digital signal processing techniques.
- This effect is ubiquitous in karaoke machines and is often used to assist pop singers who sing out of tune. It is also used intentionally for aesthetic effect in such pop songs as Cher's *Believe* and Madonna's *Die Another Day*.

Time stretching

- the opposite of pitch shift, that is, the process of changing the speed of an audio signal without affecting its pitch.

Resonators

- emphasize harmonic frequency content on specified frequencies.

Robotic voice effects

- are used to make an actor's voice sound like a synthesized human voice.

Synthesizer

- generate artificially almost any sound by either imitating natural sounds or creating completely new sounds.

Modulation

- to change the frequency or amplitude of a carrier signal in relation to a predefined signal. Ring modulation, also known as amplitude modulation, is an effect made famous by Doctor Who's Daleks and commonly used throughout sci-fi.

3D audio effects

- Place sounds outside the stereo basis

Audio editing software's

- **Pro Tools** is a digital audio workstation for Microsoft Windows and OS X, developed and manufactured by Avid Technology.
- **Audacity** is a free open source digital audio editor and recording computer software application, available for Windows, Mac OS X, Linux and other operating systems

Sony Sound Forge (formerly known as *Sonic Foundry Sound Forge*) is a digital audio editing suite by Sony Creative Software which is aimed at the professional and semi-professional markets.

- **Adobe Audition** (formerly **Cool Edit Pro**) is a digital audio workstation from Adobe Systems.