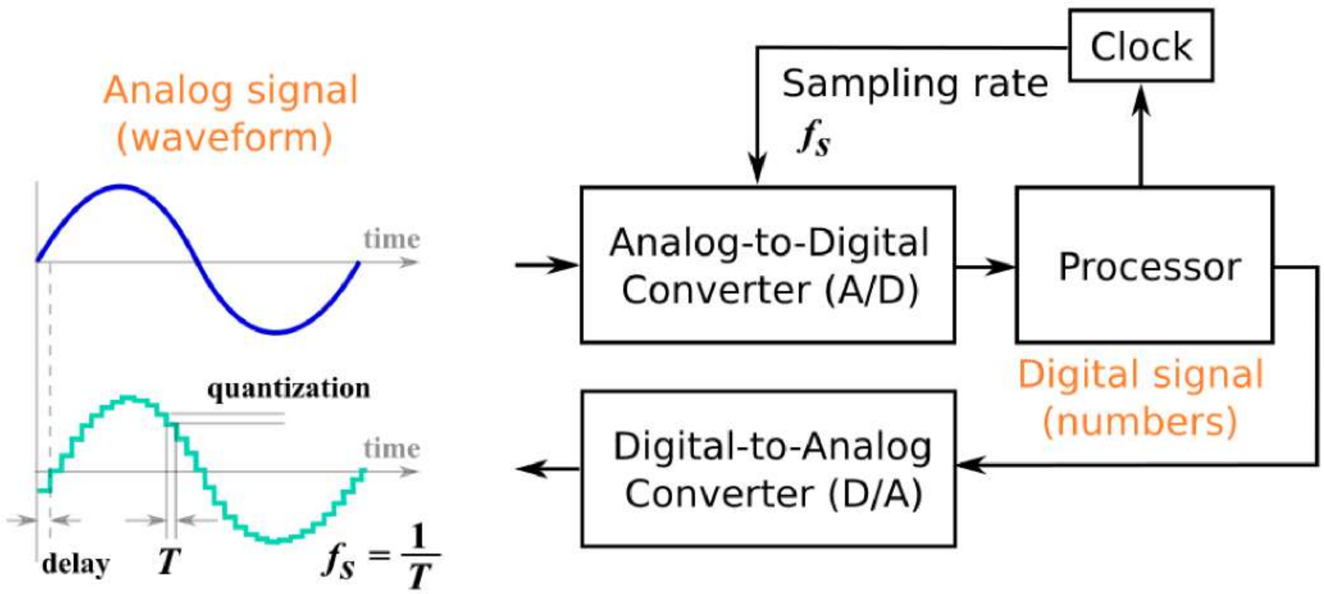
# Sampling Theory and MP3

Ying Sun

## Analog-Digital System




An analog-digital system is shown by the block diagram above. An analog signal is acquired by use of an analog-to-digital converter (A/D). The output of the A/D is the digital signal, a time-sequence of numbers. The processor can carry out digital signal processing and sends the result to the digital-to-analog converter (D/A). The output of the D/A is an analog signal but with a quantization effect on both the amplitude and the time. Due to the processing time, there is also a time delay from the input analog waveform to the output analog waveform.

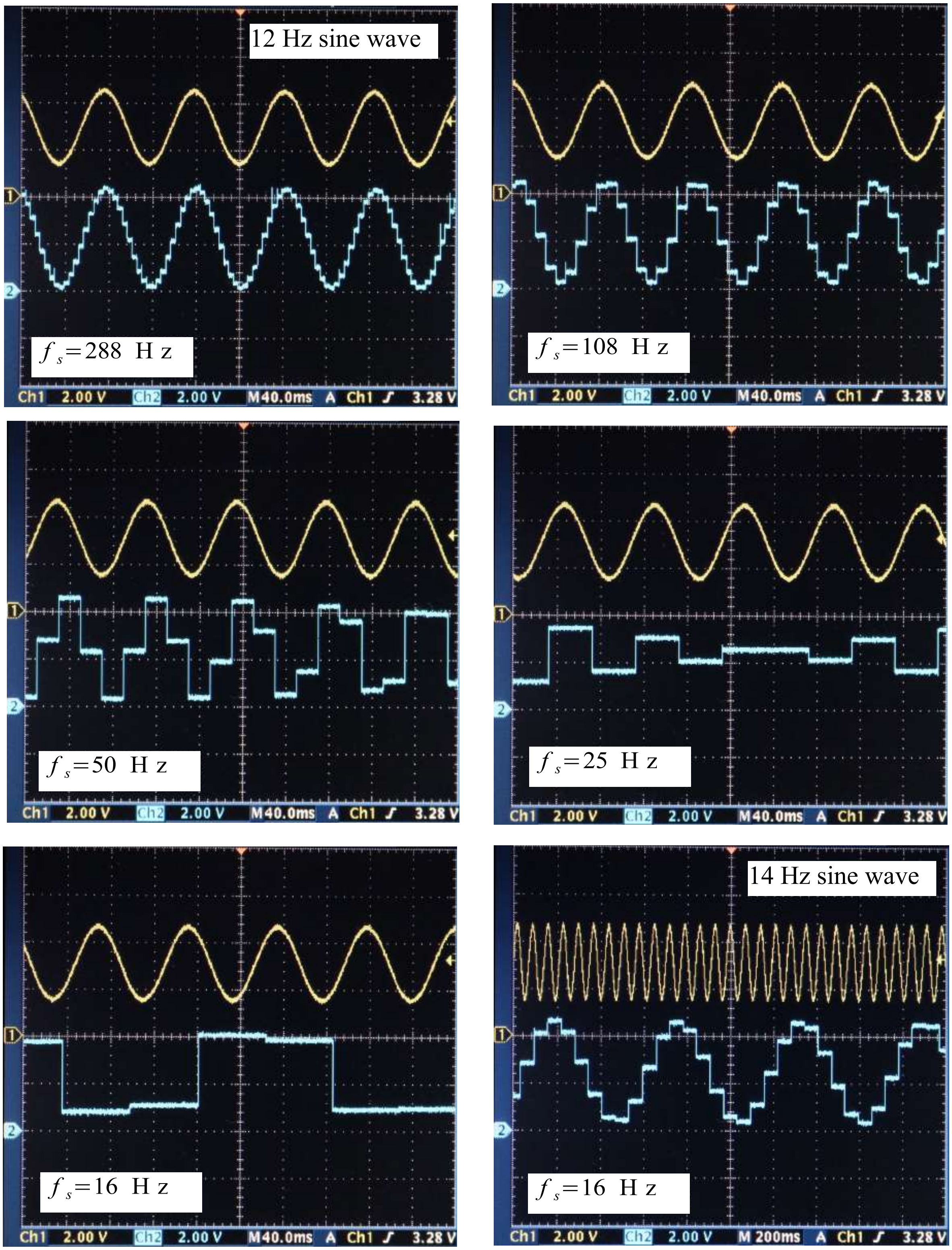
The A/D has two important parameters. First, the sampling rate $f_s$ is the frequency of taking periodical samples. The sampling interval $T$ is the time between any two adjacent sample points, and has a reciprocal relationship with the sampling frequency: $f_s = 1/T$. Second, the quantization of the signal amplitude is determined by how many bits that the A/D provides. For example, a 16-bit A/D provides a quantization of $2^{16} = 65536$ levels. In other words, the output of the A/D is a sequence of numbers, with each number between 0 and 65535. The dynamic range of the 16-bit A/D is $20 \log_{10} 65536 = 96$ dB.

You may wonder why $20 \log_{10}$, not $10 \log_{10}$. There are actually two definitions for dB. If the quantity $y$ is an amplitude, the formula $20 \log_{10} y$ is used to define dB. If the quantity $x$ is a power, the formula $10 \log_{10} x$ is used to define dB. The amplitude and power are related by a square relationship: $x = y^2$. We see that $10 \log_{10} x = 10 \log_{10} y^2 = 20 \log_{10} y$. Thus, once converted to dB, it doesn't matter whether the original quantity is amplitude or power. The resulting dB is the same.

Previously we define the sound level in dB by using the $10 \log_{10} x$. This is because the sound level $x$ is the sound intensity, which is the sound power per unit area $\mathrm{W}/m^2$. The sound intensity of $10^{-12} \mathrm{W}/m^2$ is the reference level of 0 dB.

## Examples of Samplings

Each of the following panels shows a 12-Hz input waveform (top) and the output (bottom) after sampling at a different sampling rate. The lower-right panel shows the phenomenon of aliasing, which is a miss representation of a signal frequency due to under sampling.



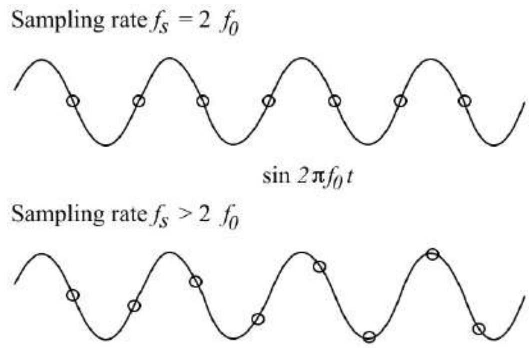
### The Nyquist-Shannon Sampling Theorem

In order to represent the highest frequency component $f_0$ in an analog signal, the sampling rate $f_s$ needs to be at least twice higher than $f_0$, i.e. $f_s > 2f_0$.

To illustrate this, consider the following example shown in the figure: A sine wave has frequency $f_0$. If we choose a sampling rate right at the Nyquist limit, which is $2f_0$, we get two sample points on each cycle. We could get all 0's and fail to represent the signal. However, if we sample at a rate > $2f_0$, we will avoid the straight line scenario and can represent that frequency.

Harry Nyquist (1889-1976, Sweden-USA)

Claude Elwood Shannon (1916-2001, USA)




### Compact Disc (CD)

CD is a digital optical disc data storage format that was co-developed by Philips and Sony and released in 1982. The standard CD has a diameter of 120 mm (4.7 in) and can hold up to about 80 minutes of uncompressed audio or about 700 MB of data (MB: mega bytes = 1,000,000 bytes; 1 byte = 8 bits). The industrial standard for the audio CD specifies a quantization level of 16 bits at a sampling rate of 44.1 KHz.

The choice of the sampling rate is a compromise between fidelity and storage space. The humans' audible range is between 20 Hz and 20 KHz. According to the Sampling Theorem, the sampling rate needs to be greater than 40 KHz in order to represent the highest audible frequency. Thus, the choice of 44.1 KHz is close to bare minimum. At the time the CD standard was established, data compression technologies were not fully developed. Thus, for 16-bit 44.1 KHz stereo music the CD has a storage capacity comparable to that of the LP record. The computation for the storage time of a CD is as follows:

700 MB / [2 byes (16 bits) × 44100 Hz × 2 channels] = 3868 s = 1 hr 6 min

(enter “=700000000/(2*44100*2)” in the spreadsheet)

### MP3

The uncompressed audio files (.wav or .aiff) are too large to be efficiently transferred over the Internet. MP3 is a coding format for digital audio that has a data compression capability. The formal name for MP3 is MPEG-1 Audio Layer III. The Moving Picture Experts Group (MPEG) is a working group of authorities that was formed by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) to set standards for audio and video compression and transmission. The MP3 is a lossy compression algorithm that uses inexact approximations and partial data discarding to represent the content. The data compression methods include discrete cosine transform, difference encoding, Huffman coding, and run-length encoding.

Uncompressed audio as stored on an audio CD has a bit rate of 1,411.2 Kbit/s (16 bits × 44100 Hz × 2 channels). The MP3 can compress to different bitrates such as 128, 160 and 192 Kbit/s, which represent compression ratios of approximately 11:1, 9:1 and 7:1, respectively.