

ANALOG TO DIGITAL CONVERSION OF SPEECH SIGNAL

Anything which carries some information is defined as a signal. Speech is a sound signal which are created by the vocal cords, travels through the vocal tract, produced at speaker's mouth and perceived by the listener ear as a pressure wave.

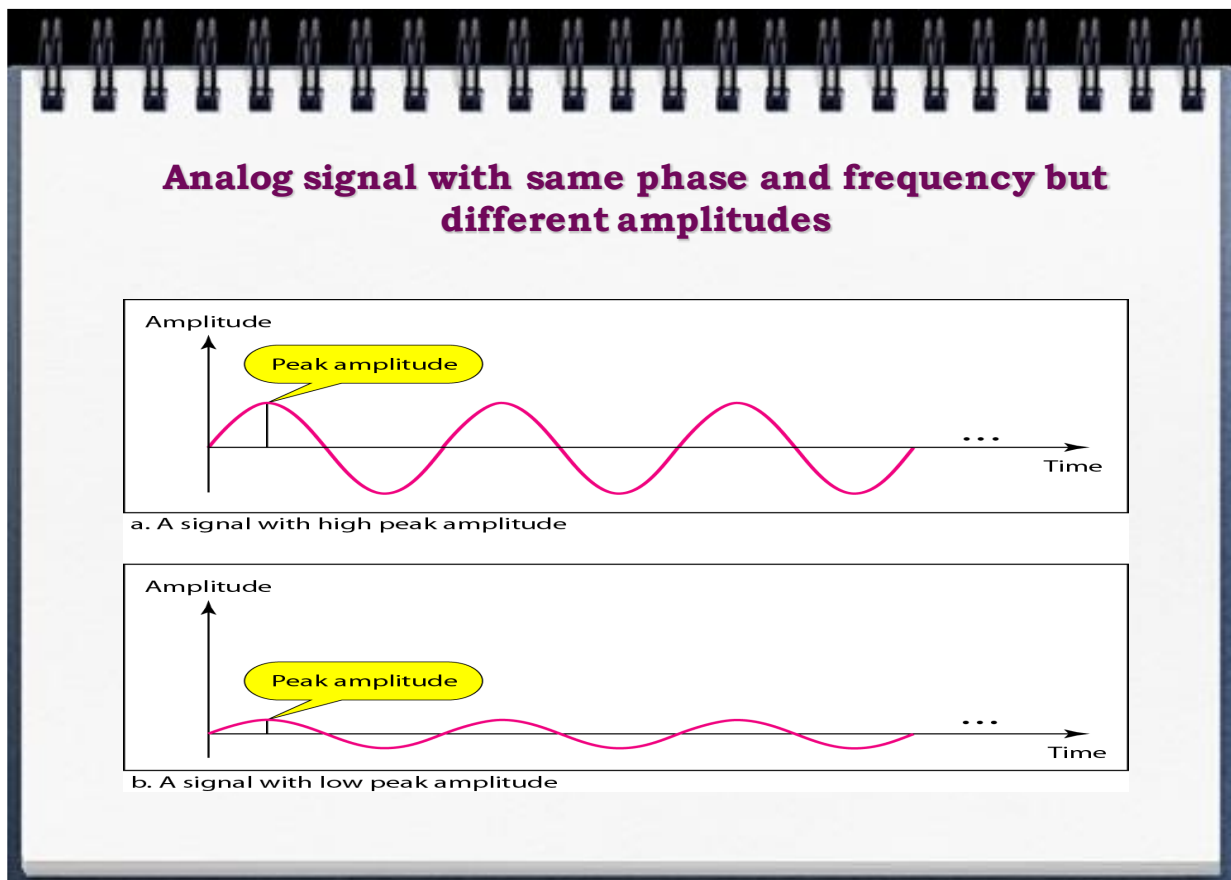
An **analog signal** refers to the information that is continuous. The analog speech signal varies continuously in its pressure and time properties. E.g. cassette tapes, vinyl records, FM and AM radio transmissions. The advantage of stored analog signal in magnetic tape recording is that it can be played back for listening and analysis. Playback is accomplished by converting the magnetic energy back to electrical energy, which in turn is converted to acoustic energy by loudspeakers or headphones.

Advantages of analog signal:

- It has potential of infinite resolution of the signal.
- Processing of analog signals is simple.

Disadvantages of analog signal:

- These signals are more susceptible to noise – as the signal is copied and re-copied or transmitted over long distances random variations occur.
- These signals are impossible to recover from noise/distortion.



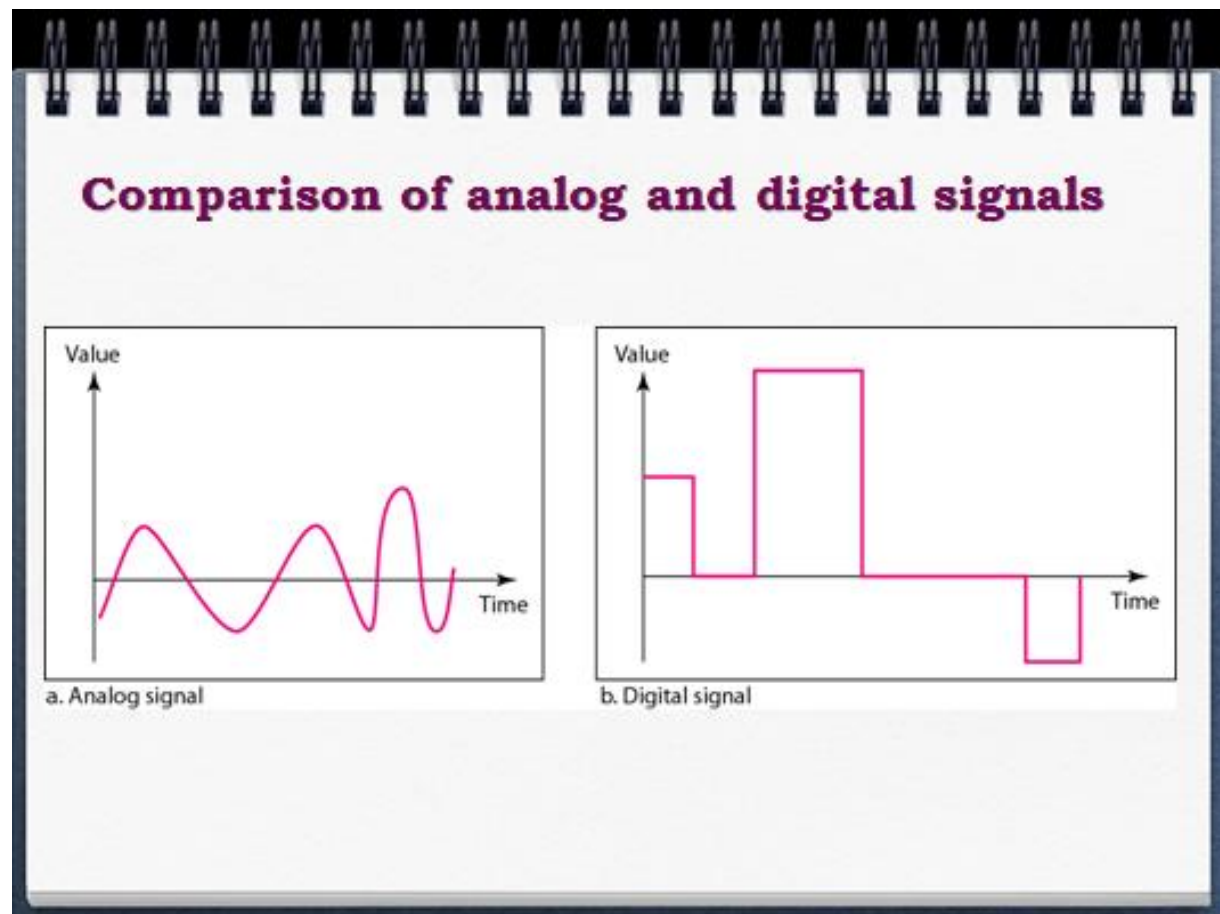
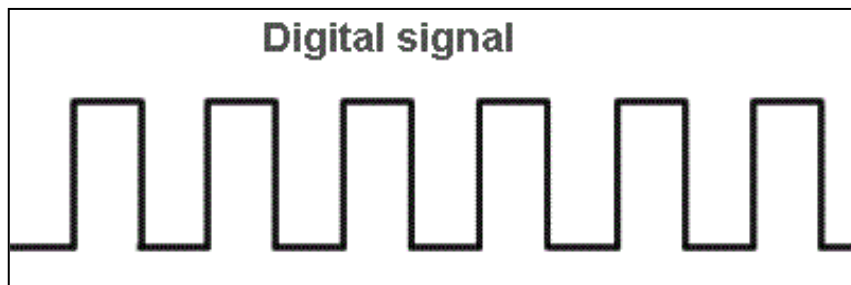
Digital signal is the one which is represented as a sequence of numbers in form of binary code. These are stream of discrete numbers that represent instantaneous amplitudes of analog signal, measured at equally spaced points in time. The digital data is represented in terms of bit and bytes.

Advantages of digital signal:

- Degradation of the digital signal (i.e. noise) can not only be detected but corrected as well.

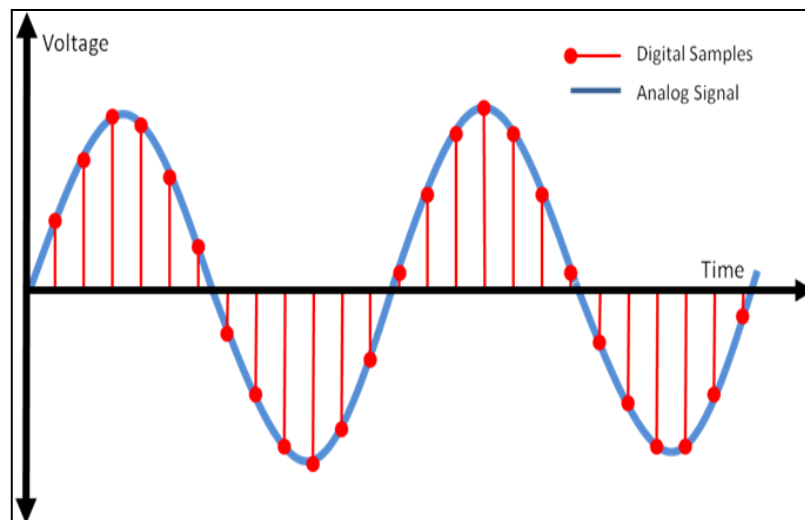
Disadvantages of digital signal:

- These signals are prone to errors (due to quantization and sampling)
- It has lower resolution than analog signals



❑ To store a speech signal in a computer, it is necessary to convert analog data into a digital format. This process of analog to digital conversion is called digitization process. While the conversion of digital data to analog representation is facilitated by reverse digitization process. Modem is used to perform such task. The digital representation of speech is very important because it permits the analysis of speech signal, employing the computational power of modern computers. Basic operations in digitizing a speech signal:

1. **Sampling:** This is a process by which an analog signal is converted into series of samples. This is basically done to convert the continuous amplitude/time signals to discrete amplitude/time signals. A sampling rate of 10 kHz means the original analog signal is sampled 10,000 times per second.



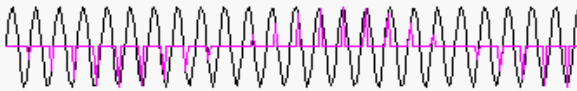
Sampling rate of the signal is chosen according to **Nyquist sampling theorem** which states that if the sampling rate is properly selected, the sampled signal contains the same information as the original signal. The sampling rate should be atleast twice the highest frequency f interest. E.g. if the highest frequency of interest is 10 kHz, then the sampling rate should be $2 \times 10\text{kHz} > 20\text{kHz}$. Sampling at much higher rate is called as **Oversampling**, which is unnecessary and will use more computer memory while sampling at rate lower than highest frequency of interest is called **Undersampling**, causing aliasing.

Aliasing

What happens if sampling rate not high enough?



A high frequency signal



sampled at too low a rate



looks like ...

This causes aliasing, meaning distortion caused in audio signal due to error in sampling rate. Different samples become indistinguishable.

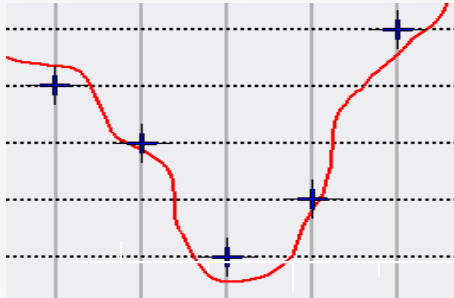
Common Sampling Rates

Sampling Rate	Uses
8kHz (8000)	Telephone
44.1 kHz (44100)	CD, DAT
48 kHz (48000)	DAT, DVD-Video
96 kHz (96000)	DVD-Audio

- Quantization:** A signal is quantized when the samples determined by the sampling operation are chopped into discrete number of amplitude levels. A quantum is an increment of energy. When an analog signal is quantized, the continuous amplitude variations are converted to discrete values or increments. Higher the number of quantization levels, the more accurately the quantized signal represents the analog signal. Speech should be quantized with atleast a 12-bit conversion. 12-bit = $(2)^{12} = 4,096$ quantization level. If too few levels of quantization are used, the signal will have distortion called quantization noise.

Quantization Noise

Round-off error: difference between actual signal and quantization to integer values...



Random errors: sounds like low-amplitude noise

Common Quantization rates

Word length	Uses
8-bit integer	Low-res web audio
16-bit integer	CD, DAT, sound files
24-bit integer	DVD-Video, DVD-Audio
32-bit	Software (usually only for internal representation)

