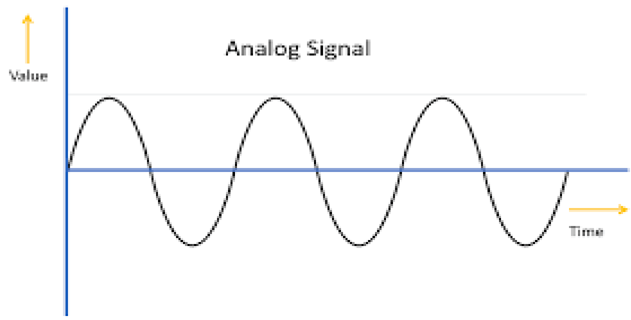
AUDIO CODECS

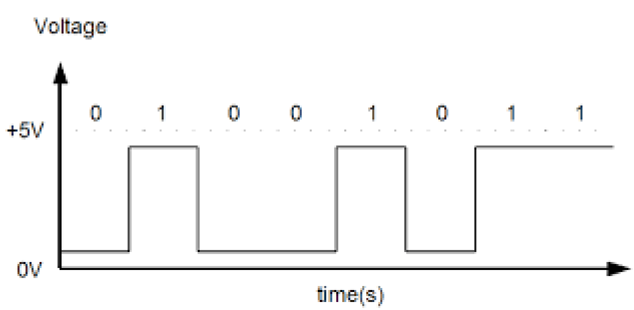
What is an analog signal?

An analog signal is a continuous and variable electrical signal that represents information as a smoothly varying voltage or current. In the context of electronics and telecommunications, analog signals are used to convey data such as audio, video, temperature, or any other continuously variable information.



What is digital signal?

A digital signal is a discrete, quantized representation of information using binary code, unlike analog signals, which are continuous and vary smoothly, digital signals are composed of distinct, discrete values.



What is a Codec?

A codec, short for coder-decoder, is a technology used to compress and decompress digital data for transmission, storage, or streaming. Codecs are commonly used in various digital media applications such as audio, video, and telecommunications.

Here's how codecs work:

- **Encoding (Compression):** In the encoding stage, the codec compresses the digital data by removing redundant or unnecessary information. This compression reduces the size of the data, making it more efficient for storage or transmission. There are two main types of compression used in codecs:
  - **Lossless Compression:** This method compresses the data without losing any information, resulting in a smaller file size but retaining all the original data.
  - **Lossy Compression:** This method removes some data that is considered less important or imperceptible to human senses, resulting in a smaller file size. However, this process may result in a slight loss of quality.
- **Decoding (Decompression)**: In the decoding stage, the codec decompresses the compressed data to its original form. This allows the digital data to be played back, displayed, or used in its intended application.

Codecs are used in a wide range of applications, including:

- **Audio Codecs**: Used to compress and decompress audio data for formats such as MP3, AAC, and FLAC.
- **Video Codecs**: Used to compress and decompress video data for formats such as H.264, H.265 (HEVC), and VP9.
- **Telecommunications**: Used to compress voice and video data for transmission over networks, such as VoIP (Voice over Internet Protocol) and video conferencing.
- **Streaming Media**: Used to compress audio and video data for streaming services, such as Netflix, YouTube, and Spotify.

Technically, analog audio itself cannot be compressed directly because compression, as it is commonly understood, involves reducing the size of digital data by encoding it more efficiently. Analog signals are continuous and do not have a digital representation with discrete values that can be compressed.

However, in practice, when people refer to compressing analog audio, they are often talking about a two-step process:

- **Analog-to-Digital Conversion (ADC):**
    - Before compression can take place, analog audio signals are first converted into digital form using an Analog-to-Digital Converter (ADC). This process involves sampling the analog signal at regular intervals and assigning discrete digital values to represent the amplitude of the signal at each sample point.
- **Digital Audio Compression:**
    - Once the analog audio has been converted to a digital format, it can be compressed using various digital audio compression algorithms. This compression reduces the file size of the digital audio data without losing too much perceptual audio quality.
    - Popular audio compression formats for this purpose include MP3, AAC, FLAC, and others.

So, while the original analog audio itself is not compressed, the digital representation of that audio, obtained through analog-to-digital conversion, can be compressed using digital compression techniques. This is a common practice in audio recording, storage, and transmission, especially in the context of digital audio technologies.

Why do we need ADC?

- Because analog signals are more susceptible to noise. Means noise free.
- Difficult to process in analog domain.
- Difficult to store in analog domain.

How ADC works?

The working of an ADC can be explained in 3 steps:

1.Sampling

2.Quantization

3.Encoding

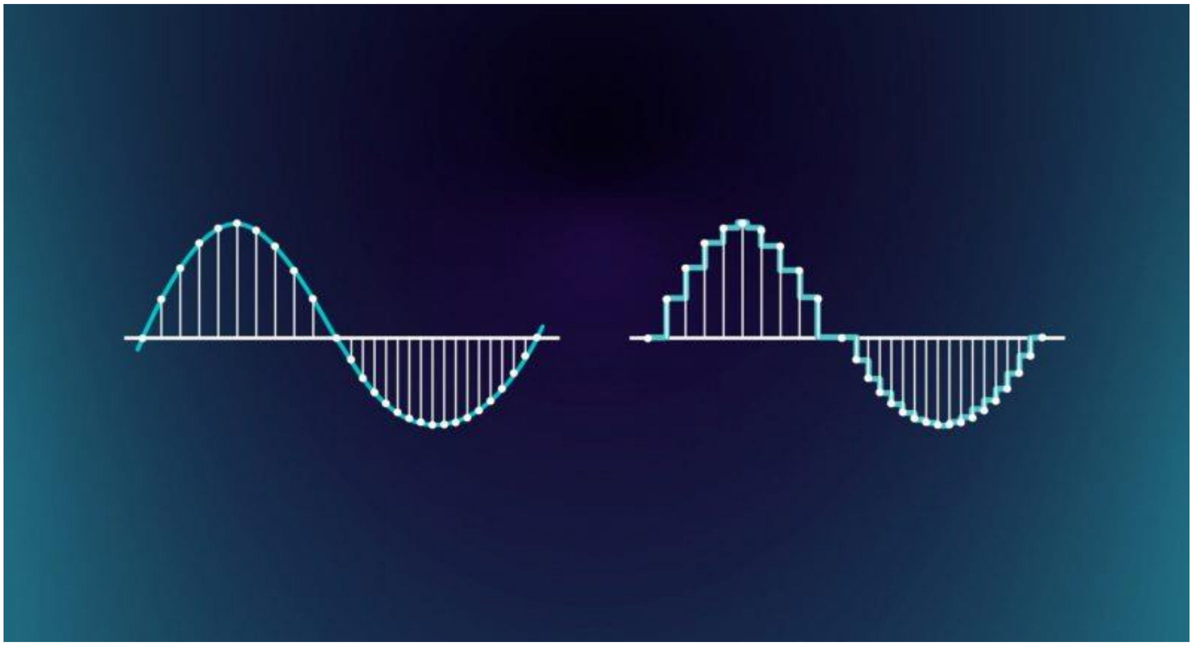
1. Sampling: As we know that analog signal is continuous in time and continuous in amplitude. So, in the sampling stage continuous time signals are converted into discrete time signals. The sampling is done using Nyquist criteria then only we can reconstruct the original analog signal.

what is Nyquist criteria?

the sampling frequency should be >= twice the frequency of the highest frequency component in the signal(bandwidth).

The Nyquist-Shannon sampling theorem, often referred to as the Nyquist theorem or Nyquist-Shannon theorem, is a fundamental principle in signal processing and communication theory. It provides guidelines for accurately sampling analog signals to avoid loss of information during the conversion to digital form.

In mathematical terms, if fs is the sampling frequency, and *fmax is* the maximum frequency component present in the analog signal, the condition for accurate reconstruction is given by $fs \geq 2 \cdot f\text{max}$



2. Quantization: in this stage the continuous amplitude is also converted into discrete.

Here we have to know ADC with how many no of bits. If ADC is 2-bit, then the quantization levels are raised to 2^n.

If ADC is n, then the quantization levels are 2^n.

Also, we have to know about the resolution. Resolution is defined as the minimum change in the input voltage that an adc can detect.

Resolution = VFS/2^n.

where VFS= full scale voltage

VFS = vmax – vmin

So, in quantization the sampled values are rounded to the closest discrete values.

3. Encoding: After quantization, the final step is to encode the quantized values into a digital format. This is typically done by representing each quantized level with a binary code. In the case of an 8-bit ADC, each level is represented by an 8-bit binary number. The encoding process assigns a binary code to each quantized level, forming a digital representation of the analog signal. The encoded digital values can then be further processed, stored, or transmitted by digital systems.

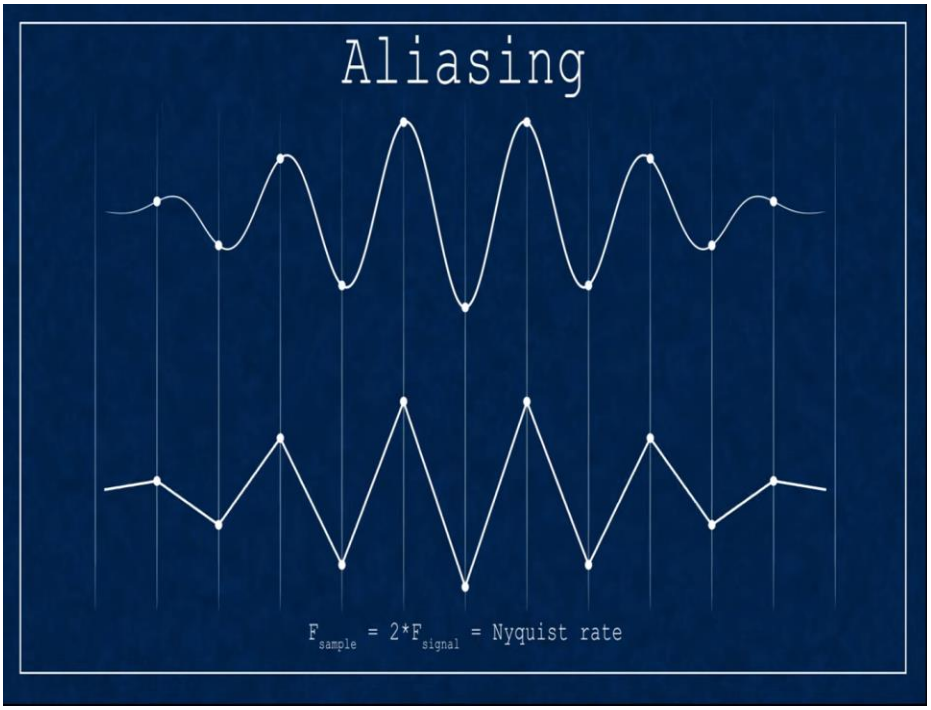
Overall, the working of an ADC involves converting continuous analog signals into discrete digital values through the stages of sampling, quantization, and encoding. It enables the conversion of real-world signals into a digital format that can be processed, analyzed, and utilized by various digital systems and applications.

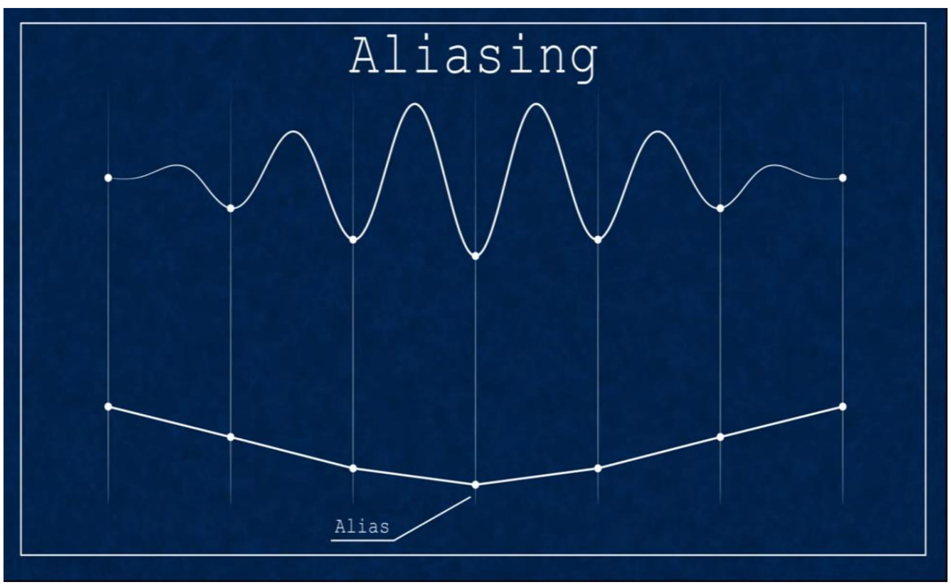
NOTE: Audio sampling rate is a fundamental parameter that directly impacts the Fidelity and accuracy of digital audio recordings. Sampling rate is the major thing in analog to digital conversion.

What is Aliasing?

Aliasing occurs when we sample the wave at a frequency less than the Nyquist sampling rate than the original signal cannot be reconstructed, and the noise or distortion will occur in the signal. Aliasing can alter or change the high frequency signal to a low frequency signal.

- **Nyquist-Shannon Sampling Theorem**: According to this theorem, to accurately reconstruct a continuous signal from its samples, the sampling rate must be at least twice the highest frequency present in the signal. This is known as the Nyquist frequency.
- **Frequency Folding**: When the sampling rate is too low, frequencies higher than the Nyquist frequency fold back into the audible range as lower frequencies. This folding is what causes aliasing.





What is Noise ?

un*wanted and random sounds or disturbances that are not part of the intended audio signal.*

How to eliminate Noise?

By using Active Noise Cancellation Technoloy...

Also, by using passive noise technology

Active Noise Cancellation (ANC) and Passive Noise Cancellation are two different approaches to reducing or eliminating unwanted ambient sounds in audio devices. Let's explore the key differences between active and passive ANC:

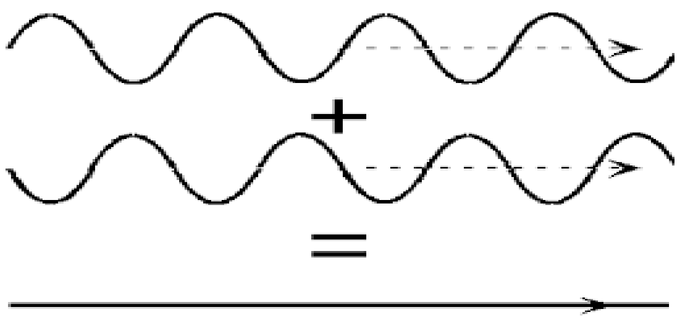
- **Active Noise Cancellation (ANC):**
  Technology: Active Noise Cancellation uses electronic components, including microphones and processing units, to actively generate anti-noise signals that cancel out external noises.
  Mechanism: ANC works by capturing external sounds with microphones, generating inverted sound waves (anti-noise), and mixing them with the original audio content to cancel out unwanted noise through destructive interference.
  Effectiveness: Effective against constant, low-frequency background noises, such as engine hums or air conditioning sounds.
  Power Requirement: Requires a power source, usually provided by a built-in battery, as the electronic components need energy to operate.
  Examples: Many high-end headphones and earphones feature ANC technology to provide a quieter listening experience in noisy environments.

- **Passive Noise Cancellation:**
  Technology: Passive Noise Cancellation relies on the physical design and materials of the audio device to block or reduce external sounds.
  Mechanism: Achieved through the use of physical barriers, sound-absorbing materials, and the design of the earcups or ear tips to physically block or attenuate external noises.
  Effectiveness: More effective at blocking higher-frequency sounds and is generally effective against a broad range of frequencies. However, it may not be as effective against constant, low-frequency noises as active noise cancellation.
  Power Requirement: Does not require a power source, making it more energy-efficient than active noise cancellation.
  Examples: In-ear monitors with a snug fit, over-ear headphones with closed-back designs, and well-designed ear tips that create a seal in the ear canal provide passive noise cancellation.

Active Noise Technology:

Noise is removed using destructive interference.

Destructive interference:



Destructive interference is a phenomenon that occurs when two or more waves meet at a point in space and combine in such a way that their amplitudes subtract from each other, leading to a reduction or cancellation of the overall amplitude at that particular point. It occurs when the peaks of one wave coincide with the troughs of another wave.

Active Noise Cancellation Technology:

Active Noise Cancellation (ANC) is a technology used to reduce or cancel out unwanted ambient sounds by generating sound waves that are specifically designed to counteract the incoming noise. It's commonly used in headphones, earbuds, and other audio devices to provide a quieter and more immersive listening experience, especially in noisy environments.

Here's how active noise cancellation works:

1. Microphones: ANC-equipped devices are equipped with one or more microphones that pick up the ambient sound from the environment, including background noises like engine sounds, chatter, or air conditioning.

2. Noise Analysis: The captured ambient sound is analyzed by an ANC processor to determine the frequency and amplitude characteristics of the unwanted noise.

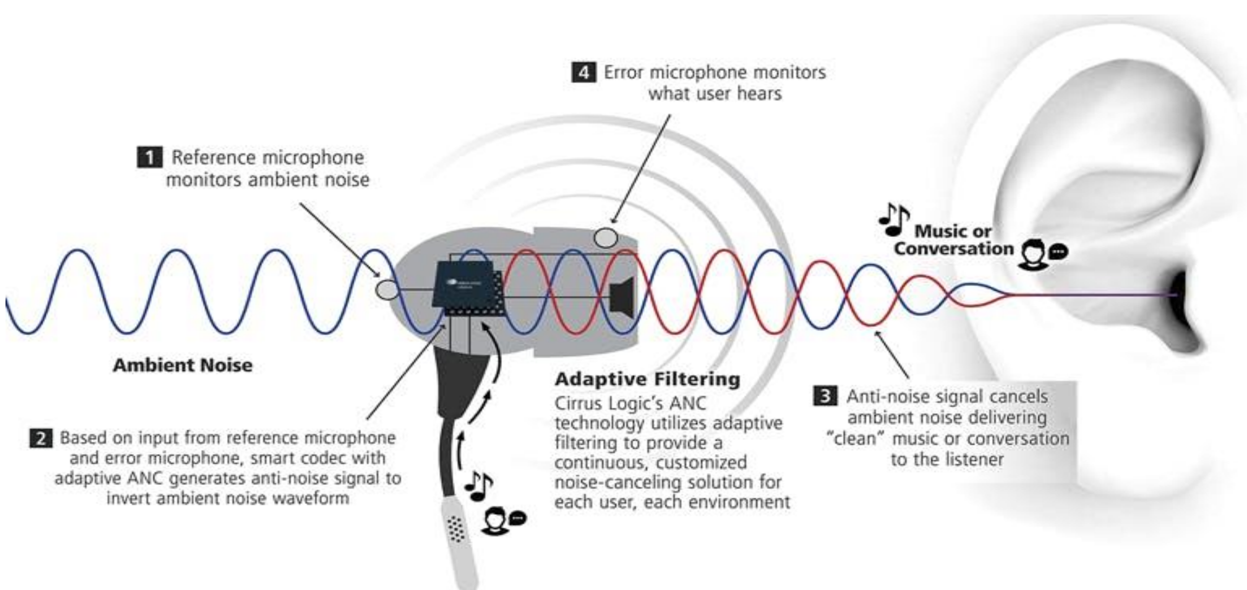
3. **Anti-Noise Generation**: The ANC processor then generates sound waves known as "anti-noise" or "counter-noise." These sound waves are created with the exact opposite phase of the incoming noise. When the anti-noise waves are combined with the original noise, they interfere destructively, effectively canceling out the unwanted noise.

4. Combining Signals: The generated anti-noise sound waves are combined with the audio signal that the listener wants to hear (such as music or speech). This combined signal is then played through the headphones or earbuds.

5. Adaptive Filtering: ANC systems often use adaptive filtering techniques to continuously adjust the anti-noise signals based on changes in the surrounding noise environment. This ensures effective cancellation even when the noise characteristics change.

6. Quiet Environment: As a result, the listener experiences a quieter environment since the unwanted noise is reduced or canceled out by the anti-noise signals.

Active Noise Cancellation (ANC) can be implemented using different methods, and feed-forward cancellation, feedback cancellation, and hybrid cancellation are three common approaches. Each method has its own advantages and limitations, and the choice of implementation depends on factors such as the design goals, the type of ambient noise, and the desired performance characteristics.



- **Feed-Forward Cancellation:**

  Principle: In feed-forward ANC, microphones are placed on the outside of the earcup or earpiece to capture the incoming ambient noise before it reaches the ear.

  Operation: The captured noise signal is then processed, and an anti-noise signal is generated. This anti-noise signal is then combined with the original audio signal before it reaches the ear, creating destructive interference to cancel out the unwanted noise.

  Advantages:

  - Effective for canceling constant, low-frequency noises.
  - Quick response to changes in ambient noise.

  Limitations:

  - May not be as effective against certain types of transient or high-frequency noises.
  - Requires accurate and fast processing to maintain effectiveness.

- **Feedback Cancellation:**

  Principle: In feedback ANC, microphones are placed inside the earcup or earpiece to capture the sound reaching the listener's ear.

  Operation: The captured sound signal is processed, and an anti-noise signal is generated. This anti-noise signal is then combined with the original audio signal before it reaches the ear, achieving destructive interference to cancel out the noise reaching the ear.

  Advantages:

  - Effective against a wide range of frequencies, including transient and high-frequency noises.
  - Can be less affected by changes in the environment compared to feed-forward.

  Limitations:

  - May introduce a time delay in the cancellation process.
  - Can be less responsive to sudden changes in ambient noise.

- **Hybrid Cancellation:**

  Principle: Hybrid ANC combines elements of both feed-forward and feedback cancellation for a more comprehensive noise-canceling solution.

  Operation: By utilizing both external and internal microphones, hybrid ANC aims to address the limitations of each method. The feed-forward component targets external noise, while the feedback component addresses noise reaching the listener's ear.

  Advantages:

  - Enhanced overall performance by leveraging the strengths of both feed-forward and feedback approaches.
  - Improved adaptability to various types of ambient noise.

  Limitations:

- Complexity in design and implementation.

Increased power consumption compared to individual methods.

The choice between feed-forward, feedback, or hybrid ANC depends on factors such as the desired performance, the characteristics of the ambient noise in the environment, and the design goals of the audio device. High-end headphones and earphones often employ hybrid ANC to achieve effective noise cancellation across a broad frequency range and in diverse environments.

What are AUDIO FILE FORMATS?

Audio file formats are utilized to store and organize audio data in digital form.

1.WAV(waveform):

Supports high quality audio.

Uncompressed format, hence larger file sizes make it less suitable for limited storage spaces.

Commonly used for professional audio applications.

2.MP3:

The most widely used audio file format, known for its efficient compression algorithm.

Designed to reduce file size
Lossy compression technology

Used in audio players, devices and os.

Audio compression

Audio compression is a technique used to reduce the size of audio files without significantly affecting the perceived audio quality. It works by encoding the audio in a way that removes unnecessary or redundant information, reducing the overall file size.

Lossless Compression:

In lossless compression, the original data can be perfectly reconstructed from the compressed version. No information is lost during the compression process.

Examples: ZIP
FLAC

Lossy Compression:
In lossy compression, some data is discarded during the compression process, resulting in a smaller file size but a slight loss of quality.

Examples: MPEG
JPEG

3.AAC (advanced audio coding)

Developed as the successor to MP3, AAC provides higher sound quality at lower bit rates.
Possesses improved compression efficiency, delivering better audio fidelity.
Commonly used in online streaming platforms, digital television, and mobile devices.
Apple uses AAC as the default format for its iTunes Store.,youtube

4.FLAC (Free lossless Audio Codec):

FLAC (Free Lossless Audio Codec) is a lossless compression format designed to reduce the file size of audio data without sacrificing any quality.
Used in various operating systems, media players.

5.OGG (ogg vorbis):
Open-source and free format.
Lossy compression with good audio quality.
Often used for streaming and online distribution like gaming, web applications and multimedia development.

NOTE: Choosing the appropriate format is crucial to achieving the desired balance between audio quality and file size.

The "best" format depends on factors such as:

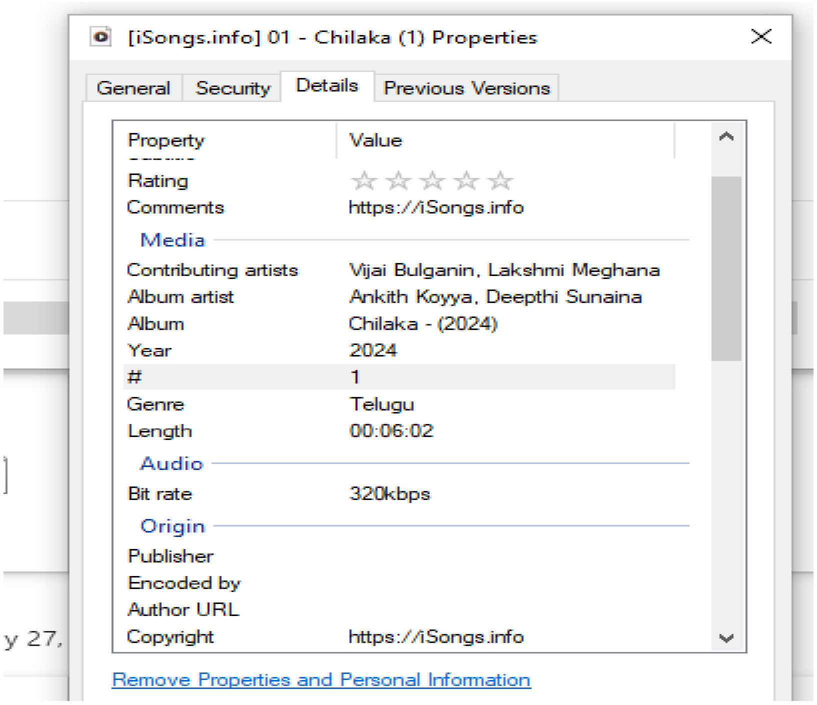
- **Use Case:** Streaming, archiving, editing, etc.
- **Audio Quality:** Whether lossless or lossy compression is acceptable.
- **File Size:** Storage and bandwidth considerations.
- **Device Compatibility:** Some formats may be more widely supported on different devices and platforms.

Type of format	Compression mode	Audio quality	File size
Wav	No compression	High quality	Large size
MP3	lossy	Acceptable quality	Good balance between quality and size
AAC	lossy	high	small
FLAC	lossless	high	Large compared to mp3
OGG	lossy	high	small

What is Bitrate?

Audio bitrate, often simply referred to as "bitrate," is a measure of the amount of data used to represent one second of audio in a digital file. It is typically expressed in bits per second (bps) or kilobits per second (kbps). Bitrate is a crucial parameter in digital audio encoding, representing the level of compression applied to the audio data.
Here for this Chilaka song the bitrate is 320kpbs means for one second it will store 320kbps amount of data.
The higher the bitrate the higher the audio quality and higher the file size.

Bitrate of 320kbps audio



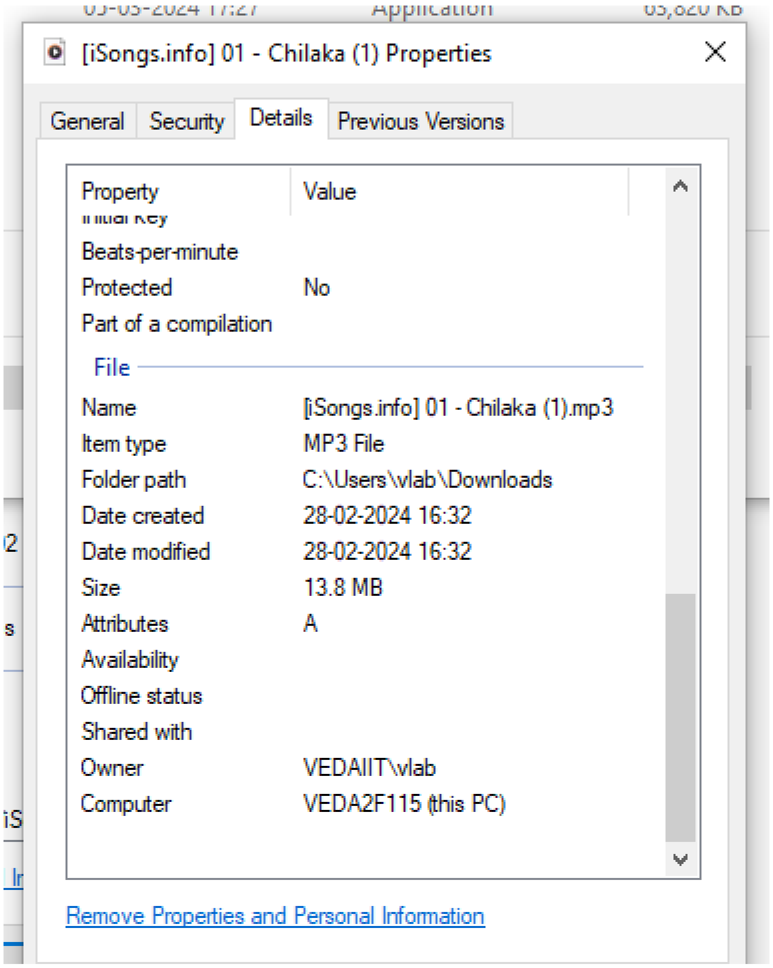
The size of 320kbps bitrate file will be larger

Means in one second it will store 320 kbps of data

Audio quality will be good

File size is larger

File size of 320 kbps audio



Here the size of file is 13.8 mb nearly 14 mb

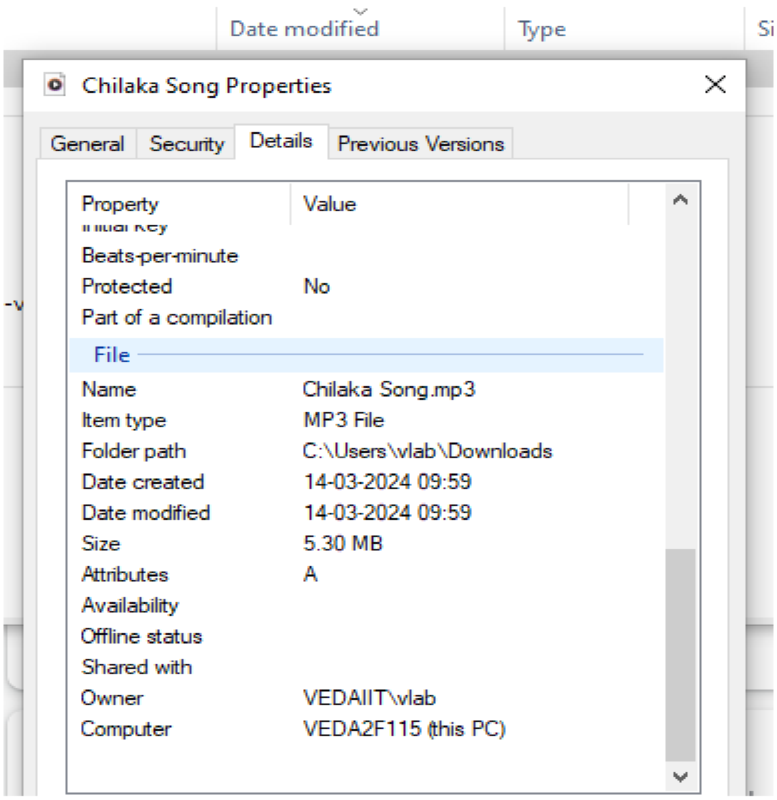
Now let's see 128kbps bit rate audio

Means in one second it will store 128 kbps of data

Compared to 320 kbps the audio quality and file size will be less for 128 kbps

So, bitrate is very crucial for audio quality and the file size

File size of 128 kbps Audio



Here the file size for 128 kbps audio is 5.30 mb which is less than 320 kbps

In the context of digital audio, higher bitrates generally result in higher audio quality but also lead to larger file sizes. Conversely, lower bitrates result in smaller file sizes but may compromise audio quality due to greater compression.

What is Bit depth?

It refers to the number of bits used to represent each sample of audio or pixel of an image. Bit depth plays a crucial role in determining the dynamic range and precision of the digital representation.

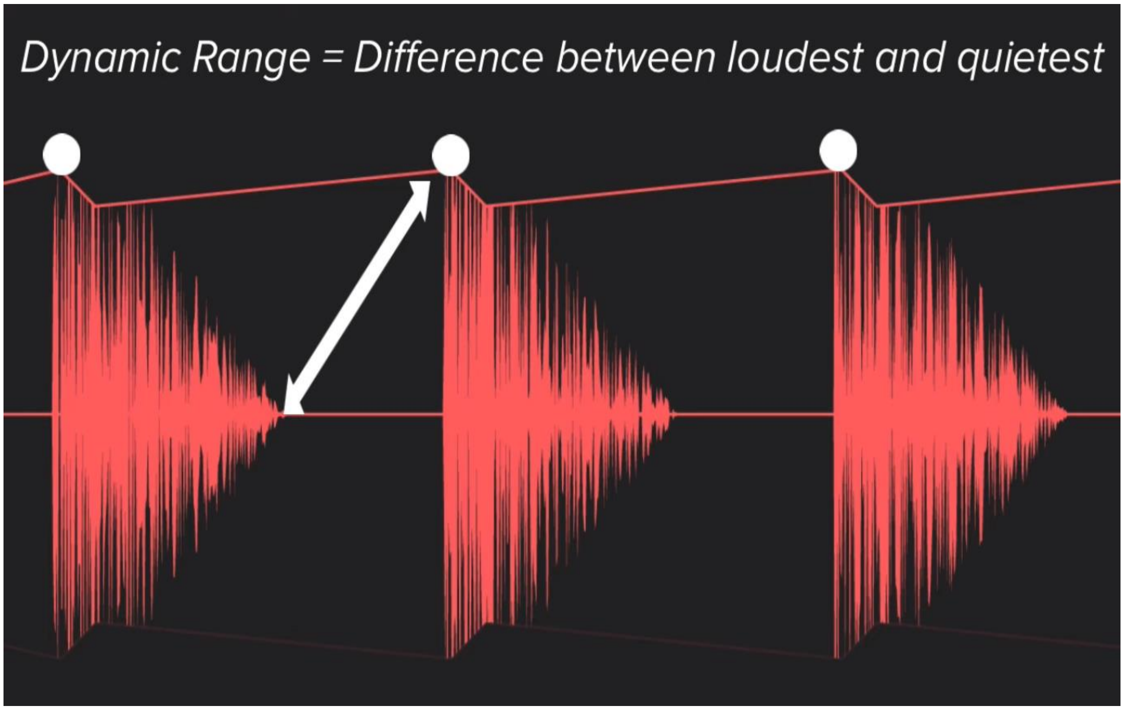
In the context of digital audio:

Bit Depth in Audio:

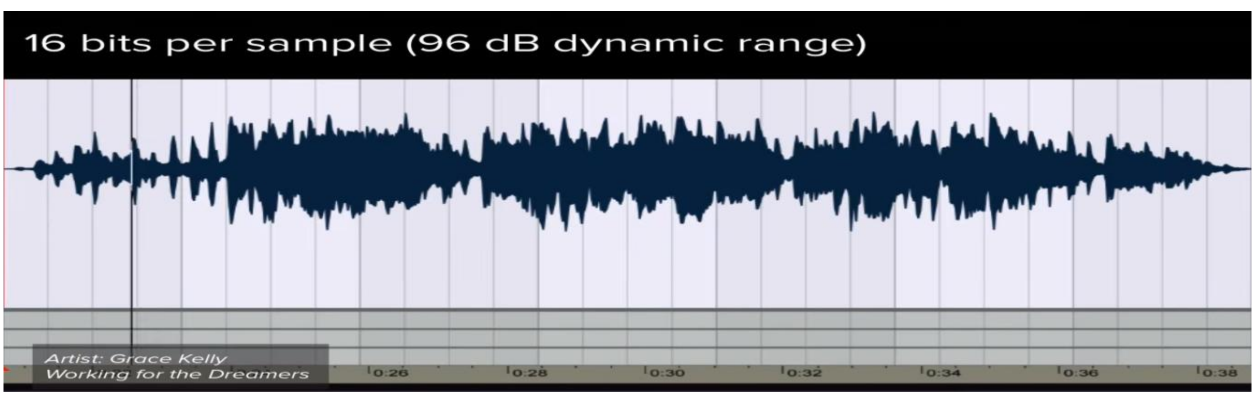
- Bit depth is the number of bits used to represent the amplitude of each sample in a digital audio signal.
- Common bit depths in audio include 8-bit, 16-bit, 24-bit, and 32-bit. The higher the bit depth, the greater the dynamic range and precision.

Dynamic Range:

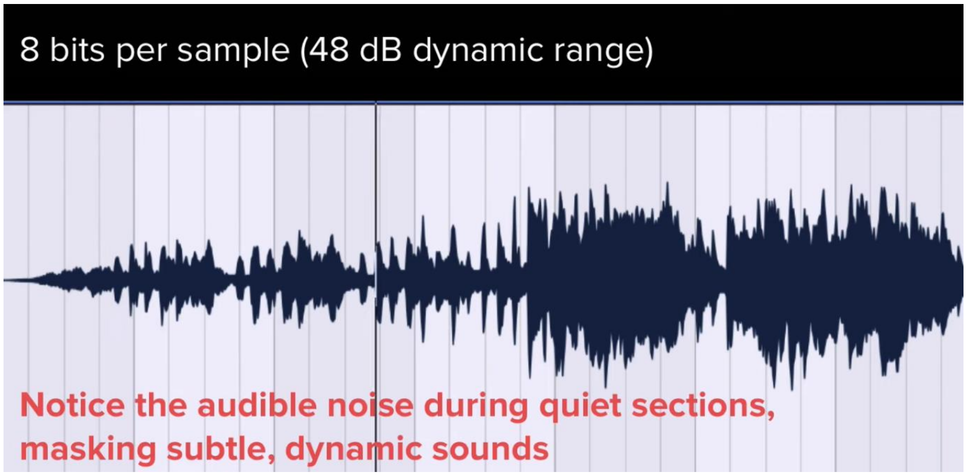
- Dynamic range is the difference between the quietest and loudest parts of an audio signal. Higher bit depths allow for a wider dynamic range, capturing more subtle variations in volume.



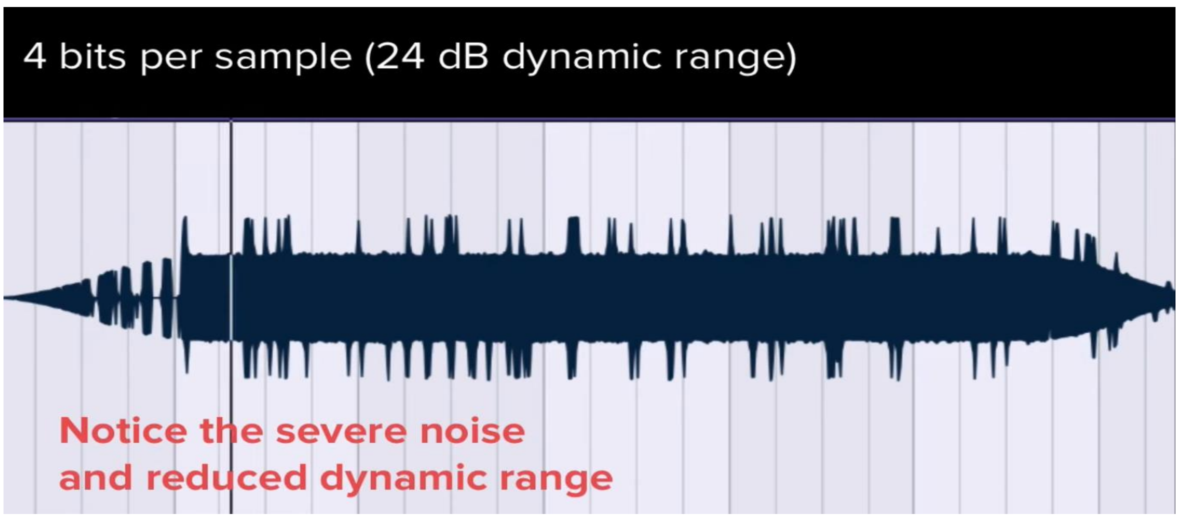
Bit depth of 16 bits:



Bit depth of 8-bits:



Bit depth of 4-bits:

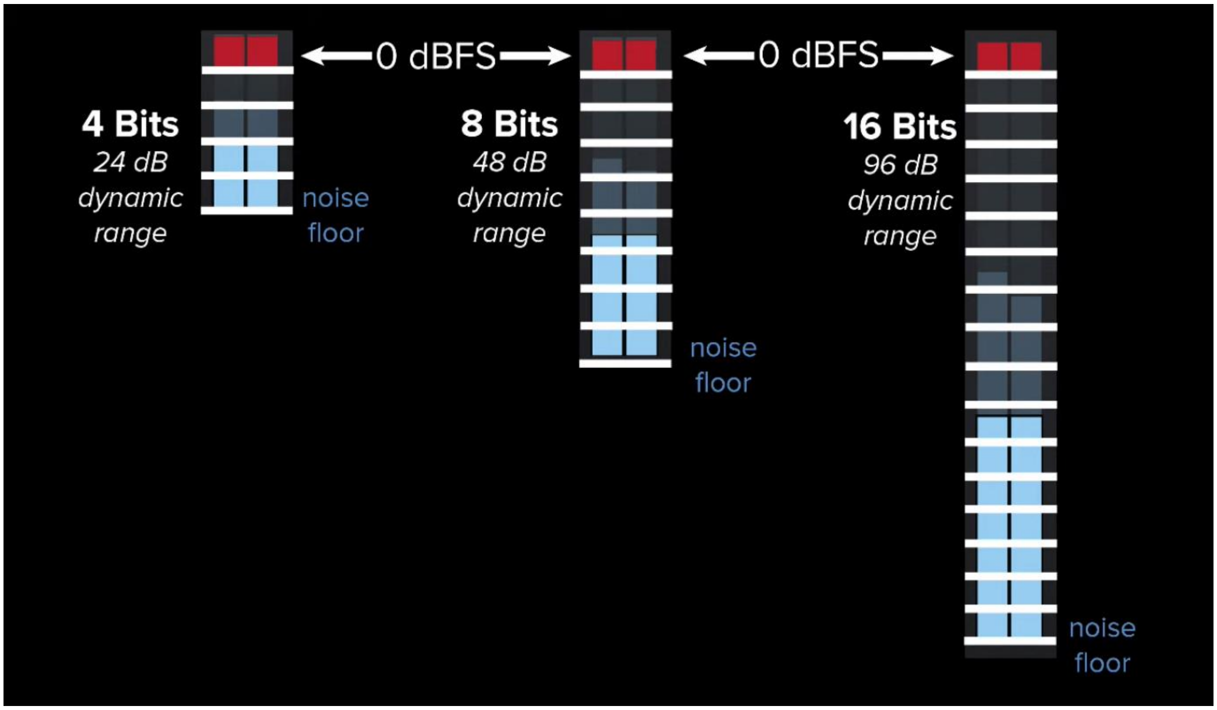


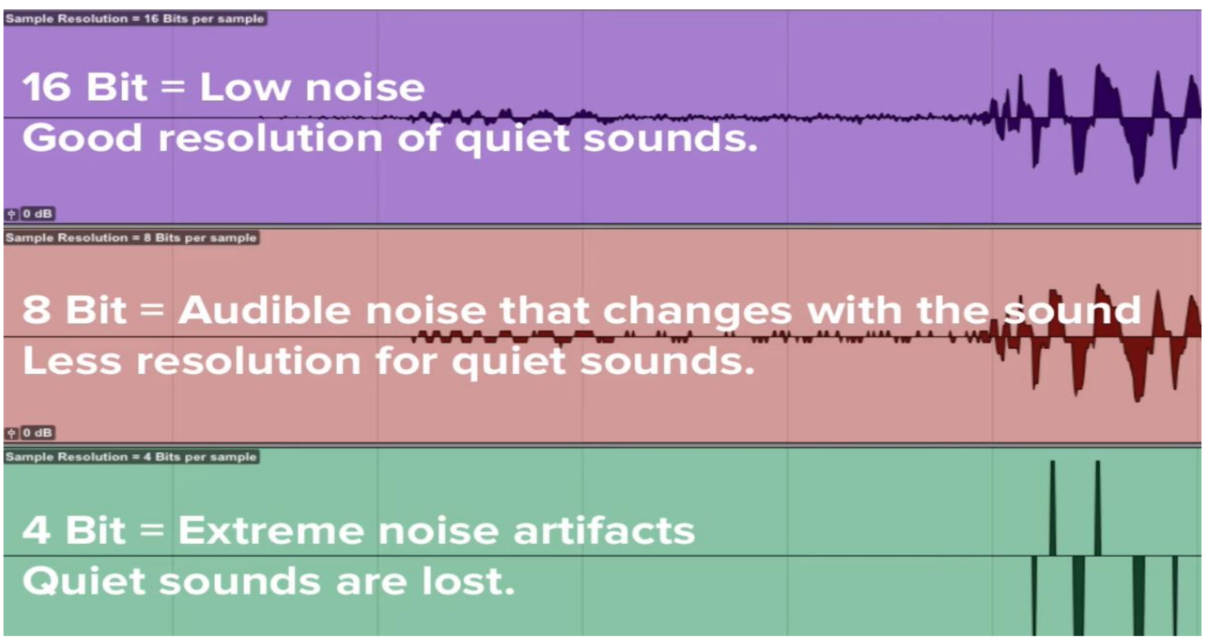
Resolution:

- Bit depth contributes to the resolution of the audio signal. Higher bit depth results in finer resolution and more accurate representation of the original analog waveform

Overview:

So, we can conclude that the higher the bit depth is, the higher the audio quality.


0 dBFS
0 dBFS
4 Bits
24 dB
dynamic
range
noise
floor
8 Bits
48 dB
dynamic
range
noise
floor
16 Bits
96 dB
dynamic
range
noise
floor


Sample Resolution = 16 Bits per sample
16 Bit = Low noise
Good resolution of quiet sounds.
0 dB
Sample Resolution = 8 Bits per sample
8 Bit = Audible noise that changes with the sound
Less resolution for quiet sounds.
0 dB
Sample Resolution = 4 Bits per sample
4 Bit = Extreme noise artifacts
Quiet sounds are lost.

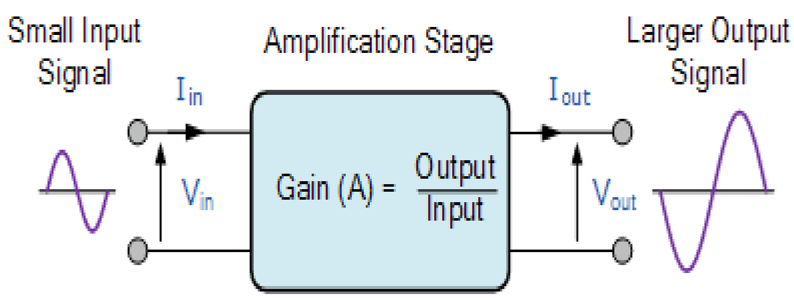
What is DAC?

The Digital-to-Analog Converter (DAC) plays a crucial role in converting digital signals into analog signals for various applications, such as audio playback. The working principle of a DAC involves translating discrete digital values into a continuous analog waveform. Here is a simplified explanation of how a DAC works:

- **Digital Input:**
  - The process begins with a digital input, typically represented by binary code (0s and 1s). This digital input can come from various sources, such as digital audio files, communication systems, or sensors.
- **Sampling:**
  - In the context of audio signals, the digital input often represents a sampled version of an analog waveform. The original analog signal is sampled at regular intervals, and each sample is represented by a binary code. The rate at which these samples are taken is known as the sampling rate.
- **Quantization:**
  - The binary code for each sample is determined by quantization. Quantization involves assigning a digital value to each sample, representing its amplitude. The number of bits used for quantization determines the resolution of the digital signal. For example, a 16-bit DAC can represent each sample with one of 2^16 (65,536) possible values.
- **Digital-to-Analog Conversion:**
  - The quantized digital values are then processed by the DAC to generate an analog output. This involves converting the discrete digital values into a continuous analog waveform that can represent the original audio signal.
- **Reconstruction Filter:**
  - In some DAC designs, a reconstruction filter may be used to smooth the output waveform and eliminate high-frequency components introduced during the sampling process. This helps in reconstructing a more accurate analog signal.
- **Analog Output:**
  - The final output of the DAC is an analog signal that can be sent to amplifiers, speakers, or headphones for playback.

What is an Audio amplifier?

An audio amplifier is an electronic device that increases the amplitude (volume) of an audio signal, making it suitable for driving speakers or headphones. The primary function of an audio amplifier is to take a weak input signal, such as that from a musical instrument, microphone, or audio source, and amplify it to a level sufficient to drive the speakers or headphones and produce sound audible to the listener.



Pre silicon testing?

Pre-silicon testing refers to the testing and verification processes performed before the actual manufacturing or fabrication of semiconductor devices, commonly known as silicon chips or ICs (Integrated Circuits). This phase of testing is crucial in identifying and resolving potential design issues, ensuring the functionality and reliability of the chip design before it goes into production. Pre-silicon testing is an integral part of the overall semiconductor design.

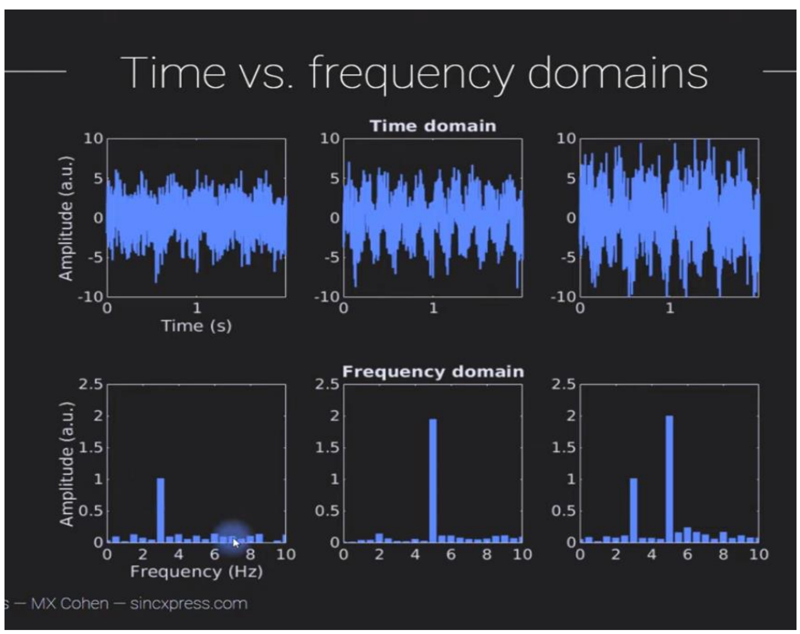
Post silicon testing?

Post-silicon testing refers to the testing and validation processes performed after semiconductor devices, such as integrated circuits (ICs) or silicon chips, have been manufactured. This phase of testing is crucial to ensure that the fabricated silicon devices meet the specifications, performance criteria, and reliability standards set during the design phase. Post-silicon testing aims to detect and address any issues that may arise during the manufacturing process, including defects, variations, and other factors that can affect the functionality and performance of the chips.

How is MP3 Audio file created?

An MP3 (MPEG Audio Layer III) audio file is created through a process called audio compression. This process involves reducing the file size of an audio recording while attempting to maintain a high level of perceived audio quality. Here's an overview of how MP3 audio files are created:

- **Sampling**: The audio signal is sampled at a certain rate (e.g., 44.1 kHz, 48 kHz) to convert the analog audio signal into a digital representation. This process captures snapshots of the audio waveform at regular intervals.
- **Quantization**: The sampled audio data is quantized to represent the amplitude of each sample using a finite number of bits. This process discretizes the amplitude values and reduces the resolution of the audio data, which introduces quantization noise.
- **Transformation:** The audio signal is transformed using a mathematical technique called the Discrete Cosine Transform (DCT). This transformation converts the audio signal from the time domain to the frequency domain, separating it into different frequency components.

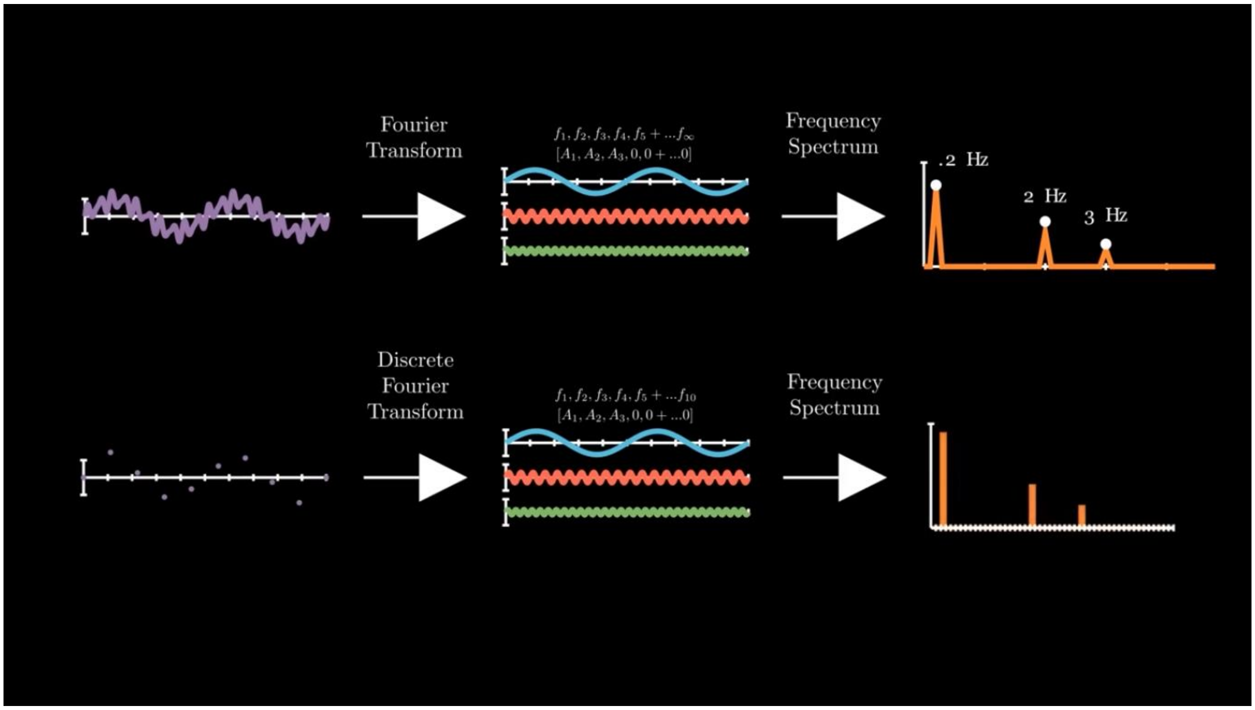


Because analyzing the signal in frequency domain is much easier than time domain

Dct is applied when input is sample points instead of continuous wave.

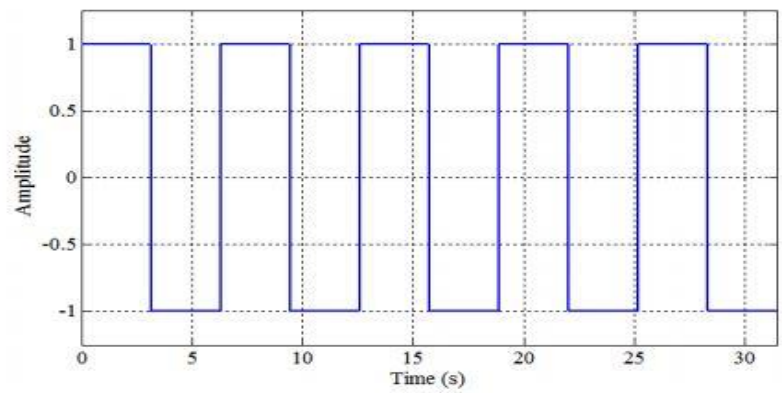
Whereas FFT is applied to continuous waves.

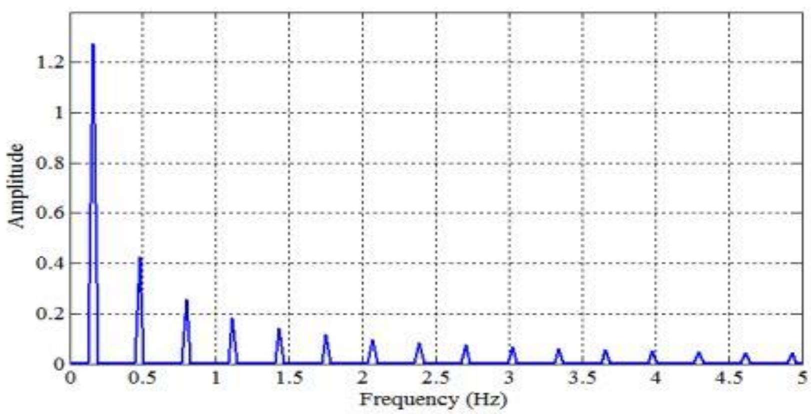
Here the audio file is in digital means represented in samples, so Dct is applied.



After applying the Discrete Cosine Transform (DCT) to a signal, the coefficients obtained represent the signal's frequency content in the transformed domain. The specific coefficients depend on the type of DCT used (e.g., DCT-I, DCT-II, DCT-III, DCT-IV) and the size of the transform (e.g., 8-point DCT, 16-point DCT, etc.).

In the case of the commonly used DCT-II (the standard DCT), the coefficients obtained represent the amplitudes of cosine functions of different frequencies that make up the original signal. The lower-frequency components are represented by the lower-numbered coefficients, while the higher-frequency components are represented by the higher-numbered coefficients.





MP3 encoding utilizes psychoacoustic models to determine **which components of the audio signal are perceptually important** and which can be discarded or encoded at lower quality. The DCT facilitates this process by transforming the audio signal into the frequency domain, where perceptual masking effects are more pronounced and easier to exploit. This allows for more accurate modeling of human auditory perception and efficient allocation of bits to encode the most important audio components.

so, here the frequencies below human hearing will be discarded to reduce the amount of file size. For example, frequencies below 50 hz are discarded because those frequencies are considered infrasonic, and which consists of very low-frequency sounds that are often not perceptible to human ears.

- **Quantization**:

1. After the DCT, the resulting frequency coefficients represent the audio signal's frequency content.
2. Quantization involves reducing the precision of these coefficients to reduce the amount of data required to represent them.
3. The goal is to discard less perceptually relevant information while preserving the overall quality of the audio signal.

- **Psychoacoustic Model Application**:
  1. A psychoacoustic model is applied to the quantized frequency coefficients.
  2. This model determines which coefficients are perceptually important and which can be discarded or encoded at lower quality.

  1) It considers auditory masking effects, tonal vs. non-tonal components, and other perceptual characteristics of human hearing.

- **Bitrate and Compression**: One of the key factors in creating an MP3 file is selecting a bit rate. The bitrate determines how many bits are used to represent a certain amount of audio data. Higher bitrates result in better audio quality but also larger file sizes. The compression algorithm used in MP3 encoding exploits psychoacoustic principles to determine which parts of the audio are less perceptible to the human ear and can be further compressed without a significant loss in perceived quality.

Entropy coding is the final stage in audio compression.

- **Entropy Coding:** The quantized and compressed audio data is then encoded using techniques like Huffman coding or arithmetic coding. These coding methods assign shorter codes to more frequently occurring values, further reducing the overall file size.

  Huffman Coding:

  Huffman coding is a variable-length prefix coding technique that assigns shorter codewords to more frequent symbols and longer codewords to less frequent symbols. It constructs an optimal prefix-free binary tree (Huffman tree) based on the probability or frequency of occurrence of each symbol in the data source. The symbols are then encoded using the codewords obtained from the Huffman tree. Huffman coding is widely used in multimedia compression standards such as JPEG (for still images) and MPEG (for video).

Understanding Huffman coding with examples:

Let's assume that we have to send a message: - BCCABBDDAECCBBAEDDCC
Size is determined by the no.of bits.
Let's see here there are 20 letters.
In computers we have to send ASCII code to send this message.
ASCII codes are 8-bits. For A (65)– 010 000 01, B (66)-010 000 10
So here total 20 letters so, 8 * 20 = 160 bits.

So, let's see if we can use our own codes instead of ASCII codes.
Measure the frequency like how many times the particular letter is present.
So, here only 5 letters 2 power 3 is 8 so we can represent in 3 bits.

character	count/frequency	code	No.of bits to send message
A	3	000	3*3 = 9
B	5	001	3*5 = 15
C	6	010	3*6 = 18
D	4	011	3*4 =12
E	2	100	3*2 = 6
Total = 5*8 = 40 bits	Total-20	Total = 15 bits	Total = 60 bits

Total letters 20 * 3 (bits) = 60 bits.
So, here we have sent the table to decode it also because it doesn't know which code is for what, so we have to sent the table also to help to decode this message.

So, 5 letters ASCII original are in 8-bits here total no.of bits to represent table are 5*8 = 40bits. And we the code we have created are 3 * 5 =15 bits

no.of bits to send message = 60 bits
no.of bits to send table = 40 + 15 = 55 bits
Total no.of bits = 60 +55 = 115 bits

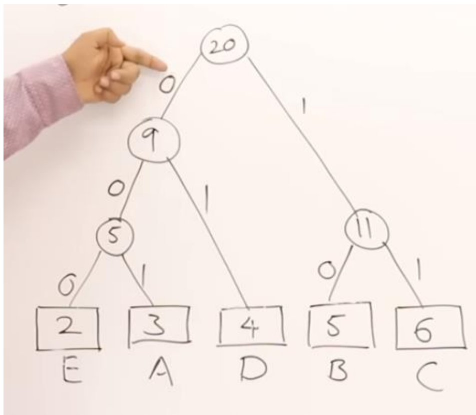
So, by using ASCII a total of 160 bits. But now it is 115 bits . This is fixed sized codes

Here the Huffman code technique comes.
It is variable sized codes.

Huffman code states that the data which is frequently present have to give less no.of bits and the data which is not frequent have to give more bits.

Huffman coding gives values by representing in binary tree.
E = 2
A = 3
D = 4
B = 5
C = 6
All the alphabets should be arranged in the increasing order of their frequency.

Merge two smallest freq and make one node.
Like that choose the minimum and make a node and obtain a tree like this.



Now assign 0 to left sub tree and 1 to the right.
Obtain the values/codes for each letter.

character	count/frequency	code	No.of bits to send message
A	3	001	3*3 = 9
B	5	10	5*2 = 10
C	6	11	6*2 = 12
D	4	01	2*4 = 8
E	2	000	3*2 = 6
Total = 5*8 = 40		Total = 12 bits	Total = 45 bits

Here, you can observe that the more frequency letters B, C, D got 2-bits less no.of bits and less frequency letters got 3-bits.
no.of bits to send message = 45 bits
no.of bits to send table = 40 + 12 = 52 bits

Total no.of bits = 45 +52 = 97 bits which reduces the file size more accurately.

- **Frame Structure:** MP3 files are divided into frames, each containing a fixed number of audio samples. Each frame includes information about how the audio data is encoded, such as bitrate, sample rate, stereo mode, and more.

- **Header Information:** Each frame begins with a header containing information about the frame's structure and the encoding parameters used. This header helps the decoding software understand how to properly interpret and reconstruct the audio data.

- **Decoding:** When playing back an MP3 file, a decoding process is performed. The decoder reverses the steps of the encoding process: it extracts the header information, performs the inverse entropy coding, dequantizes the frequency components, and applies the inverse DCT to reconstruct the audio signal in the time domain.

NOISE streams

Noise streams refer to sequences or streams of random values, often generated by algorithms or physical processes.

Pink Noise:

Pink noise, also known as "1/f noise", is a type of random signal that has equal power per octave or per one-third octave band.

Octave:

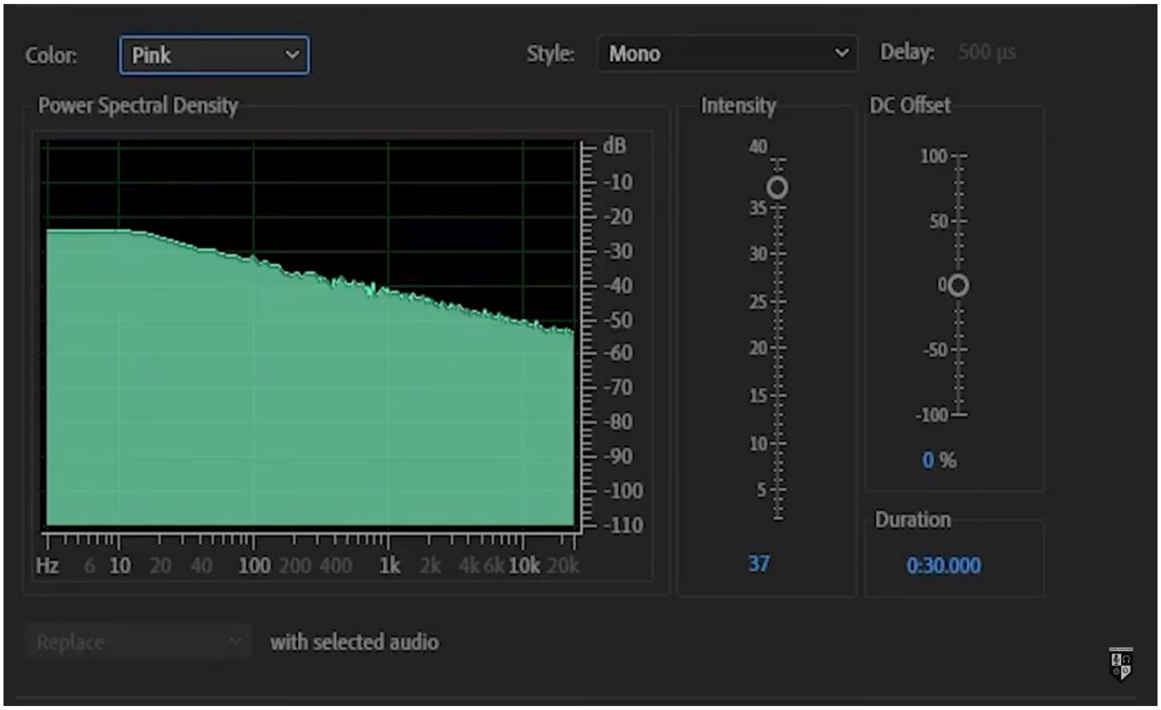
In audio, an octave refers to a doubling or halving of a frequency. It's a logarithmic scale commonly used to describe the relationship between different frequencies.

Here's how it works:

- If you start with a frequency of, say, 100 Hz, the next octave up would be 200 Hz (100 Hz doubled), then 400 Hz (200 Hz doubled), and so on.
- Similarly, the octave below 100 Hz would be 50 Hz (100 Hz halved), then 25 Hz (50 Hz halved), and so forth.

Each octave represents a doubling or halving of the frequency of the previous octave. This logarithmic scale is useful because it corresponds more closely to how humans perceive changes in pitch. For example, an increase of one octave is perceived as doubling the pitch, regardless of the starting frequency.

In other words, pink noise has more power in the lower frequencies compared to higher frequencies, resulting in a spectrum that decreases in power as the frequency increases.



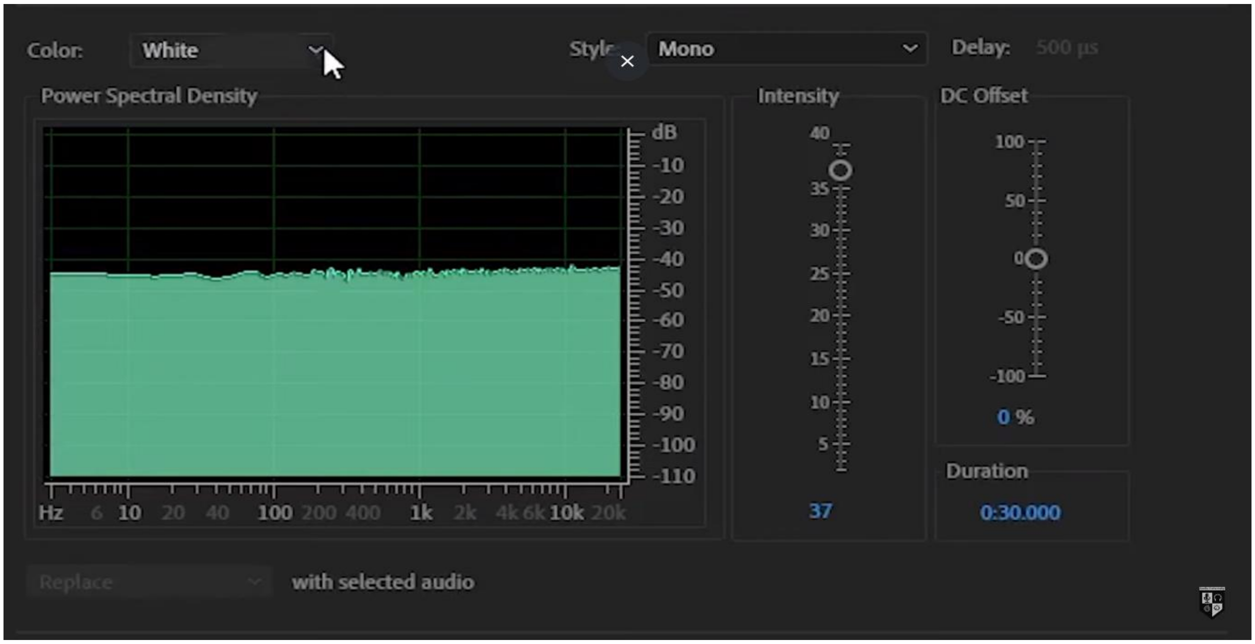
Observe the graph above the energy is high at low – frequencies and gradually energy decreasing at high frequencies.

Pink noise, with its balanced spectral characteristics, finds various uses across different fields. Here are some common applications of pink noise:

- **Audio Testing and Calibration**: Pink noise is often used in audio engineering and testing to assess the frequency response of audio equipment such as speakers, headphones, and microphones. By playing pink noise through a system and analyzing the output, engineers can identify frequency response irregularities and calibrate the equipment accordingly.
- **Room Acoustics Measurement**: Pink noise is utilized in room acoustics measurement to evaluate the reverberation time, frequency response, and other acoustic parameters of a space. By emitting pink noise into a room and analyzing the resulting sound field, acousticians can assess the room's acoustic properties and make adjustments for optimal sound quality.
- **Psychoacoustic Studies**: Pink noise is used in psychoacoustic experiments to investigate how humans perceive sound. Researchers utilize pink noise as a reference signal to study phenomena such as auditory masking, frequency discrimination, and loudness perception.

White noise:

White noise is a type of sound that contains equal amounts of frequencies within the audible range, typically ranging from 20 Hz to 20 kHz. It is called "white" because it's analogous to white light, which contains all colors of the spectrum. In white noise, each frequency band has equal power, resulting in a flat frequency response when plotted on a graph.



Observe the graph intensity is equal across all frequencies showing a flat response.

White noise is often used in various applications, including sound masking, relaxation, and improving concentration. It can also be utilized in audio engineering and scientific research to test equipment frequency response, simulate environmental conditions, or as a reference signal. Additionally, white noise is sometimes used in electronic communication systems to mask unwanted signals or to provide a constant background noise level for calibration purposes.

Babble Noise:

"Babble noise" typically refers to a type of noise characterized by the presence of multiple voices or sounds overlapping each other, often resulting in an indistinct or chaotic auditory environment. This term is commonly used in telecommunications and audio processing contexts.

In telecommunications, babble noise may refer to interference on a communication channel caused by multiple simultaneous transmissions or background noise, making it difficult for the receiver to discern individual signals clearly.

In audio processing, babble noise can refer to background chatter or crowd noise in recordings or live audio environments. It can also be used to describe the sound of multiple people speaking at once, creating a cacophony of voices.

Subjective and objective testing are two methods used in audio engineering and research to assess the quality, performance, and characteristics of audio systems, codecs, and signals. Here's a brief explanation of each:

- **Subjective Testing:**
  - **Definition:** Subjective testing involves human listeners or observers who provide qualitative feedback on the perceived quality, clarity, naturalness, and other subjective attributes of audio signals.
  - **Method:** In subjective testing, listeners are typically presented with audio samples and asked to rate or rank them based on their preferences, perceived quality, or specific criteria. This may involve listening to audio samples in controlled environments (e.g., listening rooms) or through headphones.
  - **Examples:** Double-blind listening tests, MOS (Mean Opinion Score) tests, ABX tests, MUSHRA (Multiple Stimuli with Hidden Reference and Anchor) tests, and preference tests are common subjective testing methods used in audio research and quality assessment.

- **Objective Testing:**
  - **Definition:** Objective testing uses measurement tools, algorithms, and metrics to quantitatively evaluate the performance, characteristics, and quality of audio signals without human intervention.
  - **Method:** Objective testing relies on mathematical models, signal processing techniques, and perceptual metrics to analyze audio signals and derive objective measures of performance or quality. This may involve measuring parameters such as frequency response, distortion, dynamic range, signal-to-noise ratio, and other audio quality metrics.
  - **Examples:** FFT (Fast Fourier Transform) analysis, THD (Total Harmonic Distortion) measurements, SNR (Signal-to-Noise Ratio) measurements, PESQ (Perceptual Evaluation of Speech Quality), and POLQA (Perceptual Objective Listening Quality Analysis) are examples of objective testing methods used in audio engineering and research.

Peak and RMS:

Peak

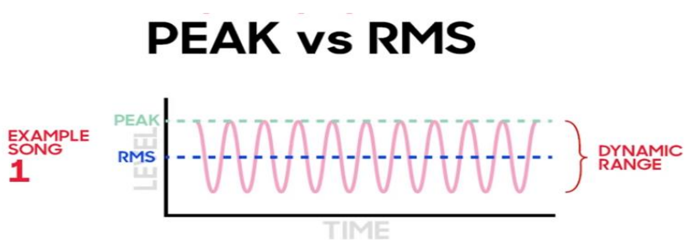
Peak levels represent the maximum amplitude reached by an audio signal at any given moment. Peak levels are measured in decibels (dB) relative to a reference level, and they provide information about the instantaneous loudness or intensity of the signal.

Use: Peak values are important for preventing clipping and distortion in audio processing and recording. They help ensure that the signal does not exceed the maximum allowable level of the recording or playback system.

Rms:

RMS stands for Root Mean Square, which is a method used to measure the average power of a waveform over a given period. RMS is commonly used to quantify the amplitude or loudness of an audio signal accurately.

RMS stands for Root Mean Square, which is a method used to measure the average power of a waveform over a given time period. RMS is commonly used to quantify the amplitude or loudness of an audio signal accurately.



For an audio signal that is represented as a series of discrete samples, the RMS value is calculated as follows:

$$\mathrm{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2}$$

Where:

- RMS is the Root Mean Square value of the signal.
- N is the total number of samples in the signal.
- x_i represents the amplitude of the ith sample.

Use: RMS is commonly used for level matching, dynamic range assessment, and volume normalization in audio processing and mastering. It is useful for ensuring consistent loudness levels between different audio tracks or sections of a recording.

SPL:

Spl means sound pressure level

In the context of audio, SPL is used to describe the loudness of audio signals. SPL is used in the design and testing of audio equipment, such as speakers and headphones, to ensure that they produce sound at safe and appropriate levels. It is also used in recording studios and live sound environments to monitor the volume of audio signals and prevent hearing damage to performers and listeners.

SPL Meter

Sound pressure level (SPL) is a measurement of the intensity or loudness of sound. It is expressed in decibels (dB) and represents the level of air pressure variations caused by a sound wave. SPL is a logarithmic scale that compares the sound pressure of a particular sound to the reference sound pressure of 20 micro pascals. Higher SPL values indicate louder sounds, while lower values indicate softer sounds. The human ear has a range of hearing between approximately 0 dB (threshold of hearing) to 130 dB (pain threshold), with normal conversation typically ranging from 60-70 db.

SNR:

SNR stands for signal to noise ratio.

SNR stands for Signal-to-Noise Ratio. In the context of an audio signal, the Signal-to-Noise Ratio refers to the ratio of the desired audio signal's power or amplitude to the power or amplitude of the unwanted noise present in the signal. It's a measure of how much stronger the audio signal is compared to the background noise.

Mathematically, the Signal-to-Noise Ratio (SNR) is expressed in decibels (dB) and is calculated using the following formula:

$$SNR_{\text{dB}} = 10 \cdot \log_{10}\left(\frac{P_{\text{signal}}}{P_{\text{noise}}}\right)$$

Where:

- SNR_{dB} is the Signal-to-Noise Ratio in decibels.
- P_{signal} is the power of the audio signal.
- P_{noise} is the power of the noise in the signal.

Resonance:

When the produced frequency matches the natural frequency of the object. Then the object vibrates and oscillates more and increases the amplitude.

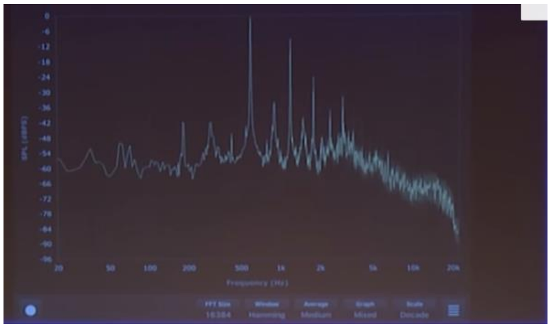
So, at resonant frequency the glass will break.

THD+N

Abbreviates as total harmonic distortion plus Noise

First off, all understand what harmonics is

A harmonic is a sound wave that has a frequency that is an integer multiple of a fundamental tone.



The spikes in the above picture are referred to as harmonics.

Harmonics are multiples of the fundamental frequency present in a complex waveform. When a periodic waveform contains harmonics, its shape deviates from a pure sine wave, and additional frequencies are present in the signal at integer multiples of the fundamental frequency.

Total Harmonic Distortion plus Noise (THD+N) is a metric used to quantify the level of distortion and noise present in an audio signal compared to the original input signal. While harmonics themselves can be useful in various applications, excessive levels of harmonics, along with other types of distortion and noise, can degrade the quality of the audio signal. THD+N provides a comprehensive measure of these distortions and noise combined.

Here's a breakdown of what THD+N represents:

- **Total Harmonic Distortion (THD):** THD measures the percentage of harmonic distortion present in the audio signal relative to the amplitude of the fundamental frequency. Harmonic distortion occurs when the input signal passes through a nonlinear system or device, causing the generation of harmonics that were not present in the original signal.
- **Noise:** In addition to harmonic distortion, audio signals can also be affected by various types of noise, such as thermal noise, electronic interference, and environmental noise. Noise adds unwanted components to the audio signal, reducing its fidelity and clarity.
- **THD+N:** THD+N combines both harmonic distortion and noise into a single metric, providing a comprehensive measure of the overall distortion and noise level present in the audio signal. It quantifies the total amount of signal degradation relative to the original input signal.

THD+N is typically expressed as a percentage or in decibels (dB) and is commonly used in audio engineering and testing to evaluate the performance of audio equipment, such as amplifiers, DACs (Digital-to-Analog Converters), and audio interfaces. Lower THD+N values indicate less distortion and noise, reflecting higher audio quality and fidelity.

Measuring THD+N involves comparing the output signal from the audio device under test to an ideal reference signal and analyzing the resulting distortion and noise components. Test signals, such as sine waves or multi-tone signals, are often used for THD+N measurements, and specialized equipment, such as audio analyzers, are employed for accurate measurement and analysis.

PESQ:

Perceptual Evaluation of Speech Quality (PESQ) is used for objectively assessing the quality of speech signals in telecommunications and audio communication systems. It provides a quantitative measure of how speech quality is perceived after undergoing various transmission or processing stages, such as compression, encoding, or transmission over communication networks. Here's a step-by-step process of how PESQ is used:

- **Selection of Reference and Degraded Signals:**
  - The process begins by selecting a reference speech signal, which is typically a clean, high-quality version of the speech signal.
  - A degraded speech signal is also chosen, which is a processed or altered version of the reference signal. This signal may be subjected to various impairments such as noise, compression artifacts, packet loss, or channel distortion.
- **Preprocessing (Optional):**
  - Preprocessing steps may be applied to both the reference and degraded signals to ensure they are properly aligned and normalized for comparison. This may include resampling, filtering, or level adjustment to match the characteristics of the signals.
- **Calculation of PESQ Score:**
  - The PESQ algorithm compares the reference and degraded signals and calculates a quality score, often represented as a Mean Opinion Score (MOS). The MOS typically ranges from 1 (worst) to 5 (best), with higher scores indicating better speech quality.
  - PESQ analyzes perceptual features of the signals, such as speech intelligibility, clarity, and naturalness, to assess the perceived degradation in speech quality caused by factors like noise, compression, and transmission errors.
- **Standardization:**
  - PESQ is standardized by the International Telecommunication Union (ITU) as ITU-T Recommendation P.862. This ensures consistency and interoperability across different implementations and applications of PESQ.
- **Interpretation of Results:**
  - The PESQ score obtained from the comparison provides an objective measure of the perceived quality of the degraded speech signal. A higher PESQ score indicates better speech quality, while a lower score indicates more significant degradation.

- The results of the PESQ evaluation can be used to assess the performance of communication systems, codecs, algorithms, or transmission channels. They help in optimizing system parameters, diagnosing problems, and ensuring high-quality speech communication experiences for users.

PESQ AND POLQA:

Perceptual Evaluation of Speech Quality (PESQ) and Perceptual Objective Listening Quality Analysis (POLQA) are both standardized algorithms and methods used for objectively assessing the quality of speech signals in telecommunications and audio communication systems. While they share similarities in their goals, there are some key differences between PESQ and POLQA:

- **Methodology:**
  - **PESQ:** PESQ evaluates speech quality by comparing a reference (original) speech signal with a degraded (processed) version of the same signal. It quantifies the perceived degradation in speech quality caused by factors such as noise, compression, and transmission errors. PESQ operates based on a perceptual model that simulates the human auditory system's response to speech signals.
  - **POLQA:** POLQA also evaluates speech quality by comparing a reference speech signal with a degraded version. However, POLQA uses an enhanced perceptual model and advanced signal processing techniques to provide more accurate and reliable assessments, particularly for modern communication systems such as VoIP and wideband audio codecs. POLQA is designed to address limitations of PESQ, particularly in handling wideband audio and non-linear processing.

- **Frequency Range:**
  - **PESQ:** PESQ was originally designed to assess speech quality in narrowband communication systems, typically covering frequencies up to 4 kHz. While it can be used for wideband speech, its performance may be limited in such cases.
  - **POLQA:** POLQA is optimized for wideband and super-wideband speech signals, covering a broader frequency range than PESQ. It is capable of accurately assessing speech quality in modern communication systems that utilize wideband audio codecs and high-definition voice services.

Overall, while both PESQ and POLQA serve similar purposes of objectively evaluating speech quality, POLQA represents a more advanced and refined approach, particularly suited for wideband speech and modern communication systems. However, PESQ remains relevant and widely used, especially in applications where wideband speech is not a primary consideration.

POLQA stands for Perceptual Objective Listening Quality Analysis. It is an objective measurement algorithm used to assess the quality of speech signals, particularly in the context of telecommunications and audio communication systems. POLQA is widely recognized as one of the most accurate and reliable methods for evaluating the perceived quality of speech signals in real-world environments.
POLQA measures the quality of speech signals based on human perception, taking into account factors such as speech intelligibility, naturalness, and overall listening experience. It analyzes the signal using advanced psychoacoustic models to simulate how the human auditory system perceives speech quality.

POLQA is commonly used in telecommunications, VoIP (Voice over Internet Protocol), mobile communication, and audio codec development. It provides valuable insights into the performance of communication systems and helps engineers optimize system parameters to enhance speech quality and user experience.

Overall, POLQA plays a crucial role in ensuring high-quality speech communication in various applications, contributing to improved customer satisfaction and better user experiences in telecommunication services and audio communication systems.

Difference of Noise and distortion:
Distortion in a signal is the alteration or change of the shape or some other characteristic of the waveform. In contrast, noise is an external random signal added to the original signal.

Watchdog Timer?

The watchdog timer acts as a fail-safe mechanism that can save a system from becoming permanently unresponsive.

The watchdog timer is a hardware component designed to monitor the proper functioning of a system and initiate corrective measures if the system fails to respond within a specified timeframe.

Here's what typically happens when a watchdog timer reaches 0:

- **Timeout Occurs:**
    - The watchdog timer is initially set with a specific timeout period or countdown interval. During normal operation, the system's software is responsible for periodically resetting or "feeding" the watchdog timer within this interval.
- **Failure to Feed:**
    - If, for any reason, the system fails to feed the watchdog timer within the specified interval, the countdown timer within the watchdog hardware continues to decrease.
- **Watchdog Timer Reaches 0:**
    - When the countdown timer reaches 0 (zero), it indicates that the system has not been properly responding or feeding the watchdog within the expected timeframe.
- **Watchdog Action:**
    - Upon reaching 0, the watchdog timer initiates a predefined action. The most common action is to trigger a system reset. This reset is intended to recover the system from a potentially malfunctioning or unresponsive state.
- **System Reboot:**
    - The system is forced to reboot as a result of the watchdog timer reaching 0. During the reboot process, the system goes through its startup sequence, initializing hardware, loading the operating system, and resuming normal operation.

The purpose of the watchdog timer is to add a layer of fault tolerance to a system. It automatically detects and responds to situations where the system becomes unresponsive, hangs, or experiences other failures. The watchdog timer provides a fail-safe mechanism to recover the system from such conditions and ensure continued operation.

OTP (one time programmable)?

in embedded systems, OTP (One-Time Programmable) refers to a type of **non-volatile memory** that can be programmed only once, and its content cannot be modified or erased thereafter. OTP memory is commonly used to store data that needs to be preserved across power cycles or system resets, such as configuration settings, security keys, or device-specific information.

Applications in Embedded Systems:

- OTP memory is commonly used in embedded systems for storing device-specific information, serial numbers, calibration data, security keys, and other configuration settings. These are values that need to be set during manufacturing and remain fixed throughout the device's lifetime.
- Security and Tamper Resistance: OTP memory can be used to store cryptographic keys and other security-related information. Since the data is non-modifiable, it can resist tampering attempts aimed at altering or stealing the stored information.

- Configuration and Calibration: OTP memory is commonly used to store configuration settings and calibration data for devices and systems. These settings can be set during manufacturing or initialization and remain fixed throughout the device's lifetime.

Cryptographic keys are essential components in cryptography, a field that deals with secure communication and information protection. Cryptographic keys are used to perform various cryptographic operations, such as encryption, decryption, digital signatures, and authentication. These keys are crucial for ensuring the confidentiality, integrity, and authenticity of data in secure communication systems.

Application Examples: OTP memory is used in various applications such as microcontroller boot code, device identification, secure key storage, system configuration, and more.

FPGA

FPGA stands for Field-Programmable Gate Array. It is a type of integrated circuit (IC) that can be programmed and configured by the user after manufacturing. FPGAs are designed to be highly flexible and can be reconfigured to perform various tasks, making them valuable tools for a wide range of applications, from digital logic design to signal processing and beyond.

Reconfigurability: FPGAs can be programmed and reprogrammed to implement custom digital circuits, allowing designers to create hardware tailored to specific tasks. This makes them versatile and adaptable to changing requirements.

Hardware Description Languages (HDLs): FPGAs are typically programmed using HDLs like Verilog or VHDL. These languages allow designers to describe the desired functionality of the circuit using a hardware-centric approach.

Parallel Processing: FPGAs can be used to create custom parallel processing architectures, making them suitable for tasks that require high computational throughput, such as signal processing, cryptography, and data analysis.

WHAT IS PACKAGE?

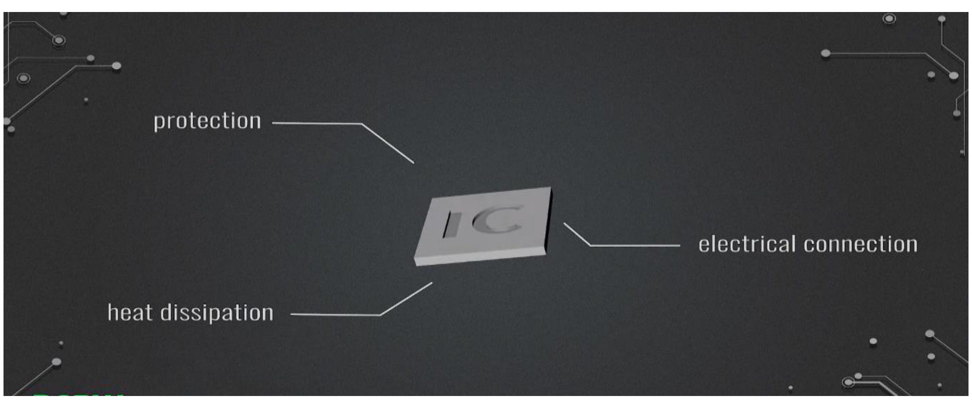
The external cover of an ic is called package.

NEED IC PACKAGES?

why ic's are packaged instead of directly soldered on the surface of the circuit board?

Basic needs of packages:

- Ics requires protection
- Electrical connection
- Heat dissipation



To meet the above needs ic packages serves as a crucial role

These packages not only safeguard the ics from environmental factors but also provide the necessary electrical connections and heat dissipation.

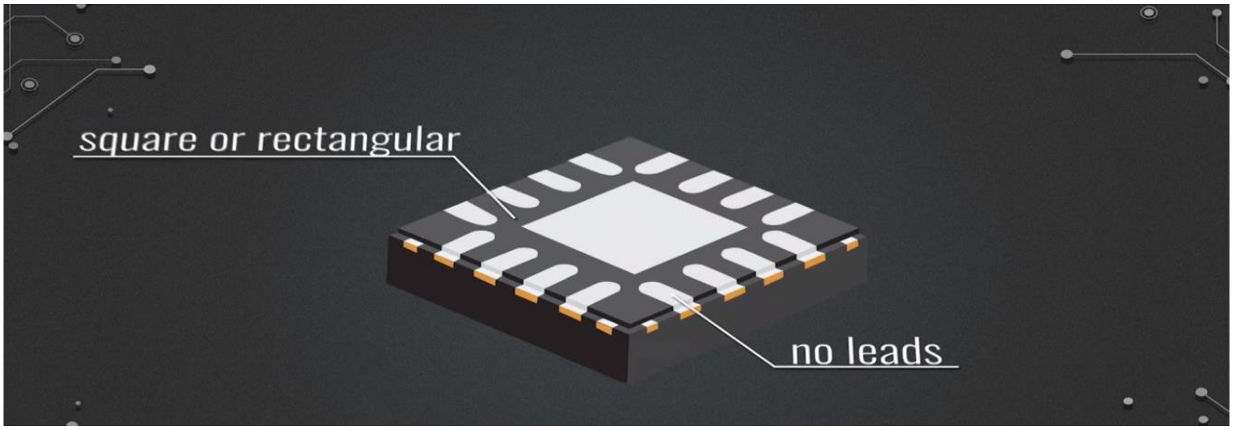
Types of packages:

1.QFN

2.CSP

QFN: QFN stands for quad-flat- no leads

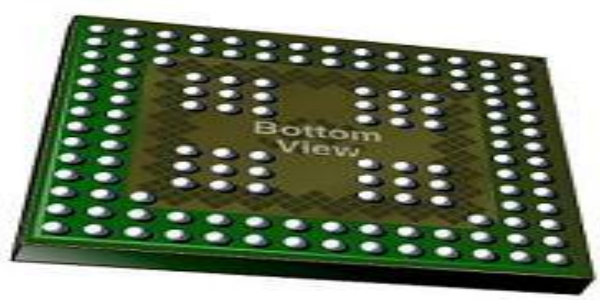
- **Package Design**: QFN packages have a square or rectangular shape and feature a flat leadless bottom surface, with the leads (or terminals) located on the sides of the package. The leads are typically arranged in a grid pattern around the perimeter of the package, which facilitates soldering to the printed circuit board
- **Size and Dimensions**: QFN packages come in a variety of sizes, ranging from very small to larger variants. The dimensions of the package are typically specified in terms of the length, width, and thickness of the body, as well as the pitch (spacing) between the leads.
- **Lead Configuration**: QFN packages can have different lead configurations, including variations with exposed pads on the bottom surface for improved thermal performance, and versions with additional leads for enhanced electrical connectivity or grounding.
- **Advantages**: QFN packages offer several advantages, including a small footprint, which allows for high-density mounting of ICs on PCBs. The absence of leads on the bottom surface also contributes to better thermal performance, as heat can dissipate more efficiently through the exposed pad or bottom surface of the package. Additionally, QFN packages are relatively easy to manufacture and assemble, which helps reduce production costs.



CSP:

Csp stands for chip scale package. Size is small closely to semiconductors.

- **Size and Form Factor**: CSP packages are designed to be as small as possible while still providing the necessary functionality for the integrated circuit. They often have dimensions similar to that of the semiconductor die itself, which allows for more compact designs and efficient use of space on printed circuit boards (PCBs).
- **Application**: CSP packages are commonly used in applications where space is limited, such as in mobile devices (smartphones, tablets), wearable electronics, IoT (Internet of Things) devices, and other compact electronic devices. They are also used in high-density applications where a large number of ICs need to be packed into a small area.
- **Advantages**: CSP packages offer several advantages, including reduced PCB footprint, improved electrical performance due to shorter interconnection lengths, and potentially lower manufacturing costs compared to larger package types. They also tend to have better thermal performance since the smaller size allows heat to dissipate more efficiently.
- **Variants**: There are several variants of CSP packages, including flip-chip CSP (where the die is flipped upside down and mounted directly onto the substrate), wafer-level CSP (where multiple dies are packaged simultaneously at the wafer level), and molded CSP (where the die is encapsulated in a plastic mold).



both QFN and CSP packages are popular choices for IC packaging, and their popularity depends on the specific requirements of the application. QFN packages are more commonly used in a wide range of industries and applications due to their versatility and cost-effectiveness, while CSP packages are favored in applications where miniaturization and high-density mounting are paramount. Ultimately, the choice between QFN and CSP packages depends on factors such as space constraints, thermal considerations, manufacturing requirements, and overall design goals.