

## MODULE-1

### INTRODUCTION TO DIGITAL SIGNAL PROCESSING

#### ❖ LEARNING OBJECTIVES

- A digital signal processing system
- The sampling process
- Discrete time sequences
- Discrete Fourier transform and Fast Fourier transform
- Digital filters
- Decimation and Interpolation
- Analysis and Design tools for DSP system-MATLAB

#### 1.1 What is DSP? [July... 2011, 8M]

- DSP is a technique of performing the mathematical operations on the signals in digital domain.
- As real time signals are analog in nature we need first convert the analog signal to digital, then we have to process the signal in digital domain and again converting back to analog domain.
- Thus ADC is required at the input side whereas a DAC is required at the output end.
- A typical DSP system is as shown in figure 1.1.



Fig 1.1 A typical DSP system

##### 1.1.1 Need for DSP

Analog signal Processing has the following drawbacks:

- They are sensitive to environmental changes
- Aging
- Uncertain performance in production units
- Variation in performance of units
- Cost of the system will be high
- Scalability

If digital Signal Processing would have been used we can overcome the above short

comings of ASP.

## 1.2 The important issues to be considered in designing and implementing a DSP system [DEC-2012, 9M]

- Complexity of algorithm: The arithmetic operations to be performed and the precision required are decided by the applications.
- Sample rate: The rate at which I/P samples are received and processed varies with the application and this rate along with the algorithm determines whether a particular dsp is suitable for a given application.
- Speed: It depends on the technology. To meet the specified through put requirements with the given sampling rate, it must be possible to operate the DSP at a particular clock rate.
- Data representation: The format and the number of bits used for data representation depend on the arithmetic precision and the dynamic range required.

## 1.3 Major features of Programmable Digital Signal Processors [Dec.. 2011, 6M]

- Multiply accumulate hardware: is the most used operation in DSP. To implement effectively the DSP has a hardware multiplier, an accumulator which holds the sum of product and an explicit multiply accumulate instruction.
- Harvard architecture: in Harvard memory architecture, there are two memory space, i.e program and Data memory. The core processor connects these memories by separate buses, allowing simultaneous access to memory. This arrangement helps to double the processor bandwidths.
- Zero overhead looping: the term Zero overhead looping means that the processor can execute loops without consuming cycles to test the value of the loop counter, performs loop counter, and decrement the loop counter.
- Special addressing: DSP processor often supports specialized addressing modes that are useful for common signal processing operations and algorithms.

## 1.4 Digital Signal Processing System [JAN-2015, 9M]

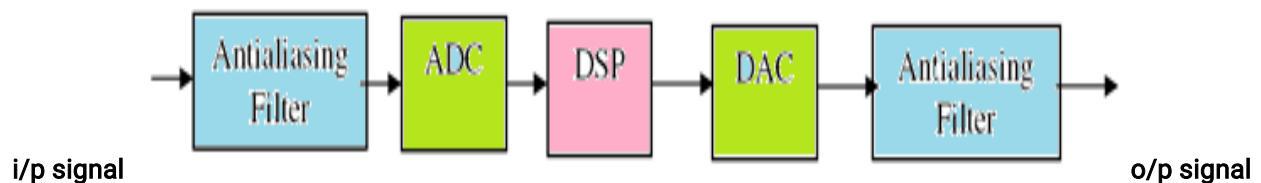


Fig 1.2 Block diagram of DSP system

- A Digital signal processor (DSP) system uses computer or a processor is used for digital signal processing.
- The DSP processor has analog front end and analog backend
- The front end contains anti aliasing filter, a sample and hold circuit and ADC feeding into the DSP
- The back end contains of a DAC to convert the digital o/p to its analog value followed by a reconstruction filter.
- Anti aliasing filter is a LPF which passes signal with frequency less than or equal to half the sampling frequency in order to avoid Aliasing effect in a samples spectrum.
- The sample and hold circuit prevent sample of the I/P of the ADC. It also hold these sample at constant level irrespective of the variation in the I/P signals.
- The analog output of DAC has staircase amplitude due to conversion process.
- Similarly at the other end, reconstruction filter is used to reconstruct the samples from the staircase output of the DAC (Figure 1.2).
- Reconstruction output removes high frequency noise due to the staircase output of the D/A CONVERTER.
- Signals that are appearing in a typical DSP are as shown in the figure 1.3.

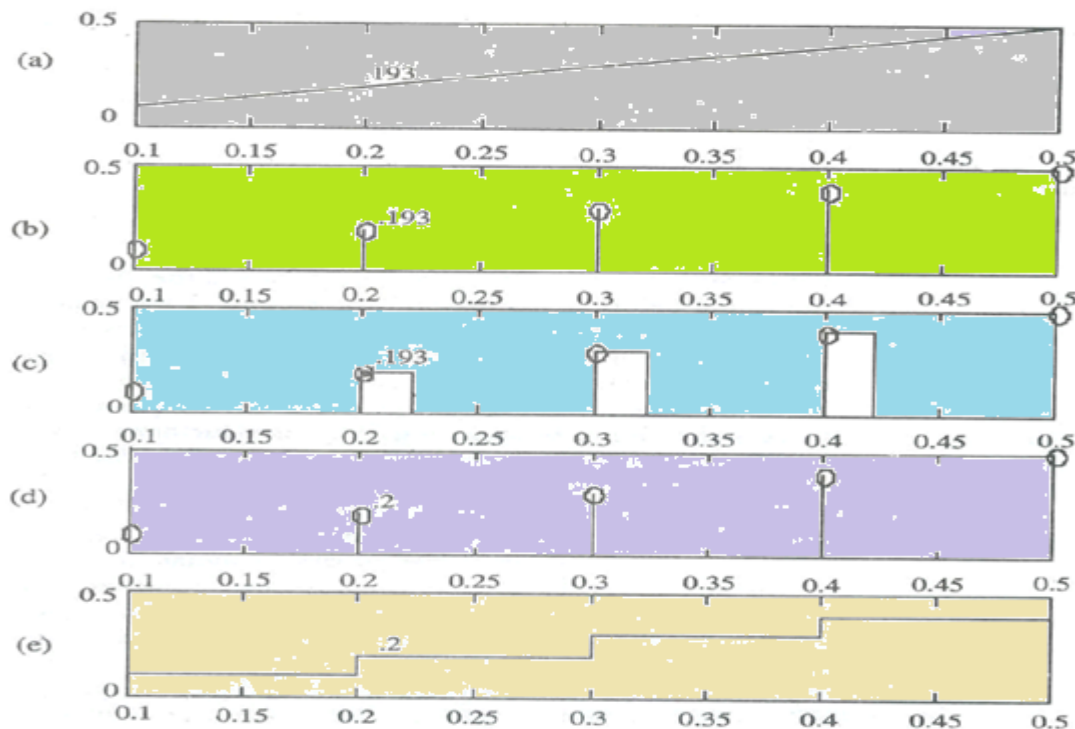


Fig:1.3 (a) continuous time signal (b) sampled signal (c) sampled data signal (d) quantized signal (e) DAC o/p signal

## 1.5 The Sampling Process

- ADC process involves sampling the signal and then quantizing the same to a digital value. In order to avoid Aliasing effect, the signal has to be sampled at a rate at least equal to the Nyquist rate.
- The condition for Nyquist Criterion is as given below,

$$f_s = 1/T = 2f_m$$

$$f_s = 1/T = 2f_m$$

Where,  $f_s$  is the sampling frequency,  
 $f_m$  is the maximum frequency component in the message signal,  
 $T$  is a sampling interval in seconds.

- If the sampling of the signal is carried out with a rate less than the Nyquist rate, the higher frequency components of the signal cannot be reconstructed properly.
- The plots of the reconstructed outputs for various conditions are as shown in figure 1.4.

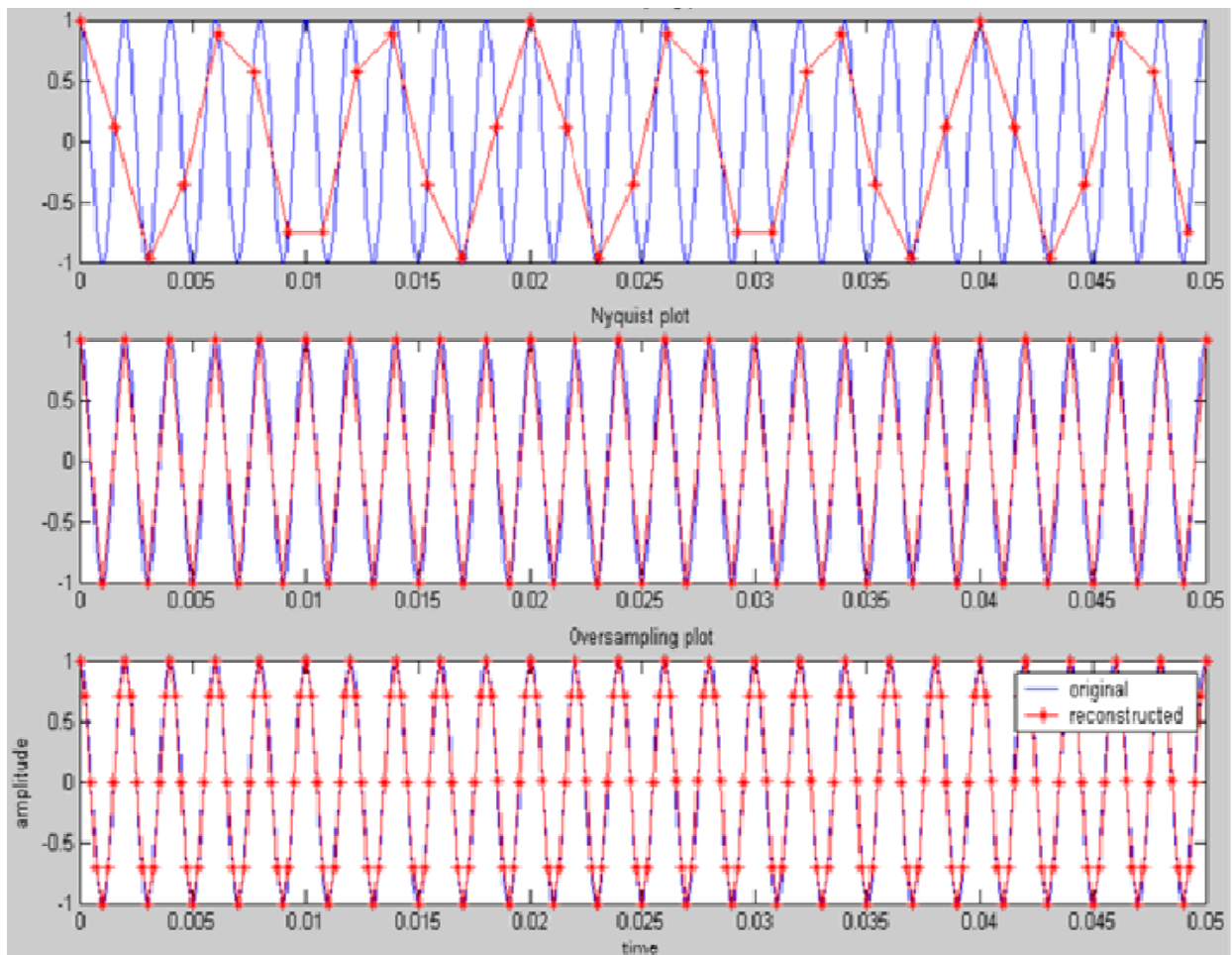


Fig:1.4 Verification of a sampling theorem

**EXAMPLE NO 1:** The signal in the following figure is to be sampled. Determine the minimum sample rate without any aliasing effect. If the signal is sampled at a rate 8 kHz, determine the cut off frequency of the anti-aliasing filter.[JUIY-2011, 6M]

**Magnitude B**

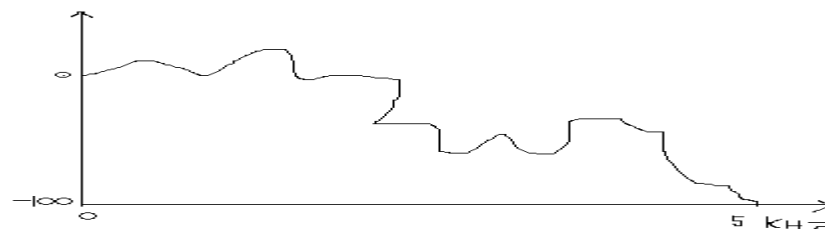


FIG:1

Freq

**SOLUTION:**

i) Sampling theorem :  $f_s = 2f_{max}$

Given:-  $f_{max} = 5k \text{ Hz}$

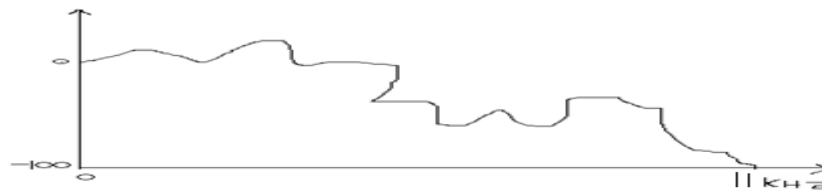
$$f_s = 2f_{\max} = 2(5 \times 10^3) \text{ Hz} = 10 \text{ kHz.}$$

ii) If  $f_s = 8 \text{ kHz}$  then  $f_{\max} = 4 \text{ kHz}$ .

Low pass filter cutoff frequency = 4 kHz.

**EXAMPLE NO 2:** The signal in the following figure is to be sampled. Determine the minimum sample rate without any aliasing effect. If the signal is sampled at a rate 16 kHz, determine the cut off frequency of the anti-aliasing filter.

**SOLUTION:**



i) Sampling theorem :  $f_s = 2f_{\max}$

Given:-  $f_{\max} = 11 \text{ kHz}$

$$f_s = 2f_{\max} = 2(11 \times 10^3) \text{ Hz} = 22 \text{ kHz.}$$

ii)  $f_s = 16 \text{ kHz}$  then  $f_{\max} = 8 \text{ kHz}$ .

Low pass filter cutoff frequency = 8 kHz.

## 1.6 Discrete Time Sequences

Consider an analog signal  $x(t)$  given by ,  $x(t) = A \cos (2\pi ft)$

if a signal is sampled at a rate of Sampling Interval  $T$ , in the above equation replacing  $t$  by  $nT$  we get,

$$x(nT) = A \cos (2\pi fnT)$$

where  $n = 0, 1, 2, \dots$  etc

For simplicity denote  $x(nT)$  as  $x(n)$

$$x(n) = A \cos (2\pi fnT)$$

where  $n = 0, 1, 2, \dots$  etc

We have  $f_s = 1/T$  also  $\theta = 2\pi fnT$

$$x(n) = A \cos(2\pi fnT) = A \cos (2\pi fn/f_s) = A \cos \theta n$$

$\theta$  is called as digital frequency.

$$\theta = 2\pi fT = 2\pi f/f_s \text{ radians}$$

$$\theta = 2\pi fT = 2\pi f/f_s \text{ radians}$$

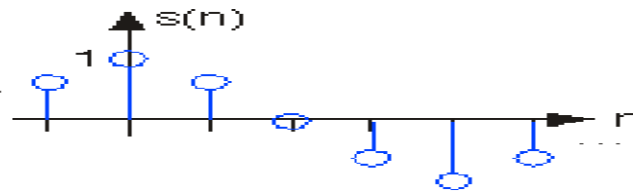


Fig 1.5 A Cosine Waveform

A sequence that repeats itself after every period  $N$  is called a periodic sequence. Consider a periodic sequence  $x(n)$  with period  $N$   $x(n) = x(n+N)$   $n = \dots, -1, 0, 1, 2, \dots$ . Frequency response gives the frequency domain equivalent of a discrete time sequence. It is denoted as

$$X(e^{j\theta}) = \sum x(n) e^{-jn\theta}$$

Where  $\theta$  is the digital frequency ranging from  $0$  to  $2\pi$ , Frequency response of a discrete sequence involves both magnitude response and phase response.

## 1.7 Discrete Fourier Transform and Fast Fourier Transform

- DFT is used to transform a time domain  $x(n)$  sequence to frequency domain  $X(K)$  sequence.
- IDFT is used to transform  $X(K)$  to  $x(n)$ .
- Algorithms for the fast computation of DFT and IDFT are known as FFT algorithms.

### 1.7.1 DFT Pair:

DFT is used to transform a time domain sequence  $x(n)$  to a frequency domain sequence  $X(K)$ . The equations that relate the time domain sequence  $x(n)$  and the corresponding frequency domain sequence  $X(K)$  are called DFT Pair and is given by,

*DFT(FFT):*

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-j\left(\frac{2\pi}{N}\right)nk} \quad (k = 0, 1, \dots, N-1)$$

*IDFT(IFFT):*

$$x(n) = \frac{1}{N} \sum_{k=0}^{N-1} X(k) \cdot e^{j\left(\frac{2\pi}{N}\right)nk} \quad (n = 0, 1, \dots, N-1)$$

Where  $N$  denotes the number of elements in the  $x(n)$  or  $X(K)$ .

### 1.7.2 THE RELATIONSHIP BETWEEN DFT AND FREQUENCY RESPONSE:

We have,

$$X(e^{j\theta}) = \sum x(n) e^{-jn\theta}$$

Also

$$X(K) = \sum x(n) e^{-j2\pi n k/N}$$

$$\therefore X(K) = X(e^{j\theta}) \text{ at } \theta = 2\pi k/N$$

- From the above expression it is clear that we can use DFT to find the Frequency response of a discrete signal. Spacing between the element of a  $X(K)$  is given as the

$$\Delta f = f_s/N = 1/NT = 1/T_0.$$

Where  $T_0$  is the signal record length.

- It is clear from the expression of  $f$  that, in order to minimize the spacing between the samples  $N$  has to be a large value.
- Although DFT is an efficient technique of obtaining the frequency response of a sequence, it requires more number of complex operations like additions and multiplications.

**EXAMPLE NO 1: An FFT is employed for determining the frequency components of a random signal. It is required that the resolution of FFT to be  $\leq 5$  Hz, for a signal with  $f_{\max} = 1.25$  kHz. Determine**

**i) Sampling interval ii) FFT length( $N$ ) as a power of 2 iii) Minimum signal record length.**

**SOLUTION:  $\Delta f = 5$  Hz,  $f_{\max} = 1.25$  KHz,  $f_s = 2 f_{\max} = 2.5$  kHz**

**i)  $T_s = 1/f_s = 0.4$  msec**

**w.r.t  $\rightarrow \Delta f = f_s/N; \quad N = \Delta f / f_s$**

$$N = (2.5 \times 10^3) / 5 = 500$$

**ii) FFT length  $N$  as the power of 2  $N = 2^{500}$  i.e  $N = 16384$**

**iii) Signal recorded length  $T_0 = NT_s = (16384)(0.4 \text{ m sec}) = 6.5536$  sec**

**EXAMPLES NO 2: The signal is filtered and sampled using the sampling rate of 8 kHz. If 512 samples of this signal are used to compute the Fourier transform  $X(K)$ , determine the frequency spacing between adjacent  $X(K)$  elements. What is the analog frequency corresponding to  $k=64, 128$ , and  $200$ . Repeat this problem using 1024 samples and an 8 kHz sampling rate.**

**Solution:**

**i) The digital frequency spacing between adjacent  $X(K)$  elements is given by  $2\pi/N$**





radians

$$\text{i.e } 2\pi/N = 2\pi/512 = 0.01227 \text{ radians}$$

The analog frequency is given by  $\Delta f = f_s / N = (8 \times 1000) / 512 = 15.625$

The analog frequency for  $k=64$ ,  $\Delta f = f_s k / N = (8 \times 1000) 64 / 512 = 1000$

The analog frequency for  $k=128$ ,  $\Delta f = f_s k / N = (8 \times 1000) 128 / 512 = 2000$

The analog frequency for  $k=200$ ,  $\Delta f = f_s k / N = (8 \times 1000) 200 / 512 = 3125$

iii) for  $N=1024$  and  $f_s = 8 \text{ kHz}$

The digital frequency spacing between adjacent  $X(K)$  elements is given by  $2\pi/N$  radians

$$\text{i.e } 2\pi/N = 2\pi/1024 = 0.0061359$$

The analog frequency is given by  $\Delta f = f_s / N = (8 \times 1000) / 1024 = 7.8$

The analog frequency for  $k=64$ ,  $\Delta f = f_s k / N = (8 \times 1000) 64 / 1024 = 500$

The analog frequency for  $k=128$ ,  $\Delta f = f_s k / N = (8 \times 1000) 128 / 1024 = 1000$

The analog frequency for  $k=200$ ,  $\Delta f = f_s k / N = (8 \times 1000) 200 / 1024 = 1562.5$

### 1.7.3 FAST FOURIER TRANSFORM (FFT)

- The direct computation of DFT and IDFT requires a larger number of complex multiplies.
- The FFT algorithms use power of 2 point and exploits the periodic nature of the complex exponential  $e^{j2\pi nk/N}$  occurring in DFT and IDFT equations.
- FFT algorithms called the radix 2 such as 2,4,8,16,.....
- The DFT requires  $N^2$  complex multipliers and  $N(N-1)$  complex additions
- The radix-2 algorithm requires  $N/2 \log_2 N$ .
- Thus many improvements over DFT were proposed. One such technique is to use the periodicity property of the twiddle factor  $e^{j2\pi / N}$ . Those algorithms were called as Fast Fourier Transform Algorithms.
- The following table depicts the complexity involved in the computation using DFT algorithms.

Table 1.1 Complexity in DFT algorithm

Operations	Number of Computations
Complex Multiplications	$N^2$
Complex Additions	$N(N-1)$
Real Multiplications	$4N^2$
Real Additions	$2N(2N-1)$
Trigonometric Functions	$2N^2$

FFT algorithms are classified into two categories via

1. Decimation in Time FFT
2. Decimation in Frequency FFT

In decimation in time FFT the sequence is divided in time domain successively till we reach the sequences of length 2. Whereas in Decimation in Frequency FFT, the sequence  $X(K)$  is divided successively. The complexity of computation will get reduced considerably in case of FFT algorithms.

## 1.8 LINEAR TIME INVARIANT SYSTEMS

- A system which satisfies superposition theorem is called as a linear system
- a system that has same input output relation at all times is called a Time Invariant System
- system which satisfy both the properties, are called LTI systems.

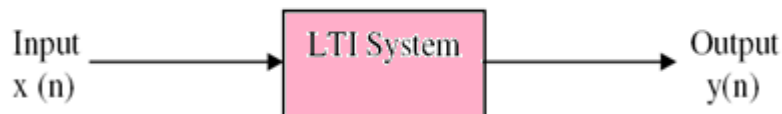


Fig 1.6 An LTI system

- LTI systems are characterized by its impulse response or unit sample response in time domain, time domain convolution can be used to determine the response of an LTI system.

- In frequency domain the system Trans function is used to represent LTI systems.

### 1.8.1 Convolution

- Convolution is the operation that related the input output of an LTI system, to its unit sample response.
- The output of the system  $y(n)$  for the input  $x(n)$  and the impulse response of the system being  $h(n)$  is given as

$$y(n) = x(n) * h(n) = \sum x(k)h(n-k)$$

$$y(n) = x(n) * h(n) = \sum x(k)h(n-k)$$

- $x(n)$  is the input of the system,  $h(n)$  is the impulse response of the system,  $y(n)$  is the output of the system.

### 1.8.2 Z Transformation

Z Transformations are used to find the frequency response of the system. The Z Transform for a discrete sequence  $x(n)$  is given by,

$$X(Z) = \sum x(n) z^{-n}$$

$$X(Z) = \sum x(n) z^{-n}$$

where  $X(Z)$  is called the Z transform of  $x(n)$

### 1.8.3 The System Function

- An LTI system is characterized by its System function or the transfer function.
- The system function of a system is the ratio of the Z transformation of its output to that of its input. It is denoted as  $H(Z)$  and is given by

$$H(Z) = Y(Z) / X(Z).$$

- The magnitude and phase of the transfer function  $H(Z)$  gives the frequency response of the system.
- From the transfer function we can also get the poles and zeros of the system by solving its numerator and denominator respectively.

## 1.9 Digital Filters

Filters are used to remove the unwanted components in the sequence. They are characterized by the impulse response  $h(n)$ . The general difference equation for an Nth order filter is given by

$$y(n) = \sum_{k=1}^N a_k y(n-k) + \sum_{k=0}^L b_k x(n-k)$$

$$y(n) = \sum_{k=1}^N a_k y(n-k) + \sum_{k=0}^L b_k x(n-k)$$



A typical digital filter structure is as shown in figure 1.7.

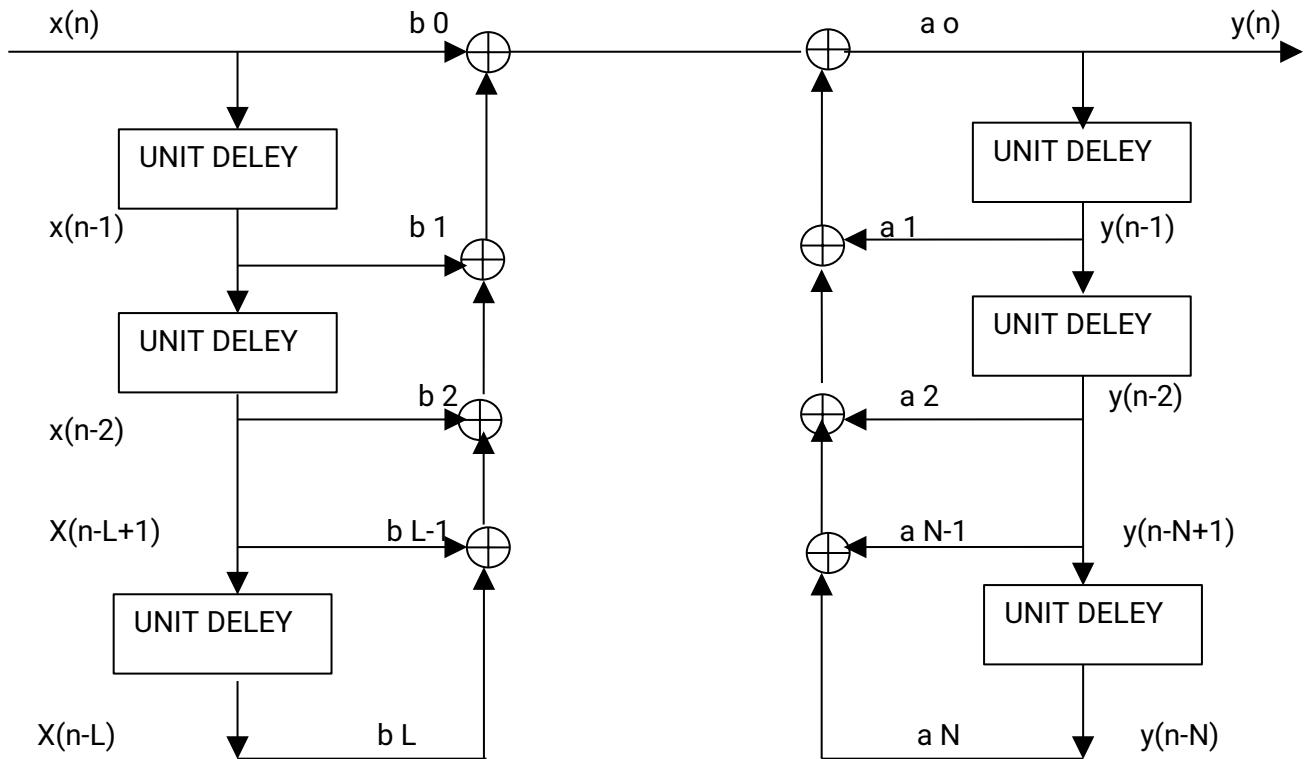


Fig 1.7 Structure of a Digital Filter

Values of the filter coefficients vary with respect to the type of the filter. Design of a digital filter involves determining the filter coefficients. Based on the length of the impulse response, digital filters are classified into two categories.

1. Finite Impulse Response (FIR)
2. Filters and Infinite Impulse Response (IIR) Filters.

### 1.9.1 FIR Filters

- FIR filters have impulse responses of finite lengths.
- In FIR filters the present output depends only on the past and present values of the input sequence but not on the previous output sequences. Thus they are non recursive hence they are inherently stable.
- FIR filters possess linear phase response. Hence they are very much applicable for the applications requiring linear phase response.
- The difference equation of an FIR filter is represented as

$$y(n) = \sum b_k x(n-k)$$

The frequency response of an FIR filter is given as

$$H(Z) = \sum b_k Z^{-k}$$

- The major drawback of FIR filters is, they require more number of filter coefficients to realize a desired response as compared to IIR filters. Thus the computational time required will also be more.
- FIR filter has no feedback hence it is stable.
- A symmetric coefficient FIR filter provides linear phase or constant group delay.

**EXAMPLE NO 1:**

A FIR Filter the equation is  $y(n) = 0.5 x(n) + 0.5 x(n-1)$  describe

- system function
- magnitude response function
- phase response function
- group delay

Plot its magnitude and phase response of a simple FIR whose output is the average of the current input  $x(n)$  and the past input  $x(n-1)$ .

**SOLUTION:**

The unit sample response of this filter is obtained by substituting  $\delta(n)$  for  $x(n)$ . thus, we have

$$\begin{aligned} h(n) &= 0.5 \delta(n) + 0.5 \delta(n-1) \\ &= [0.5 \quad 0.5] \text{ as a sequence} \end{aligned}$$

The frequency response, is obtained as

$$H(e^{j\theta}) = 0.5 + 0.5 e^{-j\theta} = e^{-j\theta/2} \cos\theta/2$$

or

$$H(Z) = 0.5 + 0.5 Z^{-1}$$

The magnitude response is given as

$$|H(e^{j\theta})| = M(\theta) = \cos\theta/2$$

The phase response is given as

$$\angle H(e^{j\theta}) = P(\theta) = -\theta/2$$

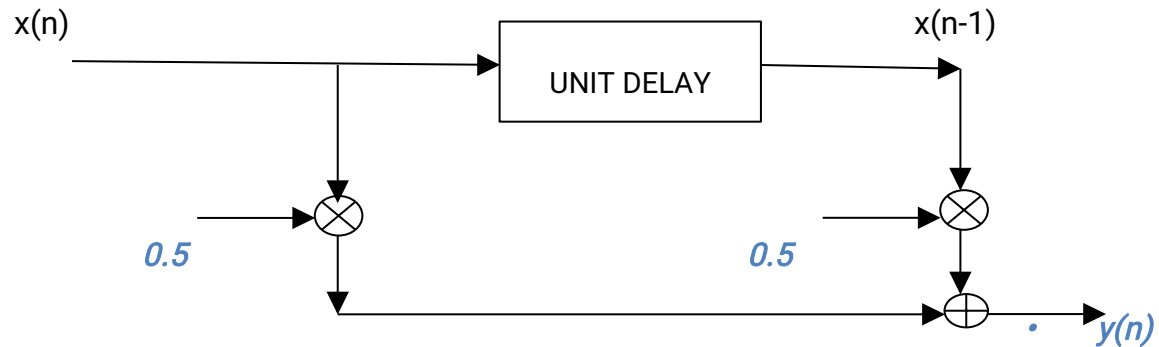
The group delay which represents the delay to various signal frequencies can be obtained by differentiating and neglecting the phase response function.

Group delay =  $\frac{1}{2}$ .

Group delay = 1/2

$$H(e^{j\theta}) = \sum b_k e^{-jk\theta}$$

BLOCK DIAGRAM

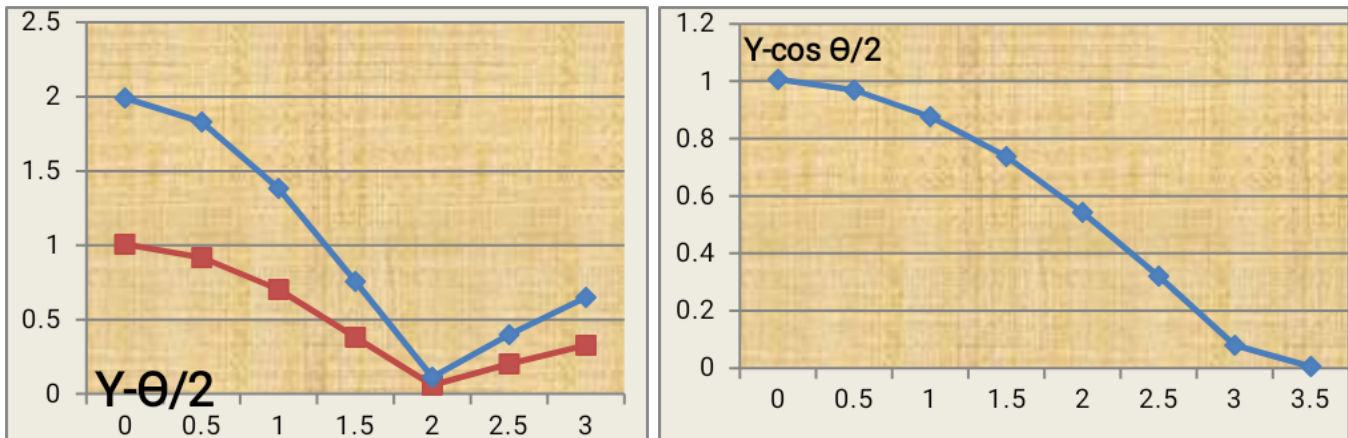


MAGNITUDE RESPONSE

$\theta$	$\cos \theta/2$
0	1
0.5	.96
1	.87
1.5	.731
2	.540
2.5	.315
3	.0707

PHASE RESPONSE

$\theta$	$-\theta/2$
0	0
0.5	-.25
1	-.50
1.5	-.75
2	-1
2.5	-1.25
3	-1.5



EXAMPLE NO 2: Implement the FIR Filter for  $y(n) = (x(n) + x(n-1) + x(n-2))/3$

Describe the

- i) system function
- ii) magnitude response function
- iii) phase response function
- iv) step response
- v) group delay

Plot its magnitude and phase response.

SOLUTION:

- i)  $y(n) = (x(n) + x(n-1) + x(n-2))/3$   
 $Y(z) = 1/3(1 + Z^{-1} + Z^{-2})X(z)$   
 $H(z) = Y(z)/X(z) = 0.33[1 + Z^{-1} + Z^{-2}]$
- ii) Magnitude response function

$$\begin{aligned} H(e^{j\theta}) &= 1/3[1 + e^{-j\theta} + e^{-j2\theta}] \\ &= e^{-j\theta}/3[e^{j\theta} + 1 + e^{-j\theta}] \\ &= e^{-j\theta}/3[\cos\theta + j\sin\theta + \cos\theta - j\sin\theta + 1] \\ &= e^{-j\theta}/3[1 + 2\cos\theta] \end{aligned}$$

Magnitude response

$$M(\theta) = |H(e^{j\theta})| = |e^{-j\theta}/3| + 2/3(|e^{-j\theta} \cos\theta|)$$

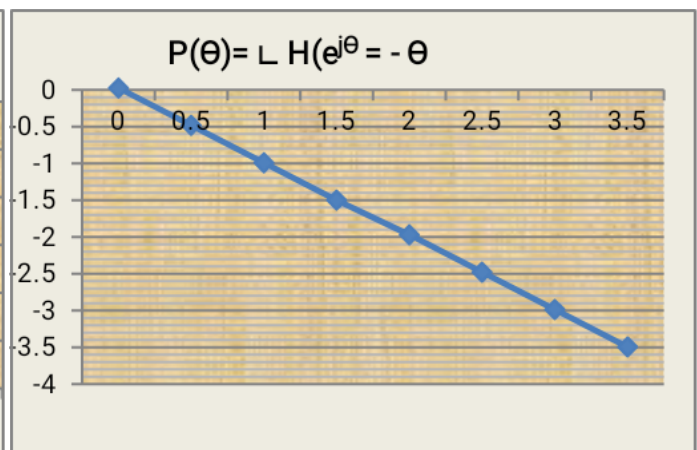
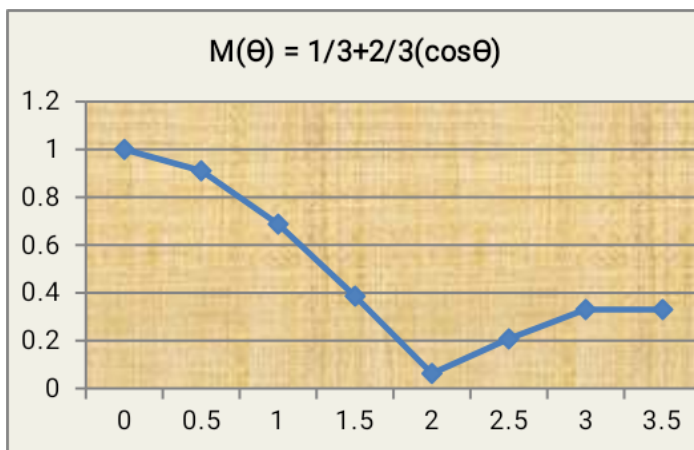
$$M(\theta) = 1/3 + 2/3(\cos\theta)$$

- iii) Phase response

$$P(\theta) = \angle H(e^{j\theta}) = -\theta$$

PLOT OF MAGNITUDE AND PHASE RESPONSE

$\theta$	$M(\theta) =  H(e^{j\theta})  = \frac{1}{3} + \frac{2}{3}(\cos\theta)$	$\theta$	$P(\theta) = \angle H(e^{j\theta}) = -\theta$
0	0.99	0	0
.5	.909	.5	-0.5
1	.686	1	-1
1.5	.376	1.5	-1.5
2	.055	2	-2
2.5	.198	2.5	-2.5
3	.323	3	-3
3.5	.329	3.5	-3.5



iv)  $Y(n) = \frac{1}{3}[x(n) + x(n-1) + x(n-2)]$

We know that  $x(n] = u(n)X(Z) = \frac{Z}{(Z-1)}$

$$Y(Z) = \frac{1}{3}[X(Z) + Z^{-1}X(Z-1) + Z^{-2}X(Z-2)]$$

$$Y(Z) = \frac{X(Z)}{3[1 + Z^{-1} + Z^{-2}]}$$

$$= \frac{1}{3}[1 + Z^{-1} + Z^{-2}][\frac{Z}{(Z-1)}]$$

$$= \frac{1}{3}[(Z + 1 + Z^{-1})/(Z-1)]$$

$$= \frac{1}{3}[(Z^2 + Z + 1)/(Z(Z-1))]$$



$$Y(Z)/Z = 1/3[(Z^2+Z+1)/(Z^2(Z-1))]$$

Applying partial fraction method

$$Y(Z)/Z = 1/3[A/Z + B/Z^2 + C/(Z-1)]$$

$$Z^2 + Z + 1 = A(Z)(Z-1) + B(Z-1) + CZ^2$$

$$\text{When } Z=0 \rightarrow 1=B(-1) \text{ i.e } B=-1.$$

$$\text{When } Z=1 \rightarrow 3=C(1) \text{ i.e } C=3.$$

$$\text{When } Z=-1 \rightarrow 1-1+1=A(-1)(-2) + B(-2) + C$$

$$1=2A -2B +C$$

$$1=2A -2(-1)+ 3$$

$$2A = -4 \text{ i.e } A= -2.$$

$$Y(Z)/Z = 1/3[A/Z + B/Z^2 + C/(Z-1)]$$

Putting the values of A= -2, B=-1, C=3.

$$Y(Z)/Z = 1/3[-2/Z - 1/Z^2 + 3/(Z-1)]$$

$$Y(Z) = 1/3[(-2) - (1/Z) + (3Z)/(Z-1)]$$

Taking inverse Z transform we get

$$Y(n) = -2/3(\delta(n)) - 1/3(\delta(n)) + u(n)$$

$$Y(n) = -2/3(\delta(n)) - 1/3(\delta(n)) + u(n)$$

v) Group delay

$$\angle H(e^{j\theta}) = P(\theta) = -\theta$$

$$dp/d\theta = -1$$

$$\text{Group delay} = 1$$

### 1.9.2 IIR Filters

Unlike FIR filters, IIR filters have infinite number of impulse response samples. They are recursive filters as the output depends not only on the past and present inputs but also on the past outputs. They generally do not have linear phase characteristics. Typical system function of such filters is given by,

$$H(Z) = (b_0 + b_1 Z^{-1} + b_2 Z^{-2} + \dots + b_L Z^{-L}) / (1 - a_1 Z^{-1} - a_2 Z^{-2} - \dots - a_N Z^{-N})$$

Stability of IIR filters depends on the number and the values of the filter coefficients. The major advantage of IIR filters over FIR is that, they require lesser coefficients compared to FIR filters for the same desired response, thus requiring less computation time.

**EXAMPLE NO 1:**

Obtain the transfer function of the IIR filter whose difference equation is given by  $y(n) = 0.9y(n-1) + 0.1x(n)$

**SOLUTION:**

$$y(n) = 0.9y(n-1) + 0.1x(n)$$

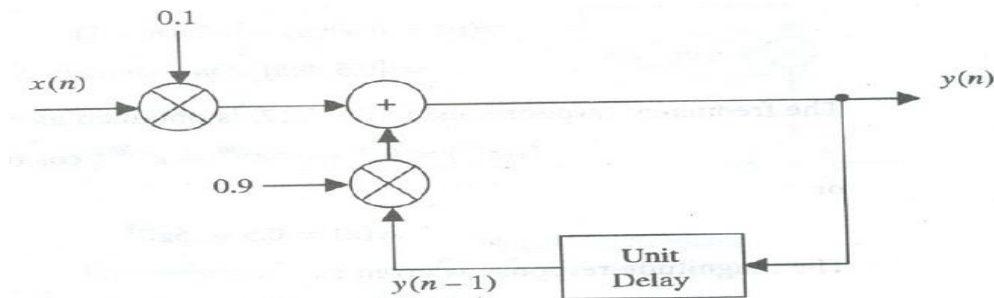
Taking Z transformation both sides  $Y(Z) = 0.9 Z^{-1} Y(Z) + 0.1 X(Z)$

$$Y(Z) [1 - 0.9 Z^{-1}] = 0.1 X(Z)$$

The transfer function of the system is given by the expression,

$$H(Z) = Y(Z)/X(Z) = 0.1 / [1 - 0.9 Z^{-1}]$$

Realization of the IIR filter with the above difference equation is as shown in figure.



**1.9.3**

**FIR Filter Design**

Frequency response of an FIR filter is given by the following expression,

$$H(e^{j\theta}) = \sum b_k e^{-jk\theta}$$

Design procedure of an FIR filter involves the determination of the filter coefficients  $b_k$ .

$$b_k = (1/2\pi) \int H(e^{j\theta}) e^{-jk\theta} d\theta$$

**1.9.4**

**IIR**

## Filter Design

IIR filters can be designed using two methods viz using windows and direct method. In this approach, a digital filter can be designed based on its equivalent analog filter. An analog filter is designed first for the equivalent analog specifications for the given digital specifications. Then using appropriate frequency transformations, a digital filter can be obtained. The filter specifications consist of passband and stopband ripples in dB and Passband and Stopband frequencies in rad/sec.

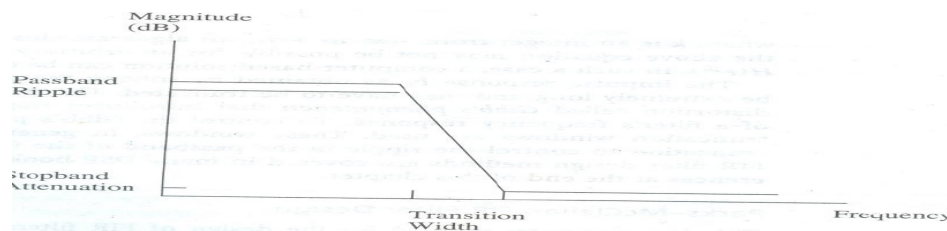


Fig 1.11 Lowpass Filter Specifications

Direct IIR filter design methods are based on least squares fit to a desired frequency response. These methods allow arbitrary frequency response specifications.

## 2.1 Decimation and Interpolation

Decimation and Interpolation are two techniques used to alter the sampling rate of a sequence. Decimation involves decreasing the sampling rate without violating the sampling theorem whereas interpolation increases the sampling rate of a sequence appropriately by considering its neighboring samples.

### 2.1.1 Decimation

Decimation is a process of dropping the samples without violating sampling theorem. The factor by which the signal is decimated is called as decimation factor and it is denoted by M. It is given by,

$$y(m) = w(mM) = \sum b_k x(mM-k) \quad \text{where } w(n) = \sum b_k x(n-k)$$

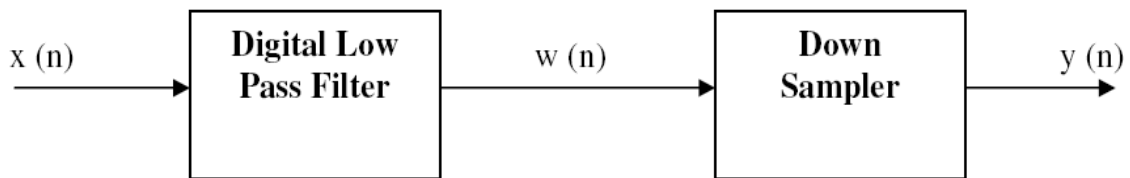


Fig 1.12 Decimation Process

### 2.1.2 Interpolation

Interpolation is a process of increasing the sampling rate by inserting new samples in between. The input output relation for the interpolation, where the sampling rate is increased by a factor  $L$ , is given as,

$$y(m) = \sum b_k w(m-k)$$

$$\text{where } w(n) = \begin{cases} x(m/L), & m=0, \pm L, \pm 2L, \dots \\ 0 & \text{Otherwise} \end{cases}$$

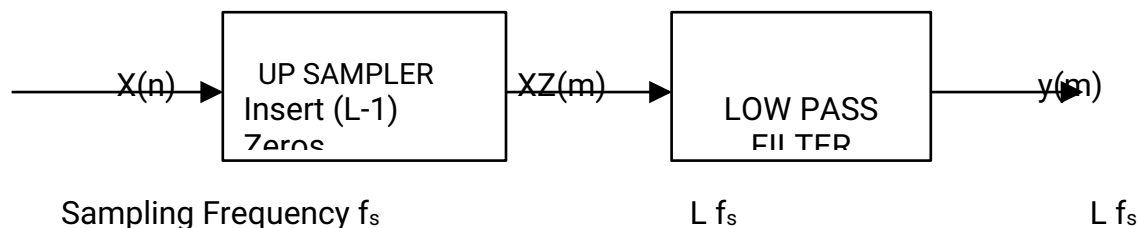


Fig 1.13 Interpolation Process

**PROBLEMS 1:** For a sequence  $x(n) = \{3, 2, 2, 4, 1, 0, -3, -2, -1, 0, 2, 3\}$  need to be decimated by a factor of 2. find the decimated sequence  $w(n)$ . after filtering with an appropriate LPF the sampling theorem let the sequence be  $w(n) = \{2.1, 2, 3.9, 1.5, 0.1, -2.9, -2, -1.1, 0.1, 1.9, 2.9\}$ .

**Solution:** The decimated sequence is obtained by dropping  $M-1$  every other sample. This gives the decimated sequence as  $M=2$ ,  $M-1=2-1=1$  sample to be dropped.

$$Y(m) = \{2, 2, 3, 1.5, 0, -2.9, -2, -1.1, 0.1, 1.9, 2.9\}$$

$$Y(m) = \{2, 1.5, -2.9, -1.1, 1.9, \}$$

**PROBLEM NO 2:** For a sequence  $x(n) = \{0, 2, 4, 6, 8, 10, 12, 14, 16\}$  need to be decimated by a factor of 2. find the decimated sequence  $w(n)$ . after filtering with an appropriate LPF

the sampling theorem let the sequence be  $w(n) = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16\}$ .

Solution: The decimated sequence is obtained by dropping  $M-1$  every other sample. This gives the decimated sequence as  $M=2$ ,  $M-1=2-1=1$  sample to be dropped.

$w(n) = \{0, 1, \cancel{2}, 3, \cancel{4}, 5, \cancel{6}, 7, \cancel{8}, 9, \cancel{10}, 11, \cancel{12}, 13, \cancel{14}, 15, \cancel{16}\}$ .

$Y(m) = \{ \underset{\uparrow}{1}, 3, 5, 7, 9, 11, 13, 15 \}$

**PROBLEM NO 3:** The sequence  $x(m) = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$  is decimated using decimation sequence  $b_k = (1, 1.5, 1)$  and a decimation factor is 2. Find the decimation sequence  $y(m)$ .

Solution: given  $x(m) = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$

$b_k = (1, \underset{\uparrow}{1.5}, 1)$   $k=3$  and varying from -1 to 1

$$w(m) = \sum_{k=-1}^{k=1} b_k x(m-k)$$

$w(m) = \sum b_k x(m-k) = b_{-1} w(m+1) + b_0 w(m) + b_1 w(m-1)$   
putting the values of  $n$  varying from 0 to 8.

We get

$$\begin{aligned} w(0) &= b_{-1} w(1) + b_0 w(0) + b_1 w(-1) = 1 \\ w(1) &= b_{-1} w(2) + b_0 w(1) + b_1 w(0) = 3.5 \\ w(2) &= b_{-1} w(3) + b_0 w(2) + b_1 w(1) = 7 \\ w(3) &= b_{-1} w(4) + b_0 w(3) + b_1 w(2) = 10.5 \\ w(4) &= b_{-1} w(5) + b_0 w(4) + b_1 w(3) = 14 \\ w(5) &= b_{-1} w(6) + b_0 w(5) + b_1 w(4) = 17.5 \\ w(6) &= b_{-1} w(7) + b_0 w(6) + b_1 w(5) = 21 \\ w(7) &= b_{-1} w(8) + b_0 w(7) + b_1 w(6) = 24.5 \\ w(8) &= b_{-1} w(9) + b_0 w(8) + b_1 w(7) = 19 \end{aligned}$$

$$w(m) = (1, 3.5, 7, 10.5, 14, 17.5, 21, 24.5, 19)$$

given  $L=2$  that means  $L-1=2-1=1$  elements to be deleted

$$w(m) = (\cancel{1}, 3.5, \cancel{7}, 10.5, \cancel{14}, 17.5, \cancel{21}, 24.5, \cancel{19})$$

$$y(m) = (3.5, 10.5, 17.5, 14.5, 24.5)$$



**ALTERNATE METHOD:**

$$x(m) = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$$

$$B_k = (1, 1.5, 1) \quad k=3 \text{ and varying from } -1 \text{ to } 1$$

$$w(m) = \sum_{k=-1}^{k=1} b_k x(m-k)$$

X(m)			0	1	2	3	4	5	6	7	8
B <sub>k</sub>								1	1.5	1	×
			0	1	2	3	4	5	6	7	8
		0	1.5	3	4.5	6	7.5	9	10.5	12	+
	0	1	2	3	4	5	6	7	8	+	+
W(n)	0	1	3.5	7	10.5	14	17.5	21	24.5	19	8

Num before 0<sup>th</sup> position deleted      num beyond range deleted

$$w(m) = (1, 3.5, 7, 10.5, 14, 17.5, 21, 24.5, 19)$$

given L=2 that means L-1=2-1=1 elements to be deleted

$$w(m) = (1, 3.5, 7, 10.5, 14, 17.5, 21, 24.5, 19)$$

$$y(m) = (3.5, 10.5, 17.5, 14.5, 24.5)$$

**PROBLEM NO:4** The sequence  $x(m) = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$  is decimated using decimation sequence  $b_k = (1/3, 2/3, 1, 2/3, 1/3)$  and a decimation factor is 3. Find the decimation sequence  $y(m)$ .

**SOLUTION:** given  $x(m) = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

$$b_k = (1/3, 2/3, 1, 2/3, 1/3) \quad k=5 \text{ and varying from } 0 \text{ to } 4$$

$$w(m) = \sum_{k=0}^{k=4} b_k x(m-k)$$

$$w(m) = \sum b_k x(m-k) = b_0 w(m) + b_1 w(m-1) + b_2 w(m-2) + b_3 w(m-3) + b_4 w(m-4)$$



putting the values of m varying from 0 to 12. We get

$$\begin{aligned}
 w(0) &= b_0 x(m-0) + b_1 x(m-1) + b_2 x(m-2) + b_3 x(m-3) + b_4 x(m-4) = 0 \\
 w(1) &= b_1 x(m-1) + b_2 x(m-2) + b_3 x(m-3) + b_4 x(m-4) + b_5 x(m-5) = 1/3 \\
 w(2) &= b_2 x(m-2) + b_3 x(m-3) + b_4 x(m-4) + b_5 x(m-5) + b_6 x(m-6) = 4/3 \\
 w(3) &= b_3 x(m-3) + b_4 x(m-4) + b_5 x(m-5) + b_6 x(m-6) + b_7 x(m-7) = 10/3 \\
 w(4) &= b_4 x(m-4) + b_5 x(m-5) + b_6 x(m-6) + b_7 x(m-7) + b_8 x(m-8) = 6 \\
 w(5) &= b_5 x(m-5) + b_6 x(m-6) + b_7 x(m-7) + b_8 x(m-8) + b_9 x(m-9) = 9 \\
 w(6) &= b_6 x(m-6) + b_7 x(m-7) + b_8 x(m-8) + b_9 x(m-9) + b_{10} x(m-10) = 12 \\
 w(7) &= b_7 x(m-7) + b_8 x(m-8) + b_9 x(m-9) + b_{10} x(m-10) + b_{11} x(m-11) = 15 \\
 w(8) &= b_8 x(m-8) + b_9 x(m-9) + b_{10} x(m-10) + b_{11} x(m-11) + b_{12} x(m-12) = 18 \\
 w(9) &= b_9 x(m-9) + b_{10} x(m-10) + b_{11} x(m-11) + b_{12} x(m-12) + b_{13} x(m-13) = 21 \\
 w(10) &= b_{10} x(m-10) + b_{11} x(m-11) + b_{12} x(m-12) + b_{13} x(m-13) + b_{14} x(m-14) = 24 \\
 w(11) &= b_{11} x(m-11) + b_{12} x(m-12) + b_{13} x(m-13) + b_{14} x(m-14) + b_{15} x(m-15) = 27 \\
 w(12) &= b_{12} x(m-12) + b_{13} x(m-13) + b_{14} x(m-14) + b_{15} x(m-15) + b_{16} x(m-16) = 30
 \end{aligned}$$

$$w(m) = (0, 1/3, 4/3, 10/3, 6, 9, 12, 15, 18, 21, 24, 27, 30)$$

given  $L=3$  that means  $L-1=3-1=2$  elements to be deleted

$$w(m) = (\cancel{0}, \cancel{1/3}, 4/3, \cancel{10/3}, \cancel{6}, 9, \cancel{12}, \cancel{15}, 18, \cancel{21}, \cancel{24}, 27, \cancel{30})$$

$$y(m) = (4/3, 9, 18, 27)$$

### ALTERNATE METHOD:

given  $x(m) = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

$b_k = (1/3, 2/3, 1, 2/3, 1/3)$   $k=5$  and varying from 0 to 4

$$w(m) = \sum_{k=0}^{k=4} b_k x(m-k)$$

$x(m) =$					0	1	2	3	4	5	6	7	8	9	10	11	12
$b_k =$												$1/3$	$2/3$	1	$2/3$	$1/3$	$\times$
					0	$1/3$	$2/3$	1	$4/3$	$5/3$	2	$7/3$	$8/3$	3	$10/3$	$11/3$	4
			0	$2/3$	$4/3$	2	$8/3$	$10/3$	4	$14/3$	$16/3$	6	$20/3$	$22/3$	8	+	+
		0	1	2	3	4	5	6	7	8	9	10	11	12	+	+	+
	0	$2/3$	$4/3$	2	$8/3$	$10/3$	4	$14/3$	$16/3$	6	$20/3$	$22/3$	8	+	+	+	+
	0	$1/3$	$2/3$	1	$4/3$	$5/3$	2	$7/3$	$8/3$	3	$10/3$	$11/3$	4	+	+	+	+
$Y(m)$	0	$1/3$	$4/3$	$10/3$	6	9	12	15	18	21	24	27	30	$86/3$	$68/3$	$35/3$	4

(elements beyond the range are deleted)

$$w(m) = (0, 1/3, 4/3, 10/3, 6, 9, 12, 15, 18, 21, 24, 27, 30)$$

given  $L=3$  that means  $L-1=3-1=2$  elements to be deleted

$$w(m) = (\cancel{0}, \cancel{1/3}, 4/3, \cancel{10/3}, \cancel{6}, 9, \cancel{12}, \cancel{15}, 18, \cancel{21}, \cancel{24}, \cancel{27}, \cancel{30})$$

$$y(m) = (4/3, 9, 18, 27)$$

### PROBLEM NO: 5

Let  $x(n) = [0 \ 3 \ 6 \ 9 \ 12]$  be interpolated with  $L=3$ . If the filter coefficients of the filters are  $b_k = [1/3 \ 2/3 \ 1 \ 2/3 \ 1/3]$ , obtain the interpolated sequence.

SOLUTION:

$L-1$  Zeros must be inserted i.e.  $L-1=3-1=2$  zeros must be inserted after every element of  $x(n)$ .

After inserting zeros,

$$w(m) = [0 \ 0 \ 0 \ 3 \ 0 \ 0 \ 6 \ 0 \ 0 \ 9 \ 0 \ 0 \ 12]$$

$$b_k = [1/3 \ 2/3 \ 1 \ 2/3 \ 1/3] \text{ We have,}$$

$k$  values varies from -2 to +2





substituting in the equation below we get

$$y(m) = \sum_{k=-2}^{k=2} b_k w(m-k)$$

$$y(m) = \sum b_k w(m-k) = b_{-2} w(m+2) + b_{-1} w(m+1) + b_0 w(m) + b_1 w(m-1) + b_2 w(m-2)$$

Substituting the values of m, as m varies from 0 to 12, we get

$$\begin{aligned} y(0) &= b_{-2} w(2) + b_{-1} w(1) + b_0 w(0) + b_1 w(-1) + b_2 w(-2) = 0 \\ y(1) &= b_{-2} w(3) + b_{-1} w(2) + b_0 w(1) + b_1 w(0) + b_2 w(-1) = 1 \\ y(2) &= b_{-2} w(4) + b_{-1} w(3) + b_0 w(2) + b_1 w(1) + b_2 w(0) = 2 \\ y(3) &= b_{-2} w(5) + b_{-1} w(4) + b_0 w(3) + b_1 w(2) + b_2 w(1) = 3 \\ y(4) &= b_{-2} w(6) + b_{-1} w(5) + b_0 w(4) + b_1 w(3) + b_2 w(2) = 4 \\ y(5) &= b_{-2} w(7) + b_{-1} w(6) + b_0 w(5) + b_1 w(4) + b_2 w(3) = 5 \\ y(6) &= b_{-2} w(8) + b_{-1} w(7) + b_0 w(6) + b_1 w(5) + b_2 w(4) = 6 \\ y(7) &= b_{-2} w(9) + b_{-1} w(8) + b_0 w(7) + b_1 w(6) + b_2 w(5) = 7 \\ y(8) &= b_{-2} w(10) + b_{-1} w(9) + b_0 w(8) + b_1 w(7) + b_2 w(6) = 8 \\ y(9) &= b_{-2} w(11) + b_{-1} w(10) + b_0 w(9) + b_1 w(8) + b_2 w(7) = 9 \\ y(10) &= b_{-2} w(12) + b_{-1} w(11) + b_0 w(10) + b_1 w(9) + b_2 w(8) = 10 \\ y(11) &= b_{-2} w(13) + b_{-1} w(12) + b_0 w(13) + b_1 w(12) + b_2 w(11) = 11 \\ y(12) &= b_{-2} w(14) + b_{-1} w(13) + b_0 w(12) + b_1 w(11) + b_2 w(10) = 12 \end{aligned}$$

Similarly we get the remaining samples as,  $y(n) = [0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ 10 \ 11 \ 12]$

### ALTERNATE METHOD (TABULATION METHOD)

L-1 Zeros must be inserted i.e L-1=3-1=2 zeros must be inserted after every element of x(n).

After inserting zeros,

$$w(m) = [0 \ 0 \ 0 \ 3 \ 0 \ 0 \ 6 \ 0 \ 0 \ 9 \ 0 \ 0 \ 12]$$

$$b_k = [1/3 \ 2/3 \ 1 \ 2/3 \ 1/3] \text{ We have,}$$

k values varies from -2 to +2

$Y(m)$ ↓	$W(m)$ ↓	$b_{-2}=1/3$	$b_{-1}=2/3$	$b_0=1$	$b_1=2/3$	$b_2=1/3$
$Y(-2)=0$	$W(0)=0$	0	0	0	0	0
$Y(-1)=0$	$W(1)=0$	0	0	0	0	0
$Y(0)=0$	$W(2)=0$	0	0	0	0	0
$Y(1)=1$	$W(3)=3$	1	2	3	2	1
$Y(2)=2$	$W(4)=0$	0	0	0	0	0
$Y(3)=3$	$W(5)=0$	0	0	0	0	0
$Y(4)=4$	$W(6)=0$	2	4	6	4	2
$Y(5)=5$	$W(7)=0$	0	0	0	0	0
$Y(6)=6$	$W(8)=3$	0	0	0	0	0
$Y(7)=7$	$W(9)=0$	3	6	9	6	3
$Y(8)=8$	$W(10)=0$	0	0	0	0	0
$Y(9)=9$	$W(11)=3$	0	0	0	0	0
$Y(10)=10$	$W(12)=0$	4	8	12	8	4
			$Y(11)=11$	$Y(12)=12$		

$$y(n) = [0 \ 0 \ 0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ 10 \ 11 \ 12 \ 8 \ 4]$$

(these elements r beyond range of elements i.e 0 to 12 i.N=13 above delet it)

(total number of elements of both sequence before zero are to be deleted i.e 0 +2=2)

Therefore required sequence after interpolation is as shown below i.e

$$y(n) = [0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ 10 \ 11 \ 12]$$

### ALTERNATE METHOD:

L-1 Zeros must be inserted i.e L-1=3-1=2 zeros must be inserted after every element of x(n).

After inserting zeros,

$$w(m) = [0 \ 0 \ 0 \ 3 \ 0 \ 0 \ 6 \ 0 \ 0 \ 9 \ 0 \ 0 \ 12]$$

$b_k = [1/3 \ 2/3 \ 1 \ 2/3 \ 1/3]$  We have,

↑  
k values varies from -2 to +2  
substituting in the equation below we get

$$y(m) = \sum_{k=-2}^{k=2} b_k w(m-k)$$

w(m)					0	0	0	3	0	0	6	0	0	9	0	0	12
b <sub>k</sub>												×	1/3	2/3	1	2/3	1/3
					0	0	0	1	0	0	2	0	0	3	0	0	4
				0	0	0	2	0	0	4	0	0	6	0	0	8	+
			0	0	0	3	0	0	6	0	0	9	0	0	12	+	+
		0	0	0	2	0	0	4	0	0	6	0	0	8	+	+	
	0	0	0	1	0	0	2	0	0	3	0	0	4	+	+	+	+
Y(m)	0	0	0	1	2	3	4	5	6	7	8	9	10	11	12	8	4

Total no,elements before zeroth element.

elements beyond range deleted.

y(n) = ( 0 1 2 3 4 5 6 7 8 9 10 11 12 )  
↑

### PROBLEM NO 6:

The signal sequence  $x(n) = (0, 4, 8, 12, 16)$  is interpolated using the interpolation filter sequence  $b_k = (1/4, 2/4, 3/4, 1, 3/4, 2/4, 1/4)$

SOLUTION:

Given  $x(n) = x(n) = (0, 4, 8, 12, 16)$

$$b_k = (1/4, 2/4, 3/4, 1, 3/4, 2/4, 1/4)$$

Interpolated factor is  $L=2$ ,  $L-1=2-1=1$  element i.e zero is to be added in a sequence  $x(n)$ .

$$w(m) = (0, 0, 4, 0, 8, 0, 12, 0, 16)$$

$$b_k = (1/4, 2/4, 3/4, 1, 3/4, 2/4, 1/4)$$

$k$  values varies from -3 to +3

substituting in the equation below we get

$$y(m) = \sum_{k=-3}^{k=3} b_k w(m-k)$$

$$y(m) = \sum b_k w(m-k) = b_{-3} w(m+3) + b_{-2} w(m+2) + b_{-1} w(m+1) + b_0 w(m) + b_1 w(m-1) + b_2 w(m-2) + b_3 w(m-3)$$

Substituting the values of  $m$ , as  $m$  varies from 0 to 8, we get

$$\begin{aligned} y(0) &= b_{-3} w(3) + b_{-2} w(2) + b_{-1} w(1) + b_0 w(0) + b_1 w(-1) + b_2 w(-2) + b_3 w(-3) = 2 \\ y(1) &= b_{-3} w(4) + b_{-2} w(3) + b_{-1} w(2) + b_0 w(1) + b_1 w(0) + b_2 w(-1) + b_3 w(-2) = 5 \\ y(2) &= b_{-3} w(5) + b_{-2} w(4) + b_{-1} w(3) + b_0 w(2) + b_1 w(1) + b_2 w(0) + b_3 w(-1) = 8 \\ y(3) &= b_{-3} w(6) + b_{-2} w(5) + b_{-1} w(4) + b_0 w(3) + b_1 w(2) + b_2 w(1) + b_3 w(0) = 12 \\ y(4) &= b_{-3} w(7) + b_{-2} w(6) + b_{-1} w(5) + b_0 w(4) + b_1 w(3) + b_2 w(2) + b_3 w(1) = 16 \\ y(5) &= b_{-3} w(8) + b_{-2} w(7) + b_{-1} w(6) + b_0 w(5) + b_1 w(4) + b_2 w(3) + b_3 w(2) = 20 \\ y(6) &= b_{-3} w(9) + b_{-2} w(8) + b_{-1} w(7) + b_0 w(6) + b_1 w(5) + b_2 w(4) + b_3 w(3) = 24 \\ y(7) &= b_{-3} w(10) + b_{-2} w(9) + b_{-1} w(8) + b_0 w(7) + b_1 w(6) + b_2 w(5) + b_3 w(4) = 23 \\ y(8) &= b_{-3} w(11) + b_{-2} w(10) + b_{-1} w(9) + b_0 w(8) + b_1 w(7) + b_2 w(6) + b_3 w(5) = 22 \end{aligned}$$

$$y(n) = (2, 5, 8, 12, 16, 20, 24, 23, 22)$$

#### ALTERNATE METHOD (TABULATION METHOD)

$$x(n) = x(n) = (0, 4, 8, 12, 16)$$

$$b_k = (1/4, 2/4, 3/4, 1, 3/4, 2/4, 1/4)$$

Interpolated factor is  $L=2$ ,  $L-1=2-1=1$  element i.e zero is to be added in a sequence  $x(n)$ .

$$x(n) = (0, 0, 4, 0, 8, 0, 12, 0, 16)$$

$$b_k = (1/4, 2/4, 3/4, 1, 3/4, 2/4, 1/4)$$

$k$  values varies from -3 to +3

$Y(m)$ ↓	$x(n)$ ↓	$b_{-3}=1/4$	$b_{-2}=2/4$	$b_{-1}=3/4$	$b_0=1$	$b_1=3/4$	$b_2=2/4$	$b_3=1/4$
$Y(-2)=0$	$W(0)=0$	0	0	0	0	0	0	0
$Y(-1)=0$	$W(1)=0$	0	0	0	0	0	0	0
$Y(0)=1$	$W(2)=4$	1	2	3	4	3	2	1
$Y(1)=2$	$W(3)=0$	0	0	0	0	0	0	0
$Y(2)=5$	$W(4)=8$	2	4	6	8	6	4	2
$Y(3)=8$	$W(5)=0$	0	0	0	0	0	0	0
$Y(4)=12$	$W(6)=12$	3	6	9	12	9	6	3
$Y(5)=16$	$W(7)=0$	0	0	0	0	0	0	0
$Y(6)=20$	$W(8)=16$	4	8	12	16	12	8	4
		$Y(7)=24$	$Y(8)=23$	$Y(9)=22$	$Y(10)=15$	$Y(11)=8$	$Y(12)=4$	

$y(m) = [0 \ 0 \ 1 \ 2 \ 5 \ 8 \ 12 \ 16 \ 20 \ 24 \ 23 \ 22 \ 15 \ 8 \ 4]$

Beyond range

(these elements are beyond range of elements i.e 0 to 8 i.e  $N=9$  above delete it)

(total number of elements of both sequence before zero are to be deleted i.e  $0+3=3$ )

Therefore required sequence after interpolation is as shown below i.e

$Y(n) = (2, 5, 8, 12, 16, 20, 24, 23, 22)$

### ALTERNATE METHOD:

$x(n) = x(n) = (0, 4, 8, 12, 16)$

$b_k = (1/4, 2/4, 3/4, 1, 3/4, 2/4, 1/4)$

Interpolated factor is  $L=2$ ,  $L-1=2-1=1$  element i.e zero is to be added in a sequence  $x(n)$ .

$w(m) = (0, 0, 4, 0, 8, 0, 12, 0, 16)$

$b_k = (1/4, 2/4, 3/4, 1, 3/4, 2/4, 1/4)$

k values varies from -3 to +3

$w(m) =$								0	0	4	0	8	0	12	0	16
$b_k =$									$\times$	$1/4$	$2/4$	$3/4$	$1$	$3/4$	$2/4$	$1/4$
								0	0	1	0	2	0	3	0	4
						0	0	2	0	4	0	6	0	8		
					0	0	3	0	6	0	9	0	12			
				0	0	4	0	8	0	12	0	16				
			0	0	3	0	6	0	9	0	12					
		0	0	2	0	4	0	6	0	8						
	0	0	1	0	2	0	3	0	4							
$y(m)$	0	0	1	2	5	8	12	16	21	24	23	22	15	8	4	

(Total number of elements of both sequence before zero are to be deleted i.e  $0 + 3 = 3$ )  
 (These elements are beyond range of elements i.e 0 to 8 i.e  $N=9$  above delete it)

$y(n) = ( \overset{\uparrow}{2} \quad 5 \quad 8 \quad 12 \quad 16 \quad 20 \quad 24 \quad 33 \quad 22 )$

### PROBLEM NO: 7

Let the given sequence  $x(n) = [0 \ 2 \ 4 \ 6 \ 8]$  be interpolated with  $L=2$ . If the filter coefficients of the filters are  $b_k = [0.5, 1, 0.5]$ , obtain the interpolated sequence.

### SOLUTION:

$L-1$  Zeros must be inserted i.e  $L-1=2-1=1$  zeros must be inserted after every element of  $x(n)$ .

After inserting zeros,

$$w(m) = [0 \ 0 \ 2 \ 0 \ 4 \ 0 \ 6 \ 0 \ 8]$$

$$b_k = [0.5 \quad \underset{\uparrow}{1} \quad 0.5] \quad \text{We have,}$$

k values varies from -1 to +1  
substituting in the equation below we get

$$y(m) = \sum_{k=-1}^{k=1} b_k w(m-k)$$

$$y(m) = \sum b_k w(m-k) = b_{-1} w(m+1) + b_0 w(m) + b_1 w(m-1)$$

Substituting the values of m, as m varies from 0 to 8, we get

$$\begin{aligned} y(0) &= b_{-1} w(1) + b_0 w(0) + b_1 w(-1) = 0 \\ y(1) &= b_{-1} w(2) + b_0 w(1) + b_1 w(0) = 1 \\ y(2) &= b_{-1} w(3) + b_0 w(2) + b_1 w(1) = 2 \\ y(3) &= b_{-1} w(4) + b_0 w(3) + b_1 w(2) = 3 \\ y(4) &= b_{-1} w(5) + b_0 w(4) + b_1 w(3) = 4 \\ y(5) &= b_{-1} w(6) + b_0 w(5) + b_1 w(4) = 5 \\ y(6) &= b_{-1} w(7) + b_0 w(6) + b_1 w(5) = 6 \\ y(7) &= b_{-1} w(8) + b_0 w(7) + b_1 w(6) = 7 \\ y(8) &= b_{-1} w(9) + b_0 w(8) + b_1 w(7) = 8 \end{aligned}$$

$$y(n) = ( \quad \underset{\uparrow}{0} \quad 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad )$$

#### ALTERNATE METHOD:

L-1 Zeros must be insertes i.e L-1=2-1=1 zeros must be inserted after every element of x(n).

After inserting zeros,

$$w(m) = [0 \ 0 \ 2 \ 0 \ 4 \ 0 \ 6 \ 0 \ 8]$$

$$b_k = [0.5 \quad \underset{\uparrow}{1} \quad 0.5] \quad \text{We have,}$$

k values varies from -1 to +1  
substituting in the equation below we get

$$y(m) = \sum_{k=-1}^{k=1} b_k w(m-k)$$

$w(m)$			0	0	2	0	4	0	6	0	8
$b_k$								$\times$	0.5	1	0.5
			0	0	1	0	2	0	3	0	4
		0	0	2	0	4	0	6	0	8	+
	0	0	1	0	2	0	3	0	4	+	+
$Y(m)$	0	0	1	2	3	4	5	6	7	8	4

$y(m) = ( 0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 )$

elements before 0 th are deleted range
 elements beyond

## Assignment no: 1 (Recommended Questions )

1. The sequence  $x(n) = [3, 2, -2, 0, 7]$ . It is interpolated using interpolation sequence  $b_k = [0.5, 1, 0.5]$  and the interpolation factor of 2. Find the interpolated sequence  $y(m)$  [june-2016, 6M, jan-2017, 8M].
2. The sequence  $x(n) = [0, 2, 4, 6, 8]$ . It is interpolated using interpolation sequence





- $b_k=[0.5,1,0.5]$  and the interpolation factor of 2. Find the interpolated sequence  $y(m)$  [june/july-2013,6M].
3. The sequence  $x(n) = [0,4,8,12,16]$ . It is interpolated using interpolation sequence  $b_k=[0.25,0.5,0.75,1, 0.75,0.5,0.25]$  and the interpolation factor of 2. Find the interpolated sequence  $y(m)$  [june/july-2013,6M].
  4. An analog signal is sampled at the rate of 8KHz. If 512 samples of this signal are used to compute DFT  $X(k)$  determine the analog and digital frequency spacing between adjacent  $X(k)$  elements. Also, determine analog and digital frequencies corresponding to  $k=60$  [june-2016,6M.jan-2017,8M].
  5. With a neat diagram explain the scheme of the DSP system. Draw the timing diagram [DEC-1013,8M.jan-2017,6M]
  6. With the help of block diagram and equation explain decimation and interpolation process also find output for sequence  $x(n) = [0,3,6,9]$  with  $b_k=[1/3,2/3,1, 2/3,1/3]$  and the interpolation factor of 2 [DEC-1013,10M].
  7. List and explain the issues that have to be considered in designing and implementing a DSP system [june/july-2013,4M. june-2016,6M].
  8. What is DSP? List the unique architectural features of DSP processor [DEC-2015,5M].
  9. What is DSP? What are the important issues to be considered in designing and implementing a DSP system? Explain in detail [DEC-2012,9M.JUNE/JULY 2014,09M].
  10. "FIR filter are linear phase filters". Justify the same with magnitude and phase plot [DEC-2015,5M]
  11. Write matlab code for design an FIR filter using Parks-McClellan method [JUNE/JULY 2014,05M]
  12. Why signal sampling is required? Explain the sampling process [DEC-2011,9M].
  13. Define decimation and interpolation process. Explain them using block diagrams and equations. [DEC-2012,6M. DEC-2013,6M.june-2014,6M].
  14. Explain common features of programmable digital signal processor [june/july-2015,8M]
  15. Implement the FIR Filter for  $y(n)=(x(n)+ x(n-1)+ x(n-2))/3$ . Describe the i) system function ii) magnitude response function iii) phase response function iv) step response v) group delay. plot its magnitude and phase response [june/july-2013,4M].
  16. Explain the processes of Decimation [june/july-2013,4M]
  17. List the major architectural features used in DSP system to achieve high speed program execution [DEC-2014,3M].

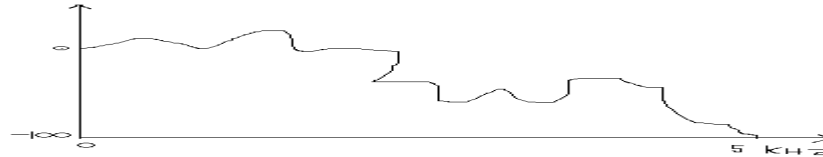


18. An FFT is employed for determining the frequency components of a random signal. It is required that the resolution of FFT to be  $\leq 5\text{Hz}$ , for a signal with  $f_{\max} = 1.25\text{ kHz}$ . Determine i) Sampling interval ii) FFT length(N) as a power of 2 iii) Minimum signal record length [DEC-2014,4M].
19. Explain the two method of sampling rate conversions used in DSP system, with suitable block diagrams and examples. Draw the corresponding spectrum [Dec 2011,8M].
20. Assuming  $X(K)$  as a complex sequence determine the number of complex real multiplies for computing IDFT using direct and Radix-2 FT algorithms.
21. With a neat diagram explain the scheme of a DSP system. (June.12,8m.Dec.10-Jan.11,June/july2011 8m. DEC-2014,9M.JUNE-2015,4M)
22. With an example explain the need for the low pass filter in decimation process. (June.12, 4m)
23. For the FIR filter  $y(n)=(x(n)+x(n-1)+x(n-2))/3$ . Determine i) System Function ii) Magnitude and phase function iii) Step response iv) Group Delay. (June.12, 8m)
24. List the major architectural features used in DSP system to achieve high speed program execution. (Dec.11, 6m).
25. Explain how to simulate the impulse responses of FIR and IIR filters. (Dec.11, 6m).
26. Explain the two method of sampling rate conversions used in DSP system, with suitable block diagrams and examples. Draw the corresponding spectrum. (Dec.11, 8m).
27. Explain with the help of mathematical equations how signed numbers can be multiplied. (July.11, 8m).
28. Describe the basic features that should be provided in the DSP architecture to be used to implement the Nth order FIR filter, where  $x(n)$  denotes the input sample,  $y(n)$  the output sample and  $h(i)$  denotes  $i^{\text{th}}$  filter coefficient.(Dec.09-Jan.10, 8m)
29. Explain the issues to be considered in designing and implementing a DSP system, with the help of a neat block diagram. (May/June10 , 6m)
30. Briefly explain the major features of programmable DSPs. (May/June10, 8m)
31. Explain the operation used in DSP to increase the sampling rate. The sequence  $x(n)=[0,2,4,6,8]$  is interpolated using interpolation sequence  $b_k=[1/2,1,1/2]$  and the interpolation factor is 2.find the interpolated sequence  $y(m)$ . (May/June10, 8m)
32. Explain with the help of mathematical equations how signed numbers can be multiplied. (Dec.10-Jan.11, 8m)
33. The sequence  $x(n) = [5,7,-2,0,3]$ .It is interpolated using interpolation sequence  $b_k=[0.5,1,0.5]$  and the interpolation factor of 2. Find the interpolated sequence  $y(m)$ .(Dec.10-Jan.11, 6m)



34. Why signal sampling is required? Explain the sampling process. (Dec.12, 5m)
35. Define decimation and interpolation process. Explain them using block diagrams and equations. (Dec.12, 6m).
36. The signal in the following figure is to be sampled. Determine the minimum sample rate without any aliasing effect. If the signal is sampled at a rate 8 kHz, determine the cut off frequency of the anti-aliasing filter.

**Magnitude B**



**FIG:1**

**Freq**

# COMPUTATIONAL ACCURACY IN DSP IMPLEMENTATIONS

## ❖ LEARNING OBJECTIVES

- Number Formats for Signals and Coefficients in DSP Systems,
- Dynamic Range and Precision
- Sources of Error in DSP Implementation

### 1.1 Number Formats for Signals and Coefficients in DSP Systems

- In a DSP signals are represented in a different ways.
- Depending on the range and precision of signals and coefficient to be represented, hardware complexity and speed requirement we will see typical formats used to represent signals and coefficient in DSP system.

- 1) FIXED-POINT FORMATE
- 2) FLOATING-POINT FORMATE
- 3) BLOCK FLOATING-POINT FORMATE

### 1.2 FIXED-POINT FORMATE

The simplest scheme of number representation is the format in which the number is represented as an integer or fraction using fixed number of bits.

$$X = -s \cdot 2^{n-1} + b_{n-2} \cdot 2^{n-2} + b_{n-3} \cdot 2^{n-3} + \dots + b_1 \cdot 2^1 + b_0 \cdot 2^0$$

where, s represents the sign of the number: s=0 for + number

s= -1 for – numbers

range of signed number is given as  $-2^{n-1}$  to  $+(2^{n-1}-1)$

s	b <sub>n-1</sub>		b <sub>2</sub>	b <sub>1</sub>	b <sub>0</sub>
n-1	n-2		2	1	0

for the fractions

$$x = -s \cdot 2^0 + b_{-1} \cdot 2^{-1} + b_{-2} \cdot 2^{-2} + \dots + b_{-(n-2)} \cdot 2^{-(n-2)} + b_{-(n-1)} \cdot 2^{-(n-1)}$$



the range is given as -1 to  $+(1-2^{-(n-1)})$

s	b <sub>1</sub>		b <sub>(n-3)</sub>	b <sub>(n-2)</sub>	b <sub>(n-1)</sub>
n-1	n-2		2	1	0

Examples: 1)

What is the range of number that can be represented in a fixed point format using 16 bits if the numbers are treated as (a) signed integers (b) signed fractions?

- Using 16 bits, the range of integers that can be represented is determined by substituting n=16 and is given as  $-2^{15}$  to  $+2^{15-1}$  i.e -32768 to + 32767.
- The range of fraction using n=16 is given as -1 to  $+(1-2^{-15})$  i.e -1 to + .999969482

Examples: 2)

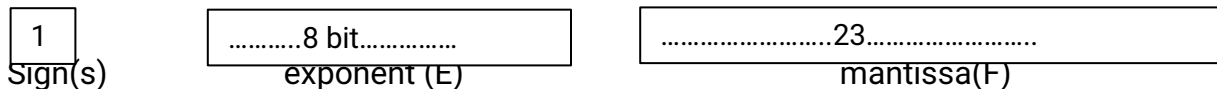
What is the range of number that can be represented in a fixed point format using 32 bits if the numbers are treated as (a) signed integers (b) signed fractions?

- Using 32 bits, the range of integers that can be represented is determined by substituting n=16 and is given as  $-2^{31}$  to  $+2^{31-1}$  i.e -32768 to +.
- The range of fraction using n=16 is given as -1 to  $+(1-2^{-31})$  i.e -1 to +

### 1.3 FLOATING POINT FORMATE

A floating point is made up of a mantissa  $M_x$  and an exponent  $E_x$  such as its value x is represented as

$$x = M_x 2^{E_x}$$



$$x = (-1) \times 2^{(E-\text{bias})} \times 1.F$$

Example:1

Find the decimal equivalent of the floating point binary number 1011000011100.  
Assume a format similar to IEEE-754 in which the MSB is the sign bit followed by 4 exponent bits followed by 8 bits for the fractional part.

Solution:

$$F = 2^{-4} + 2^{-5} + 2^{-6} = .109375$$

$$E = 2^1 + 2^2 = 6$$

Thus the value of the number is  $x = -1.109375 \times 2^{(6-7)} = -0.5546875$

Example:2

Find the decimal equivalent of the floating point binary number 1011000011100.  
Assume a format similar to IEEE-754 in which the MSB is the sign bit followed by 4 exponent bits followed by 8 bits for the fractional part.

Solution:

$$F = 2^{-4} + 2^{-5} + 2^{-6} = .109375$$

$$E = 2^1 + 2^2 = 6$$

Thus the value of the number is  $x = -1.109375 \times 2^{(6-7)} = -0.5546875$

Example:3

Find the decimal equivalent of the floating point binary number 1011000011100.  
Assume a format similar to IEEE-754 in which the MSB is the sign bit followed by 4 exponent bits followed by 8 bits for the fractional part.

Solution:

$$F = 2^{-4} + 2^{-5} + 2^{-6} = .109375$$

$$E = 2^1 + 2^2 = 6$$

Thus the value of the number is  $x = -1.109375 \times 2^{(6-7)} = -0.5546875$

Example:4

Using 16 bits for the mantissa and 8 bits for the exponent, what is the range of number that can be represented using the floating point format similar to IEEE-654?

Solution:

The most negative number will have as its mantissa  $-2 = 2^{-16}$  and as its exponent (255-127). The most negative number is, therefore

$$-1.999984741 \times 2^{128}$$

Similarly, the most positive number is

$$+1.999984741 \times 2^{128}$$

#### 1.4 BLOCK FLOATING POINT FORMAT

- An approach to increase the range and precision of the fixed point format is to use the block floating point format.
- In this approach, a group or block of fixed point numbers are represented as though they were floating point number with the same exponent value and different mantissa values.
- Mantissas are stored and handled similar to fixed point number with the same



exponent value and different mantissa values.

- The block floating point format increases the range and precision of given fixed point format by retaining as many lower order bits as is possible.

Example:1)

The following 12 bits binary fraction are to be stored in an 8 bit memory show how they can be represented in block floating point format so as to improve accuracy.

000001110011

000011110000

000000111111

000010101010

Solution:

If these fractions are represented using an 8 bit fixed point format they will be represented as

00000111

00001111

00000011

00001010

The last 4 bits of the number would have been discarded, there losing the precision corresponding to those 4 bits.

They can be written as

$01110011 \times 2^{-4}$

$00001111 \times 2^{-4}$

$00000011 \times 2^{-4}$

$00001010 \times 2^{-4}$

Eight bits of each number can be stored without discarding any bits. The block exponent is -4 and will have to be stored separately. When the numbers are read from the memory for any computation they have to be shifted by four bits positions to the right then to their original values.

### 1.5 DYNAMIC RANGE AND PRECISION

The dynamic range of a signal is the ratio of the maximum value to the minimum value that the signal can in the given number representation scheme.

The dynamic range of the signal is proportional to the number of bits used to represent it and increases by 6db for every additional bit used for the representation.