AL FURAT AL AWSAT TECHNICAL UNIVERSITY

NAJAF COLLEGE OF TECHNOLOGY

DEPARTMENT OF TECHNICAL COMMUNICATIONS ENGINEERING

DIGITAL SIGNAL PROCESSING

3rd YEAR

Academic Responsible

HAYDER S. RASHID

2015/2016

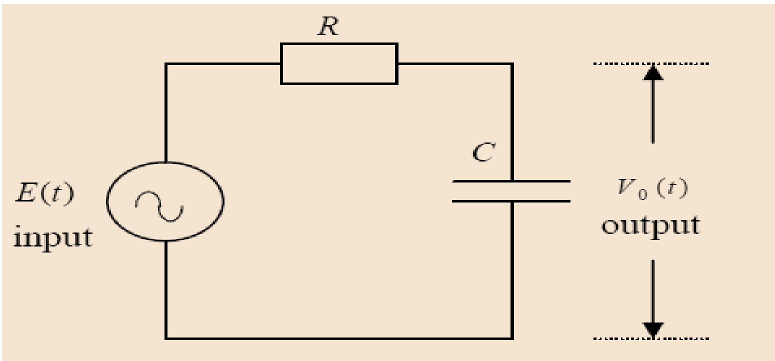
Continuous and discrete-time systems

Definition of system

A system operates on a signal (the input) to modify, transform or re-express it in another form (the output) which may be more desirable.

Example of system

The input voltage *E(t)* produces an output voltage $V_0(t)$.

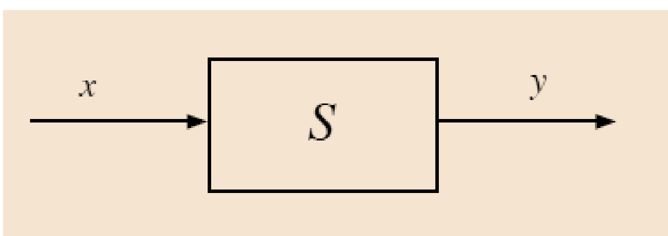


A continuous-time system is a system in which continuous-time input signals are applied and result in continuous-time output signals. A discrete-time system is a system in which discrete-time input signals are applied and result in discrete-time output signals.

Notation:

- y = Sx or y = S(x), meaning the system S acts on an input signal x to produce output signal y.
- y = Sx does not (in general) mean multiplication!

Systems often denoted by block diagram:

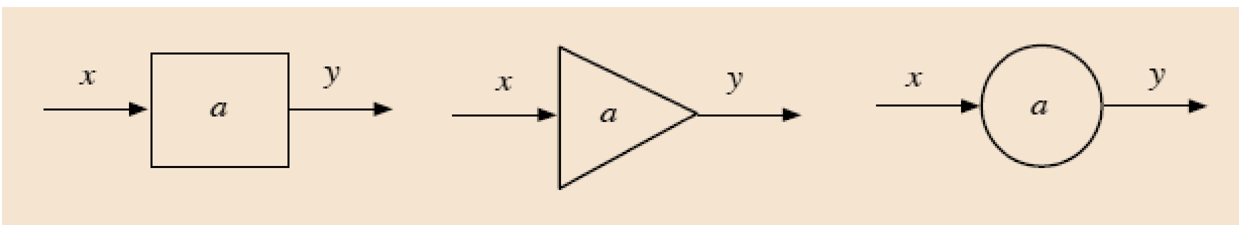


- Lines with arrows denote signals (not wires).
- Boxes denote systems; arrows show inputs & outputs.
- Special symbols denote for some systems.

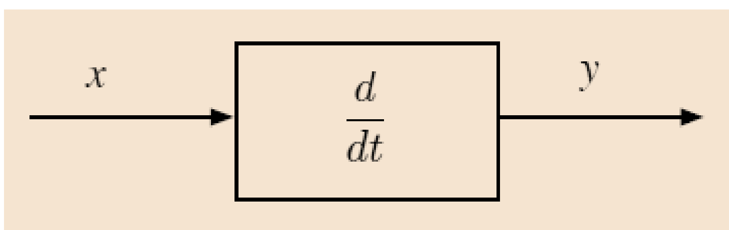
Examples (with input signal x and output signal y)

(i) Scaling system: y(t) = ax(t)

- Called an amplifier if $|a| > 1$.
- Called an attenuator if $|a| < 1$.
- Called inverting if **a** < **0**.
- ***a*** is called the gain or scale factor.
- Sometimes denoted by triangle or circle in block diagram:

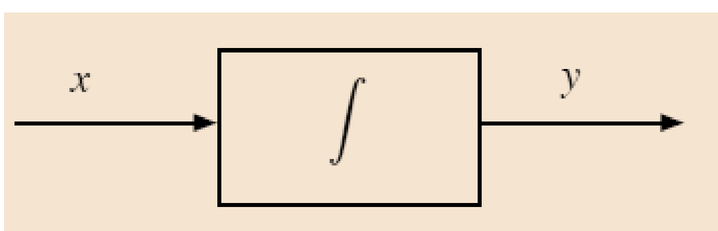


(ii) Differentiator: y(t) = x'(t)



(iii) **Integrator:** y(t) $=\int_a^t x(\tau)d\tau$, (**a** is often 0 or $-\infty$)

Common notation for integrator:



(iv) Time shift system: y(t) = x(t - T)

- called a delay system if $T > 0$
- called a predictor system if $T < 0$

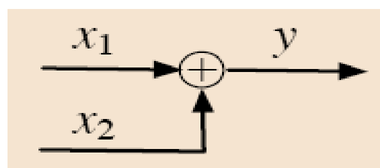
(v) Convolution system:

$$y(t) = \int x(t-\tau)h(\tau)d\tau$$

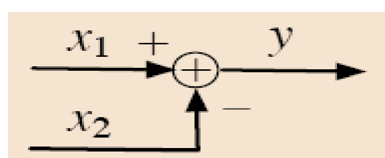
Where h is a given function (you'll be hearing much more about this!).

Examples (with multiple inputs, inputs $x_1(t)$, $x_2(t)$, and Output y(t)**)**

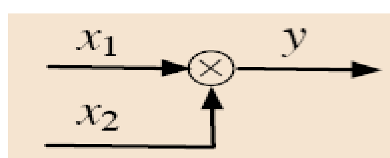
(i) Summing system: $y(t) = x_1(t) + x_2(t)$



(ii) Difference system: $y(t) = x_1(t) - x_2(t)$

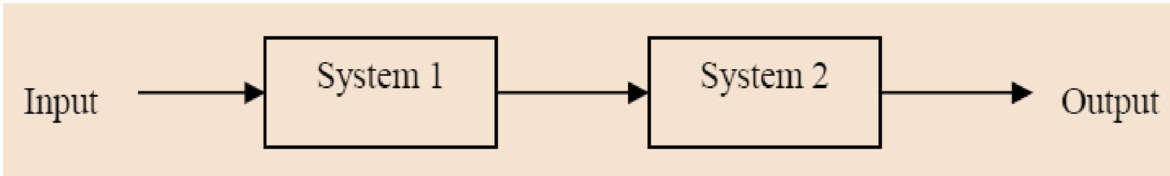


(iii) Multiplier system: $y(t) = x_1(t)x_2(t)$

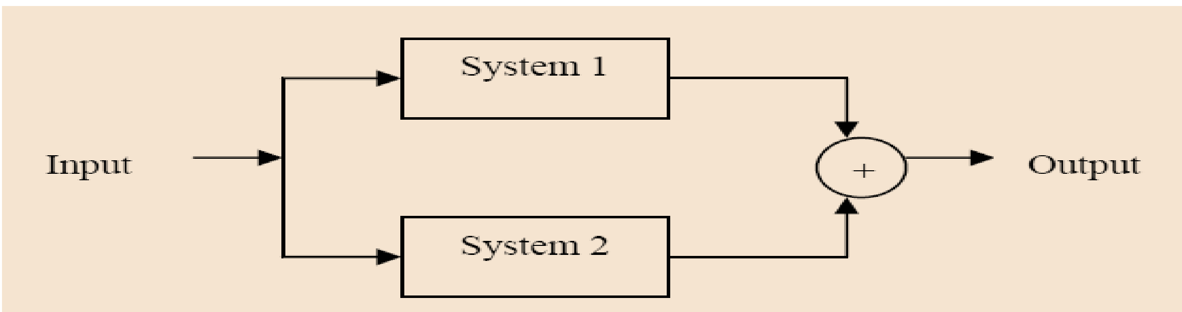


Interconnection of systems

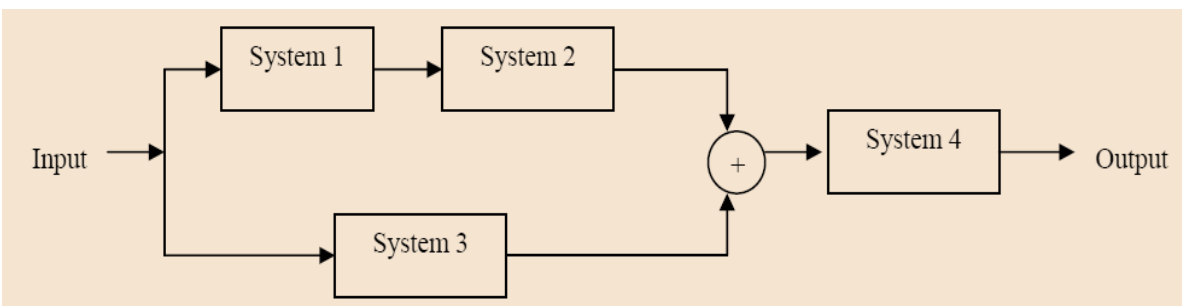
1- **Series or cascade interconnection**. The output of System 1 is the input to System 2.



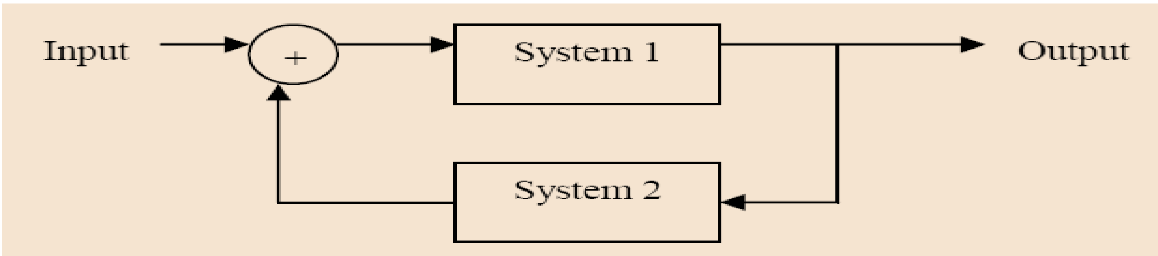
2- **Parallel interconnection**. The same input signal is applied to Systems 1 and 2.



3- **Combination of both cascade and parallel interconnection.**



4- **Feedback interconnection**. The output of System 2 is fed back and added to the external input to produce the actual input to System 1.



Basic system properties

1- Systems with and without memory

The following classification can be made.

Memoryless systems; The output at instant **t** for continuous-time systems or at instant **n** for discrete-time systems depends only on the value of input at the same instant **t** or **n**. For example $y[n] = \alpha x[n]$ or $y(t) = \alpha x(t)$.

Systems with memory; The output at instant **t** for continuous-time systems or at instant **n** for discrete-time systems depends not only on the value of input at the same instant **t** or **n**, but also on past or future values of the input. For example $y[n] = \alpha x[n] + \beta x[n-1]$.

2- Invertible and non-invertible systems

The following classification can be made.

Invertible systems. A system is called invertible, if another system exists that is called inverse that, when cascaded with the original system, yields an output equal to the input to the first system.

Example

Consider the memoryless system described by the input-output relationship $y[n] = \alpha x[n]$, $\alpha \neq 0$. This is invertible with inverse $x[n] = 1/\alpha \; y[n]$

Non-invertible systems. A system is called non-invertible, if no system exists that, when cascaded with the original system, yields an output equal to the input to the first system.

Example

Consider the memoryless system $y[n] = x^2[n]$. This is non-invertible, since from knowledge of the output we cannot determine the sign of the input.

Obviously a system is invertible if distinct inputs lead to distinct outputs.

3- Causal and non-causal systems

The following classification can be made.

Causal or non-anticipative systems. The output at instant **t** for continuous-time systems or at instant **n** for discrete-time systems depends on the value of input at the same instant **t** or **n** and on past values of the input but not on future values of the input.
For example $y[n] = \alpha x[n] + \beta x[n-1]$.
Non-causal or anticipative systems. The output at instant **t** for continuous-time systems or at instant **n** for discrete-time systems depends on future values of the input.
For example $y[n] =\alpha x[n]+ \beta x[n +1]$.

4- Stable and unstable systems

Stable systems. A stable system is one in which all bounded inputs lead to bounded outputs. For example consider the system $y[n] = e^{x[n]}$. Let ***B*** be an arbitrary positive number, and let x[n] an arbitrary signal bounded by B, i.e., $|x[n]| < B$ ***or*** $-B < x[n] < B$ for all n. In that case $e^{-B} < e^{x[n]} < e^{B}$, i.e., it is guaranteed that if the input is bounded by a positive number B, the output is bounded by a positive number e^{B}. Thus, the system is stable.

Unstable systems. An unstable system is one in which not all bounded inputs lead to bounded outputs. For example consider the system $y[n] = \sum_{k=-\infty}^{n} u[k] = (n+1)\, u[n]$, with u[n] the unit step function. In that case the input u[n] = 1 but the output y[n] increases without bound as n increases.

5- Time-invariant and time-varying systems

Time-invariant (TI) systems. A TI system is one in which if y[n] is the output when the input x[n] is applied, then $y[n-n_0]$ is the output when $x[n–n_0]$ is applied. For continuous-case we would have the signals y(t), x(t), $y(t-t_0)$, $x(t-t_0)$ respectively. Conceptually, a system is TI if the behavior and the characteristics of the system are fixed over time. For example the system $y[n] =\alpha x[n]$.

Time-varying systems. A time-varying system is one in which if y[n] is the output when the input x[n] is applied, then $y[n-n_0]$ is not necessarily the output when $x[n-n_0]$ is applied. For example consider the system $y[n] = nx[n]$. In that case if $y_1[n]$ is the output to the input $x_1[n]$, then the output to the input $x_2[n] = x_1[n-n_0]$ is $y_2[n] = nx_2[n] = nx_1[n-n_0] \neq (n-n_0)\, x_1[n-n_0] = y_1[n-n_0]$. Thus, the system is not time-invariant.

6- Linear and non-linear systems

Linear systems. A linear system is one that possesses the property of superposition, i.e., if the input consists of the weighted sum of several signals, then the output is the weighted sum of the responses of the system to each of those signals. Thus, a linear system possesses the following two properties.

1. Additivity property. If the response to $x_1(t)$ is $y_1(t)$ and the response to $x_2(t)$ is $y_2(t)$ then the response to the signal $x_1(t) + x_2(t)$ is $y_1(t) + y_2(t)$.

2. Homogeneity property. If the response to $x_1(t)$ is $y_1(t)$, then the response to the signal $\alpha x_1(t)$ is $\alpha y_1(t)$.

Non-linear systems. At least one of the above properties does not hold.

Sampling of Continuous Signal

Figure (1) below shows an analog (continuous-time) signal (solid line) defined at every point over the time axis and amplitude axis. Hence, the analog signal contains an infinite number of points. It is impossible to digitize an infinite number of points. Furthermore, the infinite points

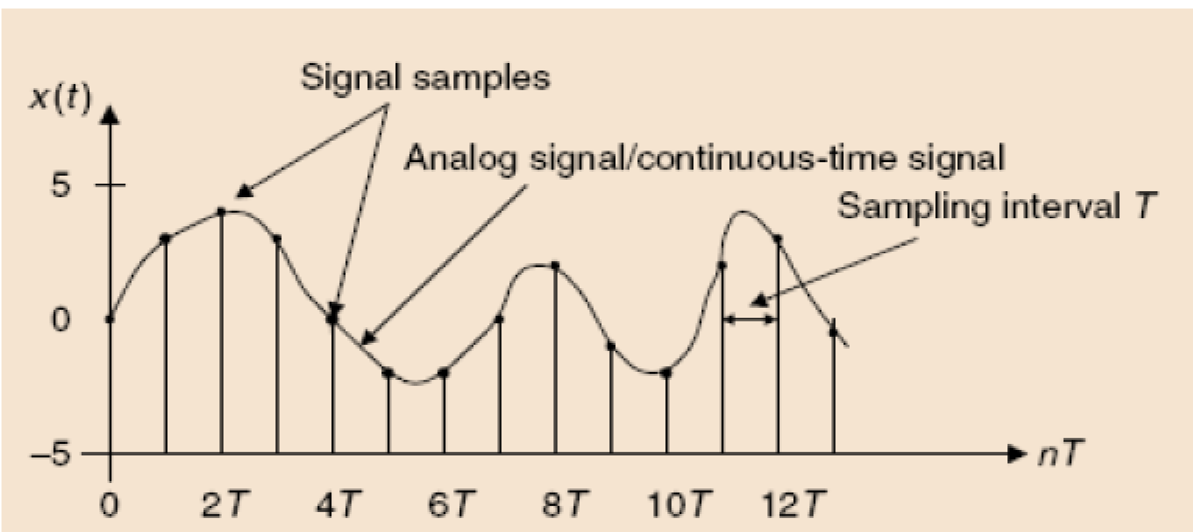


Figure 1 Display of the analog signal and display of digital samples versus the sampling time instants

are not appropriate to be processed by the digital signal (DS) processor or computer, since they require an infinite amount of memory and infinite amount of processing power for computations. Sampling can solve such a problem by taking samples at the fixed time interval, as shown in figure (1) and figure (2), where the time T represents the sampling interval or sampling period in seconds.

As shown in Figure (2), each sample maintains its voltage level during the sampling interval T to give the ADC enough time to convert it. This process is called sample and hold.

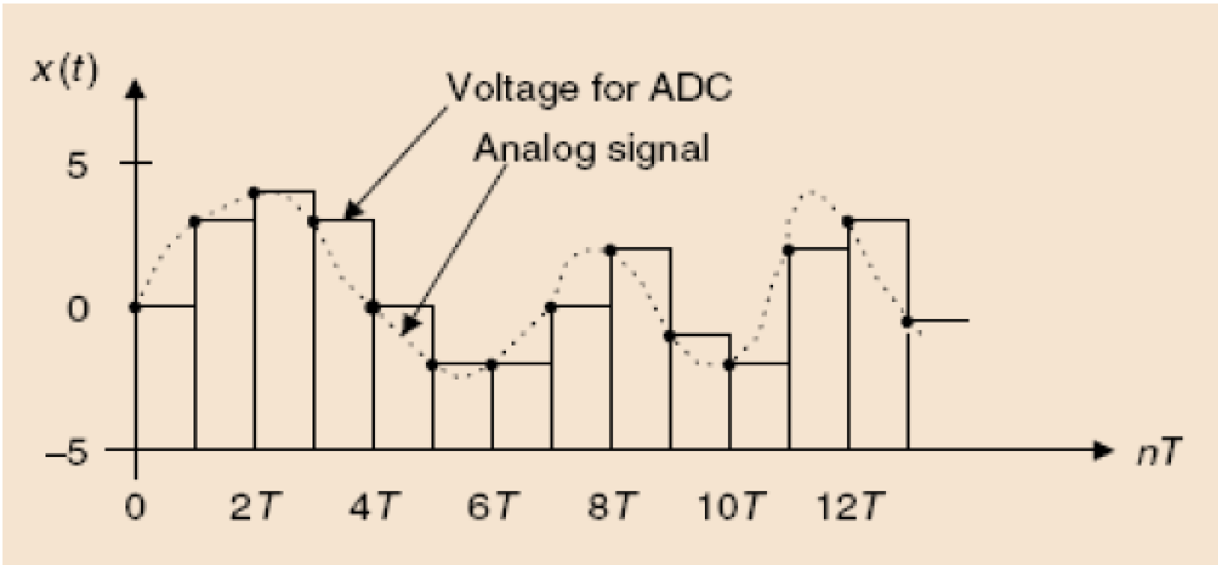


Figure 2 Sample-and-hold analog voltage for ADC

For a given sampling interval T, which is defined as the time span between two sample points, the sampling rate is therefore given by:

$$f_s = \frac{1}{T_s} \quad \text{Samples per second (Hz)} \qquad (1)$$

After the analog signal is sampled, we obtain the sampled signal whose amplitude values are taken at the sampling instants, thus the processor is able to handle the sample points. Next, we have to ensure that samples are collected at a rate high enough that the original analog signal can be reconstructed or recovered later.

In other words, we are looking for a minimum sampling rate to acquire a complete reconstruction of the analog signal from its sampled version.

If an analog signal is not appropriately sampled, aliasing will occur, which causes unwanted signals in the desired frequency band.

The sampling theorem guarantees that an analog signal can be in theory perfectly recovered as long as the sampling rate is at least twice as large as the highest-frequency component of the analog signal to be sampled. The condition is described as:

$$\boldsymbol{f_s \geq 2f_{max}} \qquad (2)$$

Where, f_{max} is the maximum-frequency component of the analog signal to be sampled. For example, to sample a speech signal containing frequencies up to 4 kHz, the minimum sampling rate is chosen to be at least 8 kHz, or 8,000 samples per second; to sample an audio signal possessing frequencies up to 20 kHz, at least 40,000 samples per second, or 40 kHz, of the audio signal are required.

Figure (3) depicts the sampled signal $x_s(t)$ obtained by sampling the continuous signal x(t) at a sampling rate of f_s samples per second.

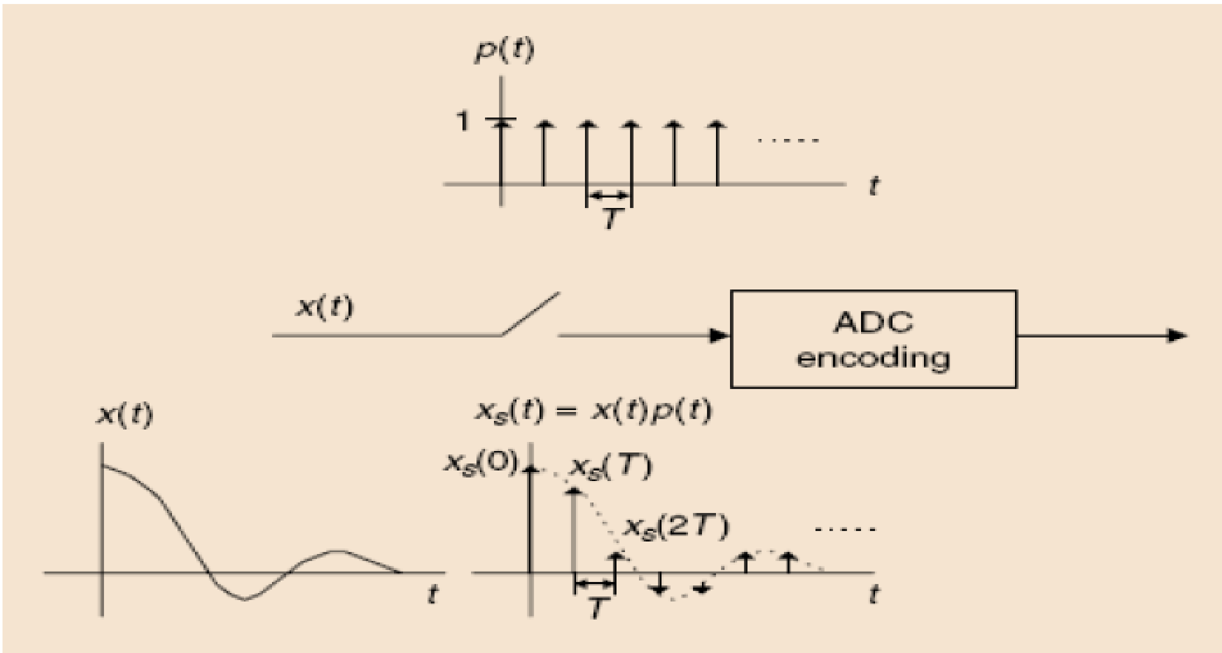


Figure 3 The simplified sampling process

Mathematically, this process can be written as the product of the continuous signal and the sampling pulses (pulse train):

$$\mathbf{x_s(t) = x(t)\, p(t)} \qquad (3)$$

Where, p(t) is the pulse train with a period $T = 1/ f_s$.

From the spectral analysis shown in Fig.(4), it is clear that the sampled signal spectrum consists of the scaled baseband spectrum centered at the origin and its replicas centered at the frequencies of $\pm nf_s$ (multiples of the sampling rate) for each of n = 1,2,3, . . . In Figure (4), three possible sketches are classified. Given the original signal spectrum X(f) plotted in Figure (4.a), the sampled signal spectrum is plotted in Figure (4.b), where, the replicas have separations between them. In Fig. (4.c), the baseband spectrum and its replicas are just connected. In Fig. (4.d), the original spectrum and its replicas are overlapped; that is, there are many overlapping portions in the sampled signal spectrum.

If applying a lowpass reconstruction filter to obtain exact reconstruction of the original signal spectrum, equation (2) must be satisfied. This fundamental conclusion is well known as **the Shannon sampling theorem**, which is formally described below:

For a uniformly sampled DSP system, an analog signal can be perfectly recovered as long as the sampling rate is at least twice as large as the highest-frequency component of the analog signal to be sampled.

Notes

1- Sampling theorem establishes a minimum sampling rate for a given bandlimited analog signal with the highest-frequency component f_{max}. If the sampling rate satisfies equation (2), then the analog signal can be recovered via its sampled values using the lowpass filter, as described in Fig. (4.b).

2- Half of the sampling frequency (f_s / 2) is usually called the Nyquist frequency (Nyquist limit), or folding frequency. The sampling theorem indicates that a DSP system with a sampling rate of f_s can ideally sample an analog signal with its highest frequency up to half of the sampling rate without introducing spectral overlap (aliasing). Hence, the analog signal can be perfectly recovered from its sampled version as described in Fig. (4.c). Fig. (4.d) shows aliasing.

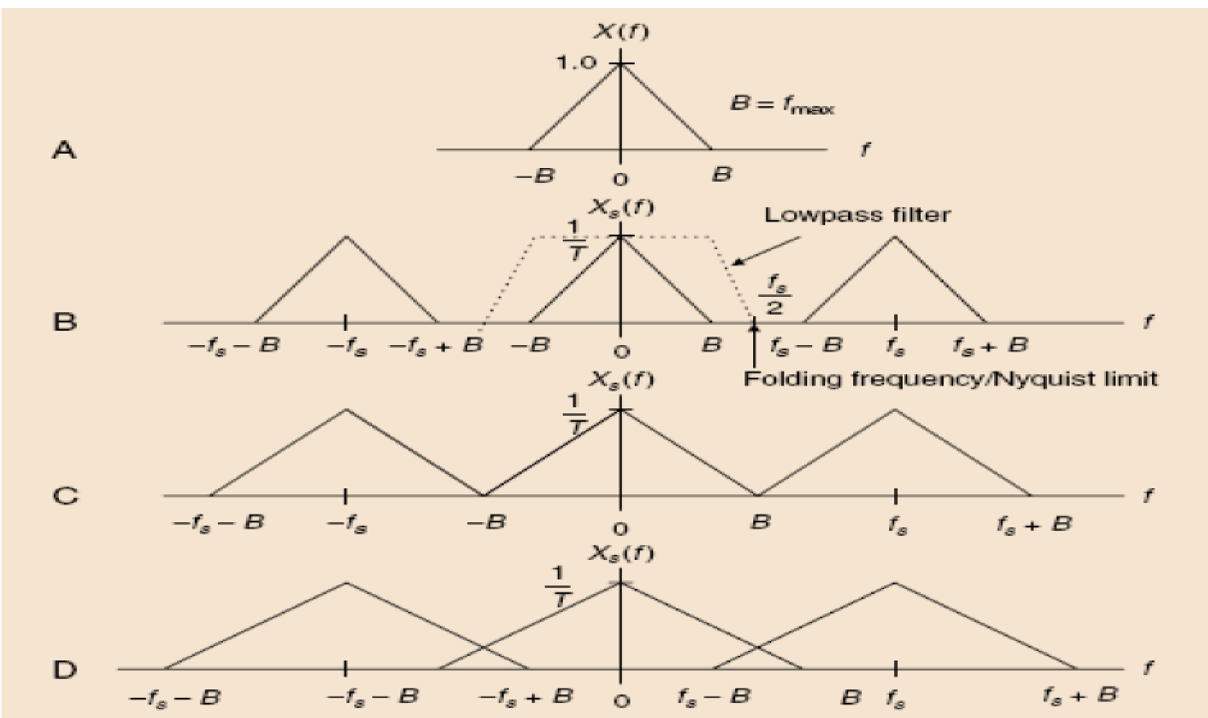


Figure 4 plots of the sampled signal spectrum

Example

Suppose that an analog signal is given as $x(t) = 5\cos(2\pi\,.\,1000t)\,, for\ t \geq 0$, and is sampled at the rate of 8,000 Hz.

a. Sketch the spectrum for the original signal.

b. Sketch the spectrum for the sampled signal from 0 to 20 kHz.

Solution:

$$5\cos(2\pi\,.\,1000t) = 5\,.\left(\frac{e^{j2\pi.1000t} + e^{-j2\pi.1000t}}{2}\right) = 2.5\,e^{j2\pi.1000t} + 2.5\,e^{-j2\pi.1000t}$$

The two-sided spectrum is plotted as shown in Fig. (5.a). After the analog signal is sampled at the rate of 8,000 Hz, the sampled signal spectrum and its replicas centered at the frequencies $\pm nf_s$, each with the scaled amplitude being 2.5/T, are as shown in Fig. (5.b)

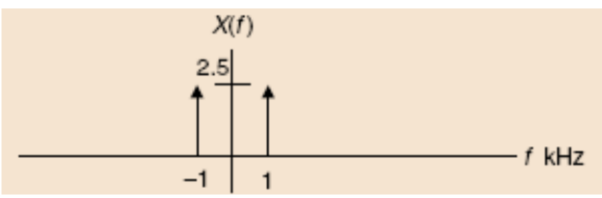


Fig.(5.a)

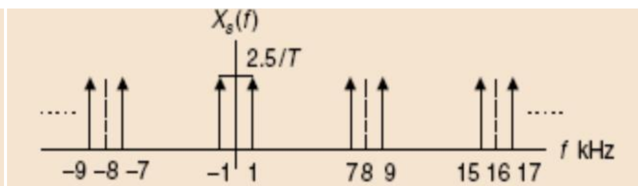


Fig.(5.b)

Notice that the spectrum of the sampled signal shown in Figure (5.b) contains the images of the original spectrum shown in Figure (5.a); that the images repeat at multiples of the sampling frequency fs (for our example, 8 kHz, 16 kHz, 24 kHz, . . .); and that all images must be removed, since they convey no additional information.

Signal Reconstruction

Two simplified steps are involved, as described in Figure (6). First, the digitally processed data y(n) are converted to the ideal impulse train ys(t), in which each impulse has its amplitude proportional to digital output y(n), and two consecutive impulses are separated by a sampling period of T; second, the analog reconstruction filter is applied to the ideally recovered sampled signal ys(t) to obtain the recovered analog signal.

The following three cases are listed for recovery of the original signal spectrum:

***Case 1*:** $f_s = 2f_{max}$

Nyquist frequency is equal to the maximum frequency of the analog signal x(t), an ideal lowpass reconstruction filter is required to recover the analog signal spectrum. This is an impractical case.

Case 2: $f_s > 2f_{max}$

In this case, there is a separation between the highest-frequency edge of the baseband spectrum and the lower edge of the first replica. Therefore, a practical lowpass reconstruction (anti-image) filter can be designed to reject all the images and achieve the original signal spectrum.

Case 3: $f_s < 2f_{max}$

This is aliasing, where the recovered baseband spectrum suffers spectral distortion, that is, contains an aliasing noise spectrum; in time domain, the recovered analog signal may consist of the aliasing noise frequency or frequencies. Hence, the recovered analog signal is incurably distorted.

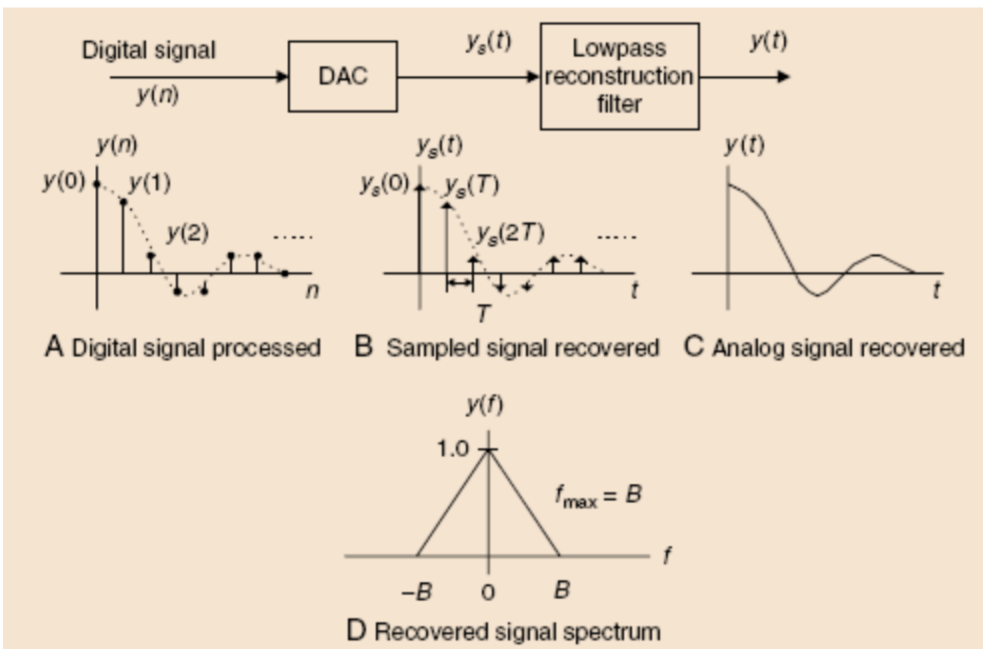


Figure 6 Signal notations at reconstruction stage

Example

Assuming that an analog signal is given by $x(t) = 5\cos(2\pi.2000t) + 3\cos(2\pi.3000t)$, $for\ t \geq 0$, and it is sampled at the rate of 8,000 Hz,

a. Sketch the spectrum of the sampled signal up to 20 kHz.

b. Sketch the recovered analog signal spectrum if an ideal lowpass filter with a cutoff frequency of 4 kHz is used to filter the sampled signal (y(n) = x(n) in this case) to recover the original signal.

Solution:

$$x(t) = \frac{3}{2}\, e^{-j2\pi.3000t} + \frac{5}{2}\, e^{-j2\pi.2000t} + \frac{5}{2}\, e^{j2\pi.2000t} + \frac{3}{2}\, e^{j2\pi.3000t}$$

The two-sided amplitude spectrum for the sinusoids is displayed in Figure (7.a). The recovered spectrum is shown in Fig. (7.b)

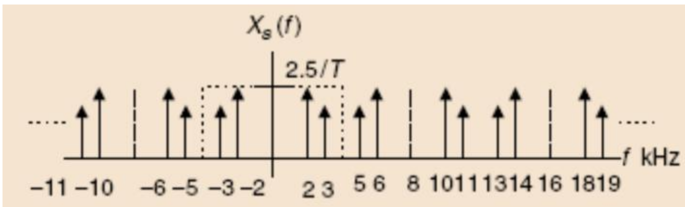


Fig.(7.a)

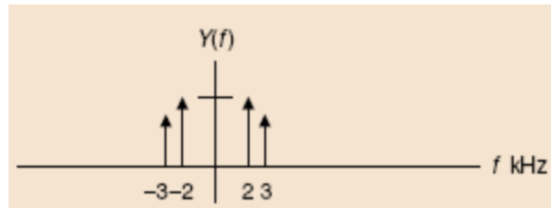


Fig.(7.b)

Aliasing noise level

Given the DSP system shown in Figure (8), where we can find the percentage of the aliasing noise level using the symmetry of the Butterworth magnitude function and its first replica. Then:-

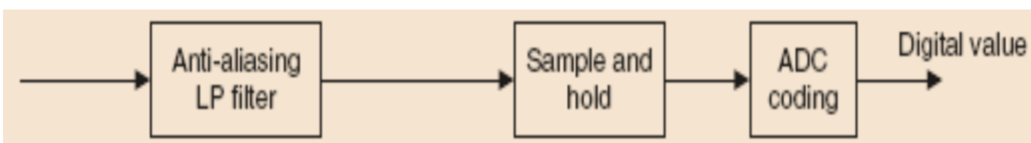


Figure 8 DSP system with anti-aliasing filter

$$\text{Aliasing noise level } \% = \frac{\sqrt{1+(f_a/f_c)^{2n}}}{\sqrt{1+(\frac{f_s-f_a}{f_c})^{2n}}}\ ,\ 0 \le f \le f_c$$

Where, n is the filter order, f_a is the aliasing frequency, f_c is the cutoff frequency, and f_s is the sampling frequency.

Example

In a DSP system with anti-aliasing filter, if a sampling rate of 8,000 Hz is used and the anti-aliasing filter is a second-order Butterworth lowpass filter with a cutoff frequency of 3.4 kHz,

a. Determine the percentage of aliasing level at the cutoff frequency.

b. Determine the percentage of aliasing level at the frequency of 1,000 Hz.

Solution:

$$f_s = 8000,\ f_c = 3400,\ \text{and } n = 2.$$

a. Since $f_a = f_c = 3400$ Hz, we compute

$$\text{aliasing noise level } \% = \frac{\sqrt{1 + \left(\frac{3.4}{3.4}\right)^{2\times 2}}}{\sqrt{1 + \left(\frac{8-3.4}{3.4}\right)^{2\times 2}}} = \frac{1.4142}{2.0858} = 67.8\%.$$

b. With $f_a = 1000$ Hz, we have

$$\text{aliasing noise level } \% = \frac{\sqrt{1 + \left(\frac{1}{3.4}\right)^{2\times 2}}}{\sqrt{1 + \left(\frac{8-1}{3.4}\right)^{2\times 2}}} = \frac{1.03007}{4.3551} = 23.05\%.$$