Tania Stathaki 811b t.stathaki@imperial.ac.uk

What is a signal?

A signal is a function of an independent variable such as time, distance, position, temperature, pressure, etc.

For example...

- Electrical Engineering

 voltages/currents in a circuit
 speech signals
 image signals
- Physics
 radiation
- Mechanical Engineering vibration studies
- Astronomy space photos



• **Biomedicine** EEG, ECG, MRI, X-Rays, Ultrasounds

- Seismology tectonic plate movement, earthquake prediction
- Economics
 - stock market data



What is DSP?

Mathematical and algorithmic manipulation of **discretized and quantized** or **naturally digital** signals in order to extract the most relevant and pertinent information that is carried by the signal.



What is a signal? What is a system? What is processing?

Signals can be characterized in several ways

Continuous time signals vs. discrete time signals (x(t), x[n]).

Temperature in London / signal on a CD-ROM.

Continuous valued signals vs. discrete signals.

Amount of current drawn by a device / average scores of TOEFL in a school over years.

-Continuous time and continuous valued : Analog signal.

-Continuous time and discrete valued: Quantized signal.

-Discrete time and continuous valued: **Sampled signal.**

-Discrete time and discrete values: **Digital signal.**

Real valued signals vs. complex valued signals.

Resident use electric power / industrial use reactive power.

Scalar signals vs. vector valued (multi-channel) signals.

Blood pressure signal / 128 channel EEG.

Deterministic vs. random signal:

Recorded audio / noise.

One-dimensional vs. two dimensional vs. multidimensional signals.

Speech / still image / video.



Systems

- For our purposes, a DSP system is one that can *mathematically manipulate (e.g., change, record, transmit, transform) digital signals*.
- Furthermore, we are not interested in processing analog signals either, even though most signals in nature are analog signals.



Various Types of Processing

Modulation and demodulation. Signal security. **Encryption and decryption.** Multiplexing and de-multiplexing. **Data compression.** Signal de-noising. **Filtering for noise reduction.** Speaker/system identification. Signal enhancement –equalization. Audio processing. Image processing – image de-noising, enhancement, watermarking. **Reconstruction. Data analysis and feature extraction. Frequency/spectral analysis.**

Filtering

• By far the most commonly used DSP operation

Filtering refers to deliberately changing the frequency content of the signal, typically, by removing certain frequencies from the signals.

For de-noising applications, the (frequency) filter removes those frequencies in the signal that correspond to noise.

In various applications, filtering is used to focus to that part of the spectrum that is of interest, that is, the part that carries the information.

- Typically we have the following types of filters
 - Low-pass (LPF) –removes high frequencies, and retains (passes) low frequencies.
 - High-pass (HPF) –removes low frequencies, and retains high frequencies.
 - **Band-pass (BPF)** –retains an interval of frequencies within a band, removes others.
 - **Band-stop (BSF)** –removes an interval of frequencies within a band, retains others.
 - Notch filter –removes a specific frequency.

A Common Application: Filtering



Components of a DSP System



Components of a DSP System



Components of a DSP System



Analog-to-Digital-to-Analog...?

- Why not just process the signals in continuous time domain? Isn't it just a waste of time, money and resources to convert to digital and back to analog?
- Why DSP? We digitally process the signals in discrete domain, because it is
 - More flexible, more accurate, easier to mass produce.
 - Easier to design.
 - System characteristics can easily be changed by programming.
 - Any level of accuracy can be obtained by use of appropriate number of bits.
 - More deterministic and reproducible-less sensitive to component values, etc.
 - Many things that cannot be done using analog processors can be done digitally.
 - Allows multiplexing, time sharing, multi-channel processing, adaptive filtering.
 - Easy to cascade, no loading effects, signals can be stored indefinitely w/o loss.
 - Allows processing of very low frequency signals, which requires unpractical component values in analog world.

Analog-to-Digital-to-Analog...?

- On the other hand, it can be
 - Slower, sampling issues.
 - More expensive, increased system complexity, consumes more power.
- Yet, the advantages far outweigh the disadvantages. Today, most continuous time signals are in fact processed in discrete time using digital signal processors.

Analog-Digital

Examples of analog technology

- photocopiers
- telephones
- audio tapes
- televisions (intensity and color info per scan line)
- VCRs (same as TV)

Examples of digital technology

• Digital computers!

In the next few slides you can see some real-life signals

Electroencephalogram (EEG) Data



Stock Market Data



Satellite image Volcano Kamchatka Peninsula, Russia



Satellite image Volcano in Alaska



Medical Images: MRI of normal brain



Medical Images: X-ray knee



Medical Images: Ultrasound Five-month Foetus (lungs, liver and bowel)



Astronomical images



Spiral Galaxy NGC 1232 - VLT UT 1 + FORS1



ESO PR Photo 37d/98 (23 September 1998)

©European Southern Observatory

- Signals represented as sequences of numbers, called **samples**
- Sample value of a typical signal or sequence denoted as *x*[*n*] with *n* being an integer in the range −∞ ≤ n ≤ ∞
- *x*[*n*] defined only for integer values of *n* and undefined for noninteger values of *n*
- Discrete-time signal represented by {*x*[*n*]}

• Here, *n*-th sample is given by

$$x[n] = x_a(t)|_{t=nT} = x_a(nT), n = \dots, -2, -1, 0, 1, \dots$$

- The spacing *T* is called the **sampling interval** or **sampling period**
- Inverse of sampling interval *T*, denoted as F_T , is called the **sampling frequency**: $F_T = (T)^{-1}$



- Two types of discrete-time signals:
 - **Sampled-data signals** in which samples are continuous-valued
 - **Digital signals** in which samples are discrete-valued
- Signals in a practical digital signal processing system are digital signals obtained by quantizing the sample values either by **rounding** or **truncation**

2 Dimensions From Continuous to Discrete: Sampling

256x256



64x64







256x256 256 levels



256x256 32 levels



256x256 256 levels



256x256 2 levels



- A discrete-time signal may be a **finite-length** or an **infinite-length sequence**
- Finite-length (also called **finite-duration** or **finite-extent**) sequence is defined only for a finite time interval: $N_1 \le n \le N_2$

where $-\infty < N_1$ and $N_2 < \infty$ with $N_1 \le N_2$

• Length or duration of the above finitelength sequence is $N = N_2 - N_1 + 1$
Discrete-Time Signals: Time-Domain Representation

• A **right-sided sequence** *x*[*n*] has zerovalued samples for *n* < *N*₁



A right-sided sequence

If N₁ ≥ 0, a right-sided sequence is called a causal sequence

Discrete-Time Signals: Time-Domain Representation

A left-sided sequence x[n] has zero-valued samples for n > N₂



A left-sided sequence

If N₂ ≤ 0, a left-sided sequence is called a anti-causal sequence

Operations on Sequences

• A single-input, single-output discrete-time system operates on a sequence, called the **input sequence**, according some prescribed rules and develops another sequence, called the output sequence, with more desirable properties



Example of an Operation on a Sequence: Noise Removal

- For example, the input may be a signal corrupted with additive noise
- A discrete-time system may be designed to generate an output by removing the noise component from the input
- In most cases, the operation defining a particular discrete-time system is composed of some **basic operations**

• **Product (modulation)** operation:

- Modulator

$$x[n] \xrightarrow{} y[n] \longrightarrow y[n]$$

 $w[n] \qquad y[n] = x[n] \cdot w[n]$

- An application is the generation of a finite-length sequence from an infinite-length sequence by multiplying the latter with a finite-length sequence called an **window sequence**
- Process called **windowing**

• Addition operation:



• Multiplication operation

- Multiplier
$$x[n] \longrightarrow y[n] \quad y[n] = A \cdot x[n]$$

- Time-shifting operation: y[n] = x[n-N]where N is an integer
- If N > 0, it is a **delay** operation

- Unit delay
$$x[n] \longrightarrow z^{-1} \longrightarrow y[n] \quad y[n] = x[n-1]$$

• If N < 0, it is an **advance** operation

- Unit advance
$$x[n] \longrightarrow z \longrightarrow y[n] \quad y[n] = x[n+1]$$

- Time-reversal (folding) operation: y[n] = x[-n]
- **Branching** operation: Used to provide multiple copies of a sequence

$$x[n] \longrightarrow x[n]$$
$$x[n]$$

Combinations of Basic Operations

• Example -



 $y[n] = \alpha_1 x[n] + \alpha_2 x[n-1] + \alpha_3 x[n-2] + \alpha_4 x[n-3]$

- Employed to generate a new sequence y[n]with a sampling rate F_T higher or lower than that of the sampling rate F_T of a given sequence x[n]
- sequence x[n]• Sampling rate alteration ratio is $R = \frac{F_T}{F_T}$
- If R > 1, the process called **interpolation**
- If R < 1, the process called **decimation**

In up-sampling by an integer factor L > 1,
 L-1 equidistant zero-valued samples are inserted by the up-sampler between each two consecutive samples of the input

sequence *x*[*n*]:

 $x_{u}[n] = \begin{cases} x[n/L], & n = 0, \pm L, \pm 2L, \cdots \\ 0, & \text{otherwise} \end{cases}$

$$x[n] \longrightarrow \uparrow L \longrightarrow x_u[n]$$

• An example of the up-sampling operation



In down-sampling by an integer factor
 M > 1, every M-th samples of the input sequence are kept and M -1 in-between samples are removed:

y[n] = x[nM]

$$x[n] \longrightarrow M \longrightarrow y[n]$$

• An example of the down-sampling operation



• Conjugate-symmetric sequence:

$$x[n] = x * [-n]$$

If *x*[*n*] is real, then it is an **even sequence**



An even sequence

• Conjugate-antisymmetric sequence: x[n] = -x * [-n]

If *x*[*n*] is real, then it is an **odd sequence**



An odd sequence

- It follows from the definition that for a conjugate-symmetric sequence {x[n]}, x[0] must be a real number
- Likewise, it follows from the definition that for a conjugate anti-symmetric sequence {y[n]}, y[0] must be an imaginary number
- From the above, it also follows that for an odd sequence {w[n]}, w[0] = 0

• Any complex sequence can be expressed as a sum of its conjugate-symmetric part and its conjugate-antisymmetric part:

$$x[n] = x_{cs}[n] + x_{ca}[n]$$

where

$$x_{cs}[n] = \frac{1}{2} (x[n] + x^{*}[-n])$$
$$x_{ca}[n] = \frac{1}{2} (x[n] - x^{*}[-n])$$

• Any real sequence can be expressed as a sum of its even part and its odd part:

$$x[n] = x_{ev}[n] + x_{od}[n]$$

where

$$x_{ev}[n] = \frac{1}{2} (x[n] + x[-n])$$
$$x_{od}[n] = \frac{1}{2} (x[n] - x[-n])$$

Classification of Sequences Based on Periodicity

- A sequence \$\tilde{x}[n]\$ satisfying \$\tilde{x}[n] = \$\tilde{x}[n + kN]\$ is called a **periodic sequence** with a **period** N where N is a positive integer and k is any integer
- Smallest value of *N* satisfying $\tilde{x}[n] = \tilde{x}[n+kN]$ is called the **fundamental period**



• A sequence not satisfying the periodicity condition is called an **aperiodic sequence**

• Total **energy** of a sequence *x*[*n*] is defined by

$$\mathcal{E}_{\mathbf{x}} = \sum_{n=-\infty}^{\infty} |x[n]|^2$$

- An infinite length sequence with finite sample values may or may not have finite energy
- A finite length sequence with finite sample values has finite energy

• The **average power** of an aperiodic sequence is defined by

$$P_{x} = \lim_{K \to \infty} \frac{1}{2K+1} \sum_{n=-K}^{K} |x[n]|^{2}$$

• We define the **energy** of a sequence x[n]over a finite interval $-K \le n \le K$ as

$$\boldsymbol{\mathcal{E}}_{x,K} = \sum_{n=-K}^{K} |\boldsymbol{x}[n]|^2$$

 The average power of a periodic sequence *x*[*n*] with a period *N* is given by

$$P_x = \frac{1}{N} \sum_{n=0}^{N-1} \left| \widetilde{x}[n] \right|^2$$

• The average power of an infinite-length sequence may be finite or infinite

• <u>Example</u> - Consider the causal sequence defined by

$$x[n] = \begin{cases} 3(-1)^n, & n \ge 0\\ 0, & n < 0 \end{cases}$$

- Note: *x*[*n*] has infinite energy
- Its average power is given by

$$P_{x} = \lim_{K \to \infty} \frac{1}{2K+1} \left(9\sum_{n=0}^{K} 1\right) = \lim_{K \to \infty} \frac{9(K+1)}{2K+1} = 4.5$$

• An infinite energy signal with finite average power is called a **power signal**

<u>Example</u> - A periodic sequence which has a finite average power but infinite energy

• A finite energy signal with zero average power is called an **energy signal**

Classification of Sequences: Deterministic-Stochastic





Other Types of Classifications

- A sequence x[n] is said to be **bounded** if $|x[n]| \le B_x < \infty$
- <u>Example</u> The sequence $x[n] = \cos 0.3\pi n$ is a bounded sequence as $|x[n]| = |\cos 0.3\pi n| \le 1$

Other Types of Classifications

- A sequence x[n] is said to be absolutely summable if $\sum_{n=-\infty}^{\infty} |x[n]| < \infty$
- <u>Example</u> The sequence

$$y[n] = \begin{cases} 0.3^n, & n \ge 0\\ 0, & n < 0 \end{cases}$$

is an absolutely summable sequence as

$$\sum_{n=0}^{\infty} \left| 0.3^n \right| = \frac{1}{1 - 0.3} = 1.42857 < \infty$$

Other Types of Classifications

- A sequence x[n] is said to be squaresummable if
- Example The sequence $h[n] = \frac{\sin 0.4n}{\pi n}$ is square-summable but not absolutely summable

• Unit sample sequence - $\delta[n] = \begin{cases} 1, & n=0 \\ 0, & n \neq 0 \end{cases}$



• Unit step sequence - $\mu[n] = \begin{cases} 1, & n \ge 0 \\ 0, & n < 0 \end{cases}$



• Real sinusoidal sequence -

 $x[n] = A\cos(\omega_o n + \phi)$

where A is the **amplitude**, ω_o is the **angular** frequency, and ϕ is the **phase** of x[n]





• Complex exponential sequence -

 $x[n] = A \alpha^n, -\infty < n < \infty$

where A and α are real or complex numbers

• If we write $\alpha = e^{(\sigma_o + j\omega_o)}, A = |A|e^{j\phi},$

then we can express

 $x[n] = |A|e^{j\phi}e^{(\sigma_o + j\omega_o)n} = x_{re}[n] + j x_{im}[n],$ where

 $x_{re}[n] = |A|e^{\sigma_o n}\cos(\omega_o n + \phi),$ $x_{im}[n] = |A|e^{\sigma_o n}\sin(\omega_o n + \phi)$

• $x_{re}[n]$ and $x_{im}[n]$ of a complex exponential sequence are real sinusoidal sequences with constant ($\sigma_o = 0$), growing ($\sigma_o > 0$), and decaying ($\sigma_o < 0$) amplitudes for n > 0



• Real exponential sequence -

$$x[n] = A\alpha^n, -\infty < n < \infty$$

where A and α are real or complex numbers



- Sinusoidal sequence $A\cos(\omega_o n + \phi)$ and complex exponential sequence $B\exp(j\omega_o n)$ are periodic sequences of period N if $\omega_o N = 2\pi r$ where N and r are positive integers
- Smallest value of *N* satisfying $\omega_o N = 2\pi r$ is the **fundamental period** of the sequence
- To verify the above fact, consider $x_1[n] = \cos(\omega_o n + \phi)$ $x_2[n] = \cos(\omega_o (n + N) + \phi)$

• Now $x_2[n] = \cos(\omega_o(n+N) + \phi)$ = $\cos(\omega_o n + \phi) \cos \omega_o N - \sin(\omega_o n + \phi) \sin \omega_o N$ which will be equal to $\cos(\omega_o n + \phi) = x_1[n]$ only if

 $\sin \omega_o N = 0$ and $\cos \omega_o N = 1$

• These two conditions are met if and only if $\omega_o N = 2\pi r$ or $\frac{2\pi}{\omega_o} = \frac{N}{r}$
- If $2\pi/\omega_o$ is a noninteger rational number, then the period will be a multiple of $2\pi/\omega_o$
- Otherwise, the sequence is **aperiodic**
- Example $x[n] = sin(\sqrt{3}n + \phi)$ is an aperiodic sequence



• Here $\omega_o = 0$

• Hence period
$$N = \frac{2\pi r}{0} = 1$$
 for $r = 0$



• Here $\omega_o = 0.1\pi$

• Hence
$$N = \frac{2\pi r}{0.1\pi} = 20$$
 for $r = 1$

- <u>Property 1</u> Consider $x[n] = \exp(j\omega_1 n)$ and $y[n] = \exp(j\omega_2 n)$ with $0 \le \omega_1 < \pi$ and $2\pi k \le \omega_2 < 2\pi (k+1)$ where k is any positive integer
- If $\omega_2 = \omega_1 + 2\pi k$, then x[n] = y[n]
- Thus, *x*[*n*] and *y*[*n*] are indistinguishable

- Property 2 The frequency of oscillation of Acos(ω_on) increases as ω_o increases from 0 to π, and then decreases as ω_o increases from π to 2π
- Thus, frequencies in the neighborhood of

 ω = 0 are called low frequencies, whereas,
 frequencies in the neighborhood of ω = π are
 called high frequencies

• Because of Property 1, a frequency ω_o in the neighborhood of $\omega = 2\pi$ k is indistinguishable from a frequency $\omega_o - 2\pi k$ in the neighborhood of $\omega = 0$

and a frequency ω_o in the neighborhood of $\omega = \pi(2k+1)$ is indistinguishable from a frequency $\omega_o - \pi(2k+1)$ in the neighborhood of $\omega = \pi$

- Frequencies in the neighborhood of $\omega = 2\pi k$ are usually called **low frequencies**
- Frequencies in the neighborhood of $\omega = \pi (2k+1)$ are usually called high frequencies
- $v_1[n] = \cos(0.1\pi n) = \cos(1.9\pi n)$ is a lowfrequency signal
- $v_2[n] = \cos(0.8\pi n) = \cos(1.2\pi n)$ is a high-frequency signal

• An arbitrary sequence can be represented in the time-domain as a weighted sum of some basic sequence and its delayed (advanced) versions



 $x[n] = 0.5\delta[n+2] + 1.5\delta[n-1] - \delta[n-2]$ $+ \delta[n-4] + 0.75\delta[n-6]$

 Often, a discrete-time sequence x[n] is developed by uniformly sampling a continuous-time signal x_a(t) as indicated below



• The relation between the two signals is

$$x[n] = x_a(t)|_{t=nT} = x_a(nT), n = \dots, -2, -1, 0, 1, 2, \dots$$

Time variable t of x_a(t) is related to the time variable n of x[n] only at discrete-time instants t_n given by

$$t_n = nT = \frac{n}{F_T} = \frac{2\pi n}{\Omega_T}$$

with $F_T = 1/T$ denoting the sampling frequency and

 $\Omega_T = 2\pi F_T$ denoting the sampling angular frequency

Hertz
The Sampling Process
• Consider the continuous-time signal

$$x(t) = A\cos(2\pi f_o t + \phi) = A\cos(\Omega_o t + \phi)$$

• The corresponding discrete-time signal is $x[n] = A\cos(\Omega_o nT + \phi) = A\cos(\frac{2\pi\Omega_o}{\Omega_T}n + \phi)$ $= A\cos(\omega_o n + \phi)$ seconds

where $\omega_o = 2\pi \Omega_o / \Omega_T = \Omega_o F$ —radians per second is the normalized digital angular frequency of x[n]

radians per sample

- If the unit of sampling period *T* is in seconds
- The unit of normalized digital angular frequency ω_o is radians/sample
- The unit of normalized analog angular frequency Ω_o is radians/second
- The unit of analog frequency f_o is hertz (Hz)

• The three continuous-time signals

 $g_1(t) = \cos(6\pi t)$ $g_2(t) = \cos(14\pi t)$ $g_3(t) = \cos(26\pi t)$

of frequencies 3 Hz, 7 Hz, and 13 Hz, are sampled at a sampling rate of 10 Hz, i.e. with T = 0.1 sec. generating the three sequences $g_1[n] = \cos(0.6\pi n)$ $g_2[n] = \cos(1.4\pi n)$

 $g_3[n] = \cos(2.6\pi n)$

 Plots of these sequences (shown with circles) and their parent time functions are shown below:



• Note that each sequence has exactly the same sample value for any given *n*

- This fact can also be verified by observing that $g_2[n] = \cos(1.4\pi n) = \cos((2\pi - 0.6\pi)n) = \cos(0.6\pi n)$ $g_3[n] = \cos(2.6\pi n) = \cos((2\pi + 0.6\pi)n) = \cos(0.6\pi n)$
 - As a result, all three sequences are identical and it is difficult to associate a unique continuous-time function with each of these sequences

• The above phenomenon of a continuoustime signal of higher frequency acquiring the identity of a sinusoidal sequence of lower frequency after sampling is called **aliasing**

- Since there are an infinite number of continuous-time signals that can lead to the same sequence when sampled periodically, additional conditions need to imposed so that the sequence {x[n]} = {x_a(nT)} can uniquely represent the parent continuous-time signal x_a(t)
- In this case, x_a(t) can be fully recovered from {x[n]}

- <u>Example</u> Determine the discrete-time signal *v*[*n*] obtained by uniformly sampling at a sampling rate of 200 Hz the continuoustime signal
 - $v_a(t) = 6\cos(60\pi t) + 3\sin(300\pi t) + 2\cos(340\pi t)$ $+ 4\cos(500\pi t) + 10\sin(660\pi t)$
- Note: $v_a(t)$ is composed of 5 sinusoidal signals of frequencies 30 Hz, 150 Hz, 170 Hz, 250 Hz and 330 Hz

- The sampling period is $T = \frac{1}{200} = 0.005$ sec
- The generated discrete-time signal *v*[*n*] is thus given by

 $v[n] = 6\cos(0.3\pi n) + 3\sin(1.5\pi n) + 2\cos(1.7\pi n)$

 $+ 4\cos(2.5\pi n) + 10\sin(3.3\pi n)$

 $= 6\cos(0.3\pi n) + 3\sin((2\pi - 0.5\pi)n) + 2\cos((2\pi - 0.3\pi)n) + 4\cos((2\pi + 0.5\pi)n) + 10\sin((4\pi - 0.7\pi)n)$

 $= 6\cos(0.3\pi n) - 3\sin(0.5\pi n) + 2\cos(0.3\pi n) + 4\cos(0.5\pi n)$ $-10\sin(0.7\pi n)$

 $= 8\cos(0.3\pi n) + 5\cos(0.5\pi n + 0.6435) - 10\sin(0.7\pi n)$

Note: v[n] is composed of 3 discrete-time sinusoidal signals of normalized angular frequencies: 0.3π, 0.5π, and 0.7π

 Note: An identical discrete-time signal is also generated by uniformly sampling at a 200-Hz sampling rate the following continuous-time signals:

 $w_a(t) = 8\cos(60\pi t) + 5\cos(100\pi t + 0.6435) - 10\sin(140\pi t)$

 $g_a(t) = 2\cos(60\pi t) + 4\cos(100\pi t) + 10\sin(260\pi t)$ $+ 6\cos(460\pi t) + 3\sin(700\pi t)$

• Recall
$$\omega_o = \frac{2\pi\Omega_o}{\Omega_T}$$

Thus if Ω_T > 2Ω_o, then the corresponding normalized digital angular frequency ω_o of the discrete-time signal obtained by sampling the parent continuous-time sinusoidal signal will be in the range – π < ω < π
 No aliasing

- On the other hand, if $\Omega_T < 2\Omega_o$, the normalized digital angular frequency will foldover into a lower digital frequency $\omega_o = \langle 2\pi\Omega_o / \Omega_T \rangle_{2\pi}$ in the range $-\pi < \omega < \pi$ because of aliasing
- Hence, to prevent aliasing, the sampling frequency Ω_T should be greater than 2 times the frequency Ω_o of the sinusoidal signal being sampled

- Generalization: Consider an arbitrary continuous-time signal $x_a(t)$ composed of a weighted sum of a number of sinusoidal signals
- x_a(t) can be represented uniquely by its sampled version {x[n]} if the sampling frequency Ω_T is chosen to be greater than 2 times the highest frequency contained in x_a(t)

- The condition to be satisfied by the sampling frequency to prevent aliasing is called the **sampling theorem**
- A formal proof of this theorem will be presented later

Discrete-Time Systems

- A discrete-time system processes a given input sequence x[n] to generates an output sequence y[n] with more desirable properties
- In most applications, the discrete-time system is a single-input, single-output system:

$$x[n] \longrightarrow \begin{array}{c} \text{Discrete-time} \\ \text{System} \end{array} \longrightarrow y[n] \\ \text{Input sequence} \\ \end{array}$$

- 2-input, 1-output discrete-time systems -Modulator, adder
- 1-input, 1-output discrete-time systems -Multiplier, unit delay, unit advance



• Accumulator -

$$y[n] = \sum_{\ell = -\infty}^{n} x[\ell] = \sum_{\ell = -\infty}^{n-1} x[\ell] + x[n] = y[n-1] + x[n]$$

- The output y[n] at time instant n is the sum of the input sample x[n] at time instant n and the previous output y[n-1] at time instant n-1, which is the sum of all previous input sample values from -∞ to n-1
- The system cumulatively adds, i.e., it accumulates all input sample values

• Accumulator - Input-output relation can also be written in the form

$$y[n] = \sum_{\ell=-\infty}^{-1} x[\ell] + \sum_{\ell=0}^{n} x[\ell]$$

= $y[-1] + \sum_{\ell=0}^{n} x[\ell], n \ge 0$

• The second form is used for a causal input sequence, in which case y[-1] is called the initial condition

• M-point moving-average system -

$$y[n] = \frac{1}{M} \sum_{k=0}^{M-1} x[n-k]$$

- Used in smoothing random variations in data
- An application: Consider

x[n] = s[n] + d[n],

where s[n] is the signal corrupted by a noise d[n]

Discrete-Time Systems:Examples $s[n] = 2[n(0.9)^n], d[n] - random signal$



- Linear interpolation Employed to estimate sample values between pairs of adjacent sample values of a discrete-time sequence
- Factor-of-4 interpolation



• Factor-of-2 interpolator -

$$y[n] = x_u[n] + \frac{1}{2}(x_u[n-1] + x_u[n+1])$$

• Factor-of-3 interpolator -

$$y[n] = x_u[n] + \frac{1}{3}(x_u[n-1] + x_u[n+2]) + \frac{2}{3}(x_u[n-2] + x_u[n+1])$$

Discrete-Time Systems: Classification

- Linear System
- Shift-Invariant System
- Causal System
- Stable System
- Passive and Lossless Systems

Linear Discrete-Time Systems

Definition - If y₁[n] is the output due to an input x₁[n] and y₂[n] is the output due to an input x₂[n] then for an input

 $x[n] = \alpha x_1[n] + \beta x_2[n]$ the output is given by $y[n] = \alpha y_1[n] + \beta y_2[n]$

• Above property must hold for any arbitrary constants α and β , and for all possible inputs $x_1[n]$ and $x_2[n]$

<u>Accumulator</u>: Linear Discrete-Time System?

- Accumulator $y_1[n] = \sum_{\ell=-\infty}^n x_1[\ell], \quad y_2[n] = \sum_{\ell=-\infty}^n x_2[\ell]$
- For an input

 $x[n] = \alpha x_1[n] + \beta x_2[n]$

the output is

$$y[n] = \sum_{\ell=-\infty}^{n} (\alpha x_1[\ell] + \beta x_2[\ell])$$

= $\alpha \sum_{\ell=-\infty}^{n} x_1[\ell] + \beta \sum_{\ell=-\infty}^{n} x_2[\ell] = \alpha y_1[n] + \beta y_2[n]$

• Hence, the above system is **linear**
<u>Causal Accumulator</u>: Linear Discrete-Time System?

- The outputs $y_1[n]$ and $y_2[n]$ for inputs $x_1[n]$ and $x_2[n]$ are given by $y_1[n] = y_1[-1] + \sum_{\ell=0}^n x_1[\ell]$ $y_2[n] = y_2[-1] + \sum_{\ell=0}^n x_2[\ell]$
- The output y[n] for an input $\alpha x_1[n] + \beta x_2[n]$ is given by

$$y[n] = y[-1] + \sum_{\ell=0}^{n} (\alpha x_1[\ell] + \beta x_2[\ell])$$

<u>Causal Accumulator</u> cont.: Linear Discrete-Time System?

- Now $\alpha y_1[n] + \beta y_2[n]$ = $\alpha (y_1[-1] + \sum_{\ell=0}^n x_1[\ell]) + \beta (y_2[-1] + \sum_{\ell=0}^n x_2[\ell])$ = $(\alpha y_1[-1] + \beta y_2[-1]) + (\alpha \sum_{\ell=0}^n x_1[\ell] + \beta \sum_{\ell=0}^n x_2[\ell])$
- Thus $y[n] = \alpha y_1[n] + \beta y_2[n]$ if $y[-1] = \alpha y_1[-1] + \beta y_2[-1]$

<u>Causal Accumulator cont.</u>: Linear Discrete-Time System?

- For the causal accumulator to be linear the condition $y[-1] = \alpha y_1[-1] + \beta y_2[-1]$ must hold for all initial conditions y[-1], $y_1[-1], y_2[-1]$, and all constants α and β
- This condition cannot be satisfied unless the accumulator is initially at rest with zero initial condition
- For nonzero initial condition, the system is **nonlinear**

A Nonlinear Discrete-Time System

• Consider

$$y[n] = x^{2}[n] - x[n-1]x[n+1]$$

 Outputs y₁[n] and y₂[n] for inputs x₁[n] and x₂[n] are given by

$$y_1[n] = x_1^2[n] - x_1[n-1]x_1[n+1]$$
$$y_2[n] = x_2^2[n] - x_2[n-1]x_2[n+1]$$

A Nonlinear Discrete-Time System cont.

• Output y[n] due to an input $\alpha x_1[n] + \beta x_2[n]$ is given by

$$y[n] = \{\alpha x_1[n] + \beta x_2[n]\}^2$$

 $-\{\alpha x_{1}[n-1]+\beta x_{2}[n-1]\}\{\alpha x_{1}[n+1]+\beta x_{2}[n+1]\}$

$$= \alpha^{2} \{ x_{1}^{2}[n] - x_{1}[n-1]x_{1}[n+1] \}$$

$$-\beta^{2} \{x_{2}^{2}[n] - x_{2}[n-1]x_{2}[n+1]\}$$

 $+\alpha\beta\{2x_1[n]x_2[n]-x_1[n-1]x_2[n+1]-x_1[n+1]x_2[n-1]\}$

A Nonlinear Discrete-Time System cont.

- On the other hand $\alpha y_1[n] + \beta y_2[n]$ $= \alpha \{x_1^2[n] - x_1[n-1]x_1[n+1]\}$ $+ \beta \{x_2^2[n] - x_2[n-1]x_2[n+1]\}$ $\neq y[n]$
- Hence, the system is **nonlinear**

Shift (Time)-Invariant System

- For a shift-invariant system, if y₁[n] is the response to an input x₁[n], then the response to an input x[n] = x₁[n n_o]
 is simply y[n] = y₁[n n_o]
 - where n_o is any positive or negative integer
- The above relation must hold for any arbitrary input and its corresponding output
- If *n* is discrete time, the above property is called **time-invariance** property

<u>Up-Sampler</u>: Shift-Invariant System?

• <u>Example</u> - Consider the up-sampler with an input-output relation given by

$$x_u[n] = \begin{cases} x[n/L], & n = 0, \pm L, \pm 2L, \dots \\ 0, & \text{otherwise} \end{cases}$$

• For an input $x_1[n] = x[n - n_o]$ the output $x_{1,u}[n]$ is given by

$$x_{1,u}[n] = \begin{cases} x_1[n/L], & n = 0, \pm L, \pm 2L, \dots \\ 0, & \text{otherwise} \end{cases}$$
$$= \begin{cases} x[(n-Ln_o)/L], & n = 0, \pm L, \pm 2L, \dots \\ 0, & \text{otherwise} \end{cases}$$

<u>Up-Sampler</u>: Shift-Invariant System?

• However from the definition of the up-sampler

$$x_{u}[n-n_{o}]$$

$$=\begin{cases} x[(n-n_{o})/L], & n=n_{o}, n_{o} \pm L, n_{o} \pm 2L, \dots \\ 0, & \text{otherwise} \end{cases}$$

$$\neq x_{1,u}[n]$$

• Hence, the up-sampler is a time-varying system

Linear Time-Invariant System

- Linear Time-Invariant (LTI) System -A system satisfying both the linearity and the time-invariance property
- LTI systems are mathematically easy to analyze and characterize, and consequently, easy to design
- Highly useful signal processing algorithms have been developed utilizing this class of systems over the last several decades

- In a **causal system**, the n_o -th output sample $y[n_o]$ depends only on input samples x[n] for $n \le n_o$ and does not depend on input samples for $n \ge n_o$
- Let y₁[n] and y₂[n] be the responses of a causal discrete-time system to the inputs x₁[n] and x₂[n], respectively

• Then

 $x_1[n] = x_2[n]$ for n < N

implies also that

 $y_1[n] = y_2[n]$ for n < N

• For a causal system, changes in output samples do not precede changes in the input samples

- Examples of causal systems:
 - $y[n] = \alpha_1 x[n] + \alpha_2 x[n-1] + \alpha_3 x[n-2] + \alpha_4 x[n-3]$ $y[n] = b_0 x[n] + b_1 x[n-1] + b_2 x[n-2]$ $+ a_1 y[n-1] + a_2 y[n-2]$ y[n] = y[n-1] + x[n]

• Examples of noncausal systems: $y[n] = x_u[n] + \frac{1}{2}(x_u[n-1] + x_u[n+1])$

- A noncausal system can be implemented as a causal system by delaying the output by an appropriate number of samples
- For example a causal implementation of the factor-of-2 interpolator is given by

$$y[n] = x_u[n-1] + \frac{1}{2}(x_u[n-2] + x_u[n])$$

Stable System

- There are various definitions of stability
- We consider here the **bounded-input**, **bounded-output (BIBO) stability**
- If y[n] is the response to an input x[n] and if $|x[n]| \le B_x$ for all values of n

then

 $|y[n]| \le B_y$ for all values of n

Stable System

• <u>Example</u> - The *M*-point moving average filter is BIBO stable:

$$y[n] = \frac{1}{M} \sum_{k=0}^{M-1} x[n-k]$$

• For a bounded input $|x[n]| \le B_x$ we have

$$|y[n]| = \left|\frac{1}{M}\sum_{k=0}^{M-1} x[n-k]\right| \le \frac{1}{M}\sum_{k=0}^{M-1} |x[n-k]|$$
$$\le \frac{1}{M}(MB_x) \le B_x$$

Passive and Lossless Systems

A discrete-time system is defined to be passive if, for every finite-energy input x[n], the output y[n] has, at most, the same energy, i.e.

$$\sum_{n=-\infty}^{\infty} |y[n]|^2 \le \sum_{n=-\infty}^{\infty} |x[n]|^2 < \infty$$

• For a **lossless** system, the above inequality is satisfied with an equal sign for every input

Passive and Lossless Systems

- Example Consider the discrete-time system defined by $y[n] = \alpha x[n - N]$ with N a positive integer
- Its output energy is given by

$$\sum_{n=-\infty}^{\infty} |y[n]|^2 \le |\alpha|^2 \sum_{n=-\infty}^{\infty} |x[n]|^2$$

• Hence, it is a passive system if $|\alpha| \le 1$ and is a lossless system if $|\alpha| = 1$

Impulse and Step Responses

- The response of a discrete-time system to a unit sample sequence {δ[n]} is called the unit sample response or simply, the impulse response, and is denoted by {h[n]}
- The response of a discrete-time system to a unit step sequence {µ[n]} is called the unit step response or simply, the step response, and is denoted by {s[n]}

Impulse Response

- <u>Example</u> The impulse response of the system
- $y[n] = \alpha_1 x[n] + \alpha_2 x[n-1] + \alpha_3 x[n-2] + \alpha_4 x[n-3]$ is obtained by setting $x[n] = \delta[n]$ resulting in

$$h[n] = \alpha_1 \delta[n] + \alpha_2 \delta[n-1] + \alpha_3 \delta[n-2] + \alpha_4 \delta[n-3]$$

• The impulse response is thus a finite-length sequence of length 4 given by

$$\{h[n]\} = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$$

Impulse Response

• <u>Example</u> - The impulse response of the discrete-time accumulator

$$y[n] = \sum_{\ell = -\infty}^{n} x[\ell]$$

is obtained by setting $x[n] = \delta[n]$ resulting in

$$h[n] = \sum_{\ell=-\infty}^{n} \delta[\ell] = \mu[n]$$

Impulse Response

- Example The impulse response $\{h[n]\}$ of the factor-of-2 interpolator $y[n] = x_u[n] + \frac{1}{2}(x_u[n-1] + x_u[n+1])$ is obtained by setting $x_u[n] = \delta[n]$ and is given by $h[n] = \delta[n] + \frac{1}{2}(\delta[n-1] + \delta[n+1])$
- The impulse response is thus a finite-length sequence of length 3:

$${h[n]} = {0.5, 1 0.5}$$

• Input-Output Relationship -

It can be shown that a consequence of the linear, time-invariance property is that an LTI discrete-time system is completely characterized by its impulse response

• Knowing the impulse response one can compute the output of the system for any arbitrary input

- Let *h*[*n*] denote the impulse response of a LTI discrete-time system
- We compute its output y[n] for the input: $x[n] = 0.5\delta[n+2] + 1.5\delta[n-1] - \delta[n-2] + 0.75\delta[n-5]$
 - As the system is linear, we can compute its outputs for each member of the input separately and add the individual outputs to determine y[n]

• Since the system is time-invariant

input output $\delta[n+2] \rightarrow h[n+2]$ $\delta[n-1] \rightarrow h[n-1]$ $\delta[n-2] \rightarrow h[n-2]$ $\delta[n-5] \rightarrow h[n-5]$

- Likewise, as the system is linear input output $0.5\delta[n+2] \rightarrow 0.5h[n+2]$ $1.5\delta[n-1] \rightarrow 1.5h[n-1]$
 - $-\delta[n-2] \rightarrow -h[n-2]$
 - $0.75\delta[n-5] \rightarrow 0.75h[n-5]$
- Hence because of the linearity property we get y[n] = 0.5h[n+2]+1.5h[n-1] -h[n-2]+0.75h[n-5]

 Now, any arbitrary input sequence x[n] can be expressed as a linear combination of delayed and advanced unit sample sequences in the form

$$x[n] = \sum_{k=-\infty}^{\infty} x[k] \delta[n-k]$$

• The response of the LTI system to an input $x[k]\delta[n-k]$ will be x[k]h[n-k]

• Hence, the response *y*[*n*] to an input

$$x[n] = \sum_{k=-\infty}^{\infty} x[k] \delta[n-k]$$

will be

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k]$$

which can be alternately written as

$$y[n] = \sum_{k=-\infty}^{\infty} x[n-k]h[k]$$

Convolution Sum

• The summation

 $y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k] = \sum_{k=-\infty}^{\infty} x[n-k]h[n]$ is called the **convolution sum** of the sequences x[n] and h[n] and represented compactly as

 $y[n] = x[n] \circledast h[n]$

Convolution Sum

- Properties -
- Commutative property:

 $x[n] \circledast h[n] = h[n] \circledast x[n]$

- Associative property : (x[n] * h[n]) y[n] = x[n] (h[n] y[n])
- Distributive property : $x[n] \circledast (h[n] + y[n]) = x[n] \circledast h[n] + x[n] \circledast y[n]$

Simple Interconnection Schemes

- Two simple interconnection schemes are:
- Cascade Connection
- Parallel Connection

 Impulse response h[n] of the cascade of two LTI discrete-time systems with impulse responses h₁[n] and h₂[n] is given by

 $h[n] = h_1[n] \circledast h_2[n]$

- <u>Note</u>: The ordering of the systems in the cascade has no effect on the overall impulse response because of the commutative property of convolution
- A cascade connection of two stable systems is stable
- A cascade connection of two passive (lossless) systems is passive (lossless)

- An application is in the development of an **inverse system**
- If the cascade connection satisfies the relation

 $h_1[n] \circledast h_2[n] = \delta[n]$

then the LTI system $h_1[n]$ is said to be the inverse of $h_2[n]$ and vice-versa

- An application of the inverse system concept is in the recovery of a signal x[n] from its distorted version x̂[n] appearing at the output of a transmission channel
- If the impulse response of the channel is known, then x[n] can be recovered by designing an inverse system of the channel

$$x[n] \longrightarrow \begin{array}{c} \text{channel} & \text{inverse system} \\ \hline k_1[n] & \hline k_2[n] & \hline h_2[n] \\ \hline h_1[n] & \hline h_2[n] & = \delta[n] \end{array}$$

- Example Consider the discrete-time accumulator with an impulse response μ[n]
- Its inverse system satisfy the condition $\mu[n] \circledast h_2[n] = \delta[n]$
- It follows from the above that $h_2[n] = 0$ for n < 0 and

$$h_2[1] = 1$$

 $\sum_{\ell=0}^{n} h_2[\ell] = 0 \text{ for } n \ge 2$
Cascade Connection

• Thus the impulse response of the inverse system of the discrete-time accumulator is given by

$$h_2[n] = \delta[n] - \delta[n-1]$$

which is called a **backward difference** system

Parallel Connection



Impulse response h[n] of the parallel connection of two LTI discrete-time systems with impulse responses h₁[n] and h₂[n] is given by

 $h[n] = h_1[n] + h_2[n]$

• Consider the discrete-time system where $h_1[n] = \delta[n] + 0.5\delta[n-1],$ $h_2[n] = 0.5\delta[n] - 0.25\delta[n-1],$ $h_{3}[n] = 2\delta[n],$ $h_1[n]$ $h_4[n] = -2(0.5)^n \mu[n]$ $h_2[n]$ $h_3[n]$ $h_4[n]$

• Simplifying the block-diagram we obtain



- Overall impulse response h[n] is given by $h[n] = h_1[n] + h_2[n] \circledast (h_3[n] + h_4[n])$ $= h_1[n] + h_2[n] \circledast h_3[n] + h_2[n] \circledast h_4[n]$
- Now,

$$h_{2}[n] \circledast h_{3}[n] = (\frac{1}{2}\delta[n] - \frac{1}{4}\delta[n-1]) \circledast 2\delta[n]$$
$$= \delta[n] - \frac{1}{2}\delta[n-1]$$

$$h_{2}[n] \circledast h_{4}[n] = (\frac{1}{2}\delta[n] - \frac{1}{4}\delta[n-1]) \circledast \left(-2(\frac{1}{2})^{n}\mu[n]\right)$$
$$= -(\frac{1}{2})^{n}\mu[n] + \frac{1}{2}(\frac{1}{2})^{n-1}\mu[n-1]$$
$$= -(\frac{1}{2})^{n}\mu[n] + (\frac{1}{2})^{n}\mu[n-1]$$
$$= -(\frac{1}{2})^{n}\delta[n] = -\delta[n]$$
• Therefore

 $h[n] = \delta[n] + \frac{1}{2}\delta[n-1] + \delta[n] - \frac{1}{2}\delta[n-1] - \delta[n] = \delta[n]$

- BIBO Stability Condition A discretetime is BIBO stable if the output sequence {y[n]} remains bounded for all bounded input sequence {x[n]}
- An LTI discrete-time system is BIBO stable if and only if its impulse response sequence {h[n]} is absolutely summable, i.e.

$$\mathbf{S} = \sum_{n = -\infty}^{\infty} |h[n]| < \infty$$

- <u>Proof</u>: Assume *h*[*n*] is a real sequence
- Since the input sequence x[n] is bounded we have

$$|x[n]| \le B_{\chi} < \infty$$

• Therefore

 $\left| y[n] \right| = \left| \sum_{k=-\infty}^{\infty} h[k] x[n-k] \right| \le \sum_{k=-\infty}^{\infty} \left| h[k] \right| \left| x[n-k] \right|$

$$\leq B_x \sum_{k=-\infty}^{\infty} \left| h[k] \right| = B_x S$$

- Thus, $S < \infty$ implies $|y[n]| \le B_y < \infty$ indicating that y[n] is also bounded
- To prove the converse, assume y[n] is bounded, i.e., $|y[n]| \le B_y$
- Consider the input given by

$$x[n] = \begin{cases} \operatorname{sgn}(h[-n]), & \text{if } h[-n] \neq 0\\ K, & \text{if } h[-n] = 0 \end{cases}$$

where sgn(c) = +1 if c > 0 and sgn(c) = -1if c < 0 and $|K| \le 1$

- Note: Since $|x[n]| \le 1$, $\{x[n]\}$ is obviously bounded
- For this input, y[n] at n = 0 is

$$y[0] = \sum_{k=-\infty}^{\infty} \operatorname{sgn}(h[k])h[k] = S \le B_{y} < \infty$$

• Therefore, $|y[n]| \le B_y$ implies $S < \infty$

• <u>Example</u> - Consider a causal LTI discrete-time system with an impulse response

$$h[n] = (\alpha)^n \, \mu[n]$$

• For this system

$$S = \sum_{n=-\infty}^{\infty} \left| \alpha^n \right| \mu[n] = \sum_{n=0}^{\infty} \left| \alpha \right|^n = \frac{1}{1 - \left| \alpha \right|}, \left| \alpha \right| < 1$$

- Therefore $S < \infty$ if $|\alpha| < 1$ for which the system is BIBO stable
- If $|\alpha| = 1$, the system is not BIBO stable

• Let $x_1[n]$ and $x_2[n]$ be two input sequences with

$$x_1[n] = x_2[n] \quad \text{for} \quad n \le n_o$$

The corresponding output samples at n = n_o of an LTI system with an impulse response {h[n]} are then given by

Causality Condition of an LTI **Discrete-Time System** $y_1[n_o] = \sum_{k=0}^{\infty} h[k] x_1[n_o - k] = \sum_{k=0}^{\infty} h[k] x_1[n_o - k]$ k=0 $k = -\infty$ + $\sum h[k]x_1[n_0 - k]$ $k = -\infty$ ∞ $y_2[n_o] = \sum h[k] x_2[n_o - k] = \sum h[k] x_2[n_o - k]$ $k = -\infty$ k=0+ $\sum h[k] x_2[n_0 - k]$ $k = -\infty$

- If the LTI system is also causal, then $y_1[n_o] = y_2[n_o]$
- As $x_1[n] = x_2[n]$ for $n \le n_o$ $\sum_{k=0}^{\infty} h[k]x_1[n_o - k] = \sum_{k=0}^{\infty} h[k]x_2[n_o - k]$
- This implies $\sum_{k=-\infty}^{-1} h[k] x_1[n_o - k] = \sum_{k=-\infty}^{-1} h[k] x_2[n_o - k]$

• As $x_1[n] \neq x_2[n]$ for $n > n_o$ the only way the condition

$$\sum_{k=-\infty}^{-1} h[k] x_1[n_o - k] = \sum_{k=-\infty}^{-1} h[k] x_2[n_o - k]$$

will hold if both sums are equal to zero, which is satisfied if

 $h[k] = 0 \quad \text{for } k < 0$

- An LTI discrete-time system is **causal** if and only if its impulse response {*h*[*n*]} is a causal sequence
- <u>Example</u> The discrete-time system defined by

 $y[n] = \alpha_1 x[n] + \alpha_2 x[n-1] + \alpha_3 x[n-2] + \alpha_4 x[n-3]$ is a causal system as it has a causal impulse response $\{h[n]\} = \{\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4\}$

• <u>Example</u> - The discrete-time accumulator defined by

$$y[n] = \sum_{\ell=-\infty}^{n} \delta[\ell] = \mu[n]$$

is a causal system as it has a causal impulse response given by

$$h[n] = \sum_{\ell=-\infty}^{n} \delta[\ell] = \mu[n]$$

• <u>Example</u> - The factor-of-2 interpolator defined by

 $y[n] = x_u[n] + \frac{1}{2}(x_u[n-1] + x_u[n+1])$ is noncausal as it has a noncausal impulse response given by

$${h[n]} = {0.5 \ 1 \ 0.5}$$

- Note: A noncausal LTI discrete-time system with a finite-length impulse response can often be realized as a causal system by inserting an appropriate amount of delay
- For example, a causal version of the factorof-2 interpolator is obtained by delaying the input by one sample period:

$$y[n] = x_u[n-1] + \frac{1}{2}(x_u[n-2] + x_u[n])$$

Finite-Dimensional LTI Discrete-Time Systems

• An important subclass of LTI discrete-time systems is characterized by a linear constant coefficient difference equation of the form

$$\sum_{k=0}^{N} d_k y[n-k] = \sum_{k=0}^{M} p_k x[n-k]$$

- *x*[*n*] and *y*[*n*] are, respectively, the input and the output of the system
- $\{d_k\}$ and $\{p_k\}$ are constants characterizing the system

Finite-Dimensional LTI Discrete-Time Systems

- The **order** of the system is given by max(*N*,*M*), which is the order of the difference equation
- It is possible to implement an LTI system characterized by a constant coefficient difference equation as here the computation involves two finite sums of products

Finite-Dimensional LTI Discrete-Time Systems

• If we assume the system to be causal, then the output *y*[*n*] can be recursively computed using

$$y[n] = -\sum_{k=1}^{N} \frac{d_k}{d_0} y[n-k] + \sum_{k=1}^{M} \frac{p_k}{d_0} x[n-k]$$

rovided $d_0 \neq 0$

provided $d_0 \neq 0$

• y[n] can be computed for all $n \ge n_o$, knowing x[n] and the initial conditions $y[n_o -1], y[n_o -2], ..., y[n_o -N]$

Based on Impulse Response Length -

• If the impulse response *h*[*n*] is of finite length, i.e.,

h[n] = 0 for $n < N_1$ and $n > N_2$, $N_1 < N_2$

then it is known as a **finite impulse response** (FIR) discrete-time system

• The convolution sum description here is $y[n] = \sum_{k=N_1}^{N_2} h[k]x[n-k]$

- The output *y*[*n*] of an FIR LTI discrete-time system can be computed directly from the convolution sum as it is a finite sum of products
- Examples of FIR LTI discrete-time systems are the moving-average system and the linear interpolators

- If the impulse response is of infinite length, then it is known as an infinite impulse response (IIR) discrete-time system
- The class of IIR systems we are concerned with in this course are characterized by linear constant coefficient difference equations

• <u>Example</u> - The discrete-time accumulator defined by

y[n] = y[n-1] + x[n]is seen to be an IIR system

• <u>Example</u> - The familiar numerical integration formulas that are used to numerically solve integrals of the form *t*

$$y(t) = \int_{0}^{t} x(\tau) d\tau$$

can be shown to be characterized by linear constant coefficient difference equations, and hence, are examples of IIR systems

• If we divide the interval of integration into *n* equal parts of length *T*, then the previous integral can be rewritten as

$$y(nT) = y((n-1)T) + \int_{(n-1)T}^{nT} x(\tau)d\tau$$

where we have set t = nT and used the notation T

$$y(nT) = \int_{0}^{nT} x(\tau) d\tau$$

- Using the trapezoidal method we can write $\int_{nT} x(\tau) d\tau = \frac{T}{2} \{ x((n-1)T) + x(nT) \}$ (n-1)T
- Hence, a numerical representation of the definite integral is given by

$$y(nT) = y((n-1)T) + \frac{T}{2} \{x((n-1)T) + x(nT)\}$$

- Let y[n] = y(nT) and x[n] = x(nT)
- Then

 $y(nT) = y((n-1)T) + \frac{T}{2} \{x((n-1)T) + x(nT)\}$ reduces to

$$y[n] = y[n-1] + \frac{T}{2} \{x[n] + x[n-1]\}$$

which is recognized as the difference equation representation of a first-order IIR discrete-time system

Based on the Output Calculation Process

- Nonrecursive System Here the output can be calculated sequentially, knowing only the present and past input samples
- Recursive System Here the output computation involves past output samples in addition to the present and past input samples

Based on the Coefficients -

- **Real Discrete-Time System** The impulse response samples are real valued
- **Complex Discrete-Time System -** The impulse response samples are complex valued

Definitions

- A measure of similarity between a pair of energy signals, x[n] and y[n], is given by the cross-correlation sequence $r_{xy}[\ell]$ defined by $r_{xy}[\ell] = \sum_{n=-\infty}^{\infty} x[n]y[n-\ell], \quad \ell = 0, \pm 1, \pm 2, ...$
- The parameter ℓ called lag, indicates the time-shift between the pair of signals

If y[n] is made the reference signal and we wish to shift x[n] with respect to y[n], then the corresponding cross-correlation sequence is given by

$$r_{yx}[\ell] = \sum_{n=-\infty}^{\infty} y[n]x[n-\ell]$$

$$= \sum_{m=-\infty}^{\infty} y[m+\ell] x[m] = r_{xy}[-\ell]$$

• Thus, $r_{yx}[\ell]$ is obtained by time-reversing $r_{xy}[\ell]$

• The autocorrelation sequence of *x*[*n*] is given by

 $r_{xx}[\ell] = \sum_{n=-\infty}^{\infty} x[n]x[n-\ell]$

obtained by setting y[n] = x[n] in the definition of the cross-correlation sequence $r_{xy}[\ell]$

• Note: $r_{xx}[0] = \sum_{n=-\infty}^{\infty} x^2[n] = E_x$, the energy of the signal x[n]

- From the relation $r_{yx}[\ell] = r_{xy}[-\ell]$ it follows that $r_{xx}[\ell] = r_{xx}[-\ell]$ implying that $r_{xx}[\ell]$ is an even function for real x[n]
- An examination of

$$r_{xy}[\ell] = \sum_{n=-\infty}^{\infty} x[n]y[n-\ell]$$

reveals that the expression for the crosscorrelation looks quite similar to that of the linear convolution
Correlation of Signals

- This similarity is much clearer if we rewrite the expression for the cross-correlation as $r_{xy}[\ell] = \sum_{n=-\infty}^{\infty} x[n]y[-(\ell - n)] = x[\ell] * y[-\ell]$
- The cross-correlation of y[n] with the reference signal x[n] can be computed by processing x[n] with an LTI discrete-time system of impulse response y[-n]

$$x[n] \longrightarrow y[-n] \longrightarrow r_{xy}[n]$$

Correlation of Signals

Likewise, the autocorrelation of x[n] can be computed by processing x[n] with an LTI discrete-time system of impulse response x[-n]

$$x[n] \longrightarrow x[-n] \longrightarrow r_{xx}[n]$$

- Consider two finite-energy sequences x[n] and y[n]
- The energy of the combined sequence $ax[n] + y[n \ell]$ is also finite and nonnegative, i.e.,

$$\sum_{n=-\infty}^{\infty} (a x[n] + y[n-\ell])^2 = a^2 \sum_{n=-\infty}^{\infty} x^2[n]$$
$$+ 2a \sum_{n=-\infty}^{\infty} x[n] y[n-\ell] + \sum_{n=-\infty}^{\infty} y^2[n-\ell] \ge 0$$

• Thus

$$a^{2}r_{xx}[0] + 2ar_{xy}[\ell] + r_{yy}[0] \ge 0$$

where $r_{xx}[0] = E_{x} > 0$ and $r_{yy}[0] = E_{y} > 0$

• We can rewrite the equation on the previous slide as

$$\begin{bmatrix} a & 1 \end{bmatrix} \begin{bmatrix} r_{xx}[0] & r_{xy}[\ell] \\ r_{xy}[\ell] & r_{yy}[0] \end{bmatrix} \begin{bmatrix} a \\ 1 \end{bmatrix} \ge 0$$

for any finite value of a

• Or, in other words, the matrix $\begin{bmatrix} r_{xx}[0] & r_{xy}[\ell] \\ r_{xy}[\ell] & r_{yy}[0] \end{bmatrix}$

• • $r_{xx}[0]r_{yy}[0] - r_{xy}^{2}[\ell] \ge 0$ $|r_{xy}[\ell]| \le \sqrt{r_{xx}[0]r_{yy}[0]} = \sqrt{E_{x}E_{y}}$

- The last inequality on the previous slide provides an upper bound for the crosscorrelation samples
- If we set y[n] = x[n], then the inequality reduces to

$$|r_{xx}[\ell]| \leq r_{xx}[0] = \mathbf{E}_x$$

- Thus, at zero lag (l = 0), the sample value of the autocorrelation sequence has its maximum value
- Now consider the case

 $y[n] = \pm b \, x[n-N]$

where *N* is an integer and b > 0 is an arbitrary number

• In this case $E_y = b^2 E_x$

• Therefore

$$\sqrt{\mathbf{E}_{x}\mathbf{E}_{y}} = \sqrt{b^{2}\mathbf{E}_{x}^{2}} = b\mathbf{E}_{x}$$

• Using the above result in $|r_{xy}[\ell]| \le \sqrt{r_{xx}[0]r_{yy}[0]} = \sqrt{E_xE_y}$ we get

$$-br_{xx}[0] \le r_{xy}[\ell] \le br_{xx}[0]$$

- The cross-correlation and autocorrelation sequences can easily be computed using MATLAB
- <u>Example</u> Consider the two finite-length sequences

$$x[n] = \begin{bmatrix} 1 & 3 & -2 & 1 & 2 & -1 & 4 & 4 & 2 \end{bmatrix}$$
$$y[n] = \begin{bmatrix} 2 & -1 & 4 & 1 & -2 & 3 \end{bmatrix}$$

• The cross-correlation sequence $r_{xy}[n]$ computed using Program 2_7 of text is plotted below



- The autocorrelation sequence $r_{xx}[\ell]$ computed using Program 2_7 is shown below
- Note: At zero lag, $r_{xx}[0]$ is the maximum



- The plot below shows the cross-correlation of x[n] and y[n] = x[n-N] for N = 4
- Note: The peak of the cross-correlation is precisely the value of the delay *N*



- The plot below shows the autocorrelation of x[n] corrupted with an additive random noise generated using the function randn
- Note: The autocorrelation still exhibits a peak at zero lag



- The autocorrelation and the crosscorrelation can also be computed using the function xcorr
- However, the correlation sequences generated using this function are the time-reversed version of those generated using Programs 2_7 and 2_8

Normalized Forms of Correlation

• Normalized forms of autocorrelation and cross-correlation are given by

$$\rho_{xx}[\ell] = \frac{r_{xx}[\ell]}{r_{xx}[0]}, \quad \rho_{xy}[\ell] = \frac{r_{xy}[\ell]}{\sqrt{r_{xx}[0]r_{yy}[0]}}$$

- They are often used for convenience in comparing and displaying
- Note: $|\rho_{xx}[\ell]| \le 1$ and $|\rho_{xy}[\ell]| \le 1$ independent of the range of values of x[n]and y[n]

- The cross-correlation sequence for a pair of power signals, x[n] and y[n], is defined as $r_{xy}[\ell] = \lim_{K \to \infty} \frac{1}{2K+1} \sum_{n=-K}^{K} x[n]y[n-\ell]$
- The autocorrelation sequence of a power signal *x*[*n*] is given by

$$r_{xx}[\ell] = \lim_{K \to \infty} \frac{1}{2K+1} \sum_{n=-K}^{K} x[n]x[n-\ell]$$

The cross-correlation sequence for a pair of periodic signals of period N, x̃[n] and ỹ[n], is defined as

$$r_{\widetilde{x}\widetilde{y}}[\ell] = \frac{1}{N} \sum_{n=0}^{N-1} \widetilde{x}[n] \widetilde{y}[n-\ell]$$

• The autocorrelation sequence of a periodic signal $\tilde{x}[n]$ of period N is given by

$$r_{\widetilde{x}\widetilde{x}}[\ell] = \frac{1}{N} \sum_{n=0}^{N-1} \widetilde{x}[n] \widetilde{x}[n-\ell]$$

- Note: Both $r_{\tilde{x}\tilde{y}}[\ell]$ and $r_{\tilde{x}\tilde{x}}[\ell]$ are also periodic signals with a period *N*
- The periodicity property of the autocorrelation sequence can be exploited to determine the period of a periodic signal that may have been corrupted by an additive random disturbance

Let x̃[n] be a periodic signal corrupted by the random noise d[n] resulting in the signal w[n] = x̃[n] + d[n]
which is observed for 0 ≤ n ≤ M −1 where M ≫ N

• The autocorrelation of w[n] is given by $r_{ww}[\ell] = \frac{1}{M} \sum_{n=0}^{M-1} w[n]w[n-\ell]$ $= \frac{1}{M} \sum_{n=0}^{M-1} (\tilde{x}[n] + d[n]) (\tilde{x}[n-\ell] + d[n-\ell])$ $= \frac{1}{M} \sum_{n=0}^{M-1} \tilde{x}[n] \tilde{x}[n-\ell] + \frac{1}{M} \sum_{n=0}^{M-1} d[n] d[n-\ell]$ $+\frac{1}{M}\sum_{n=0}^{M-1}\tilde{x}[n]d[n-\ell] + \frac{1}{M}\sum_{n=0}^{M-1}d[n]\tilde{x}[n-\ell]$ $= r_{\tilde{\gamma}\tilde{\gamma}}[\ell] + r_{dd}[\ell] + r_{\tilde{\gamma}d}[\ell] + r_{d\tilde{\gamma}}[\ell]$

- In the last equation on the previous slide, $r_{\tilde{\chi}\tilde{\chi}}[\ell]$ is a periodic sequence with a period *N* and hence will have peaks at $\ell = 0, N, 2N, ...$ with the same amplitudes as ℓ approaches *M*
- As $\tilde{x}[n]$ and d[n] are not correlated, samples of cross-correlation sequences $r_{\tilde{x}d}[\ell]$ and $r_{d\tilde{x}}[\ell]$ are likely to be very small relative to the amplitudes of $r_{\tilde{x}\tilde{x}}[\ell]$

- The autocorrelation $r_{dd}[\ell]$ of d[n] will show a peak at $\ell = 0$ with other samples having rapidly decreasing amplitudes with increasing values of $|\ell|$
- Hence, peaks of r_{ww}[l] for l > 0 are essentially due to the peaks of r_{xx}[l] and can be used to determine whether x[n] is a periodic sequence and also its period N if the peaks occur at periodic intervals

Correlation Computation of a Periodic Signal Using MATLAB

- Example We determine the period of the sinusoidal sequence x[n] = cos(0.25n),
 0 ≤ n ≤ 95 corrupted by an additive uniformly distributed random noise of amplitude in the range [-0.5, 0.5]
- Using Program 2_8 of text we arrive at the plot of $r_{ww}[\ell]$ shown on the next slide

Correlation Computation of a Periodic Signal Using MATLAB



- As can be seen from the plot given above, there is a strong peak at zero lag
- However, there are distinct peaks at lags that are multiples of 8 indicating the period of the sinusoidal sequence to be 8 as expected

Correlation Computation of a Periodic Signal Using MATLAB

• Figure below shows the plot of $r_{dd}[\ell]$



• As can be seen $r_{dd}[\ell]$ shows a very strong peak at only zero lag