Q1. Discuss, compare and contrast various curve fitting and interpolation methods

Curve Fitting

- Problem statement: Given a set of $(n+1)$ point-pairs $\{x_i, y_i\}$, $i = 0, 1, \ldots n$, find an analytic, smooth curve in the interval $[x_0, x_n]$.
- Why we perform curve fitting?
  - To get estimates at some intermediate points
  - To produce a simplified version of a more complicated function
- Methods:
  - Interpolation for clean data:
    * Lagrange
    * Newton's divided-difference
    * Splines
  - Regression for noisy data:
    * Linear, Polynomial ...

Interpolation: Direct Approach

- To fit exactly $(n+1)$ data points, use the polynomial of degree n:

$$P_n = c_0 + c_1 x \ldots + c_n x^n$$

- Find c_i by solving the linear system of equations:

$$\begin{pmatrix} 1 & x_0 & \ldots & x_0^n \\ 1 & x_1 & \ldots & x_1^n \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & \ldots & x_n^n \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_n \end{pmatrix}$$

- Caution: Not advisable to solve this system owing to the matrix-inversion!

- Do we have any inversion-free methods ?

Lagrange Interpolation

- An n-th degree Lagrange basis polynomial:

$$\phi_i(x) = \frac{\prod_{j=0,j\neq i}^{n}(x - x_j)}{\prod_{j=0,j\neq i}^{n}(x_i - x_j)} \quad i = 0, 1, \ldots n.$$

- Hence the Lagrange's interpolating polynomial is

$$P_n(x) = \sum_{i=0}^{n} c_i \phi_i(x)$$

- $\phi_i(x)$ has the property:

$$\phi_i(x_k) = \begin{cases} 1, & i = k; \\ 0, & i \neq k. \end{cases}$$

Lagrange (Cont'd)

- Using the above property, we get the coefficients

$$c_i = y_i$$

hence much simpler to find coefficients!

- Limitations:
  - Redo the whole procedure when adding/deleting a point ⇒ works bad with unknown order.
  - Divisions present in computing the Lagrange polynomial are expensive

Newton's Divided Difference Polynomial

- An i-th order Newton basis polynomial:

$$\phi_i(x) = \prod_{j=0}^{i-1}(x - x_j)$$

- The interpolating polynomial in terms of Newton's basis:

$$P_n(x) = \sum_{i=0}^{n} c_i \phi_i(x)$$

- Get the Coefficients:

$$c_i = [y_0, \ldots y_i]$$

where $[y_0, \ldots y_i]$ is the notation for an $(i+1)$-th order divided difference.

Newton (Cont'd)

- Virtues:
  - For equally spaced data points, replace the divided differences with functional differences.
  - Less arithmetic operations in writing the polynomial than that of Lagrangian.
  - Easy to add/delete a point $\Rightarrow$ works well for an unknown order
- All the above methods yield the same results for a given set of points. However, for larger n,they all suffer from the Runge's phenomenon.

Runge's Phenomenon

- Is the error is always guaranteed to diminish with increasing polynomial order? No!
- Runge observed an increasing oscillatory behavior when using polynomial interpolation with polynomials of high degree.
- Why?
  - The error between the generating function and the interpolating polynomial of order n is bounded by the n-th derivative of the generating function. For Runge-type functions (e.g., $f(x) = \frac{1}{1+25x^2}$), the magnitude of the derivative increases.

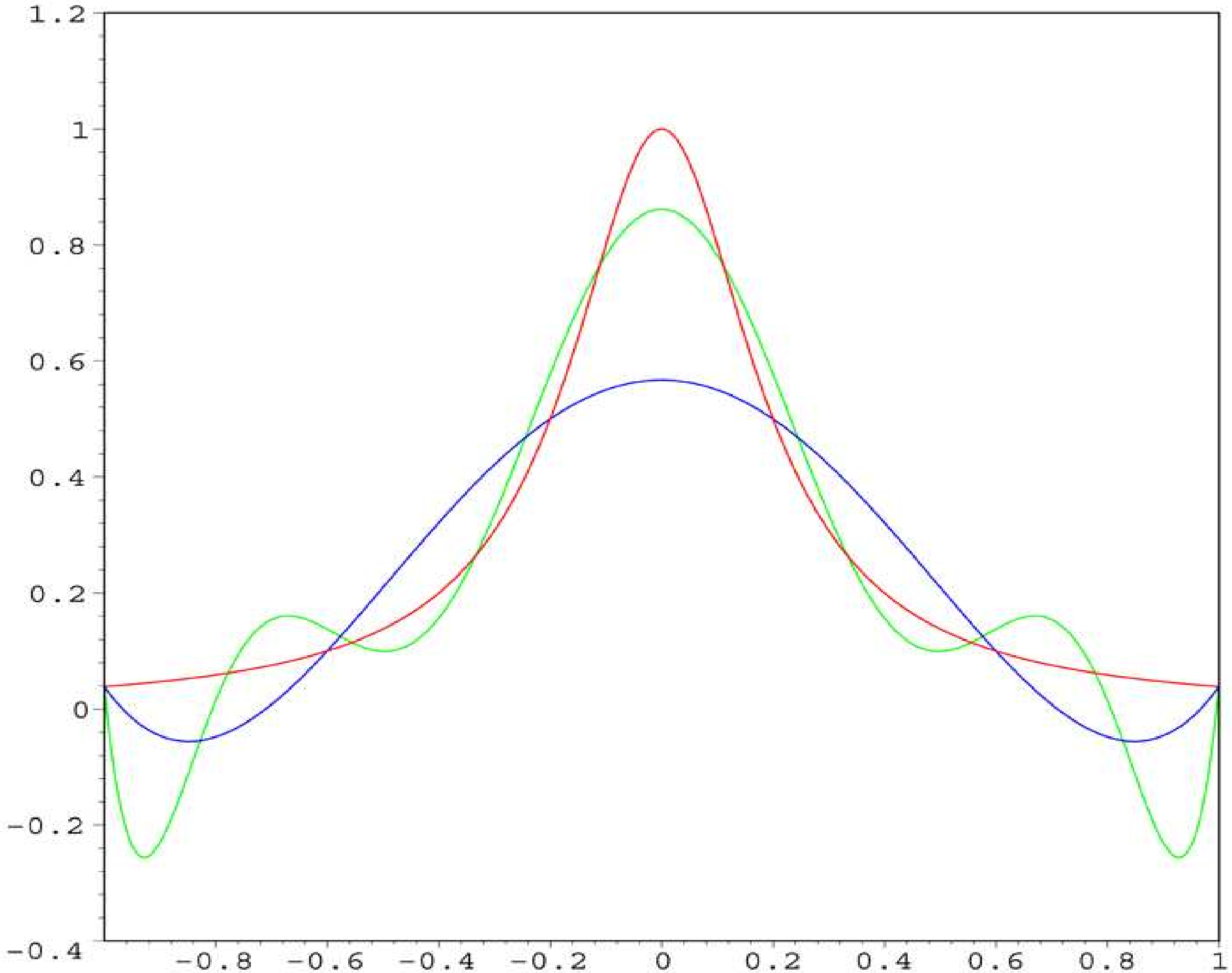
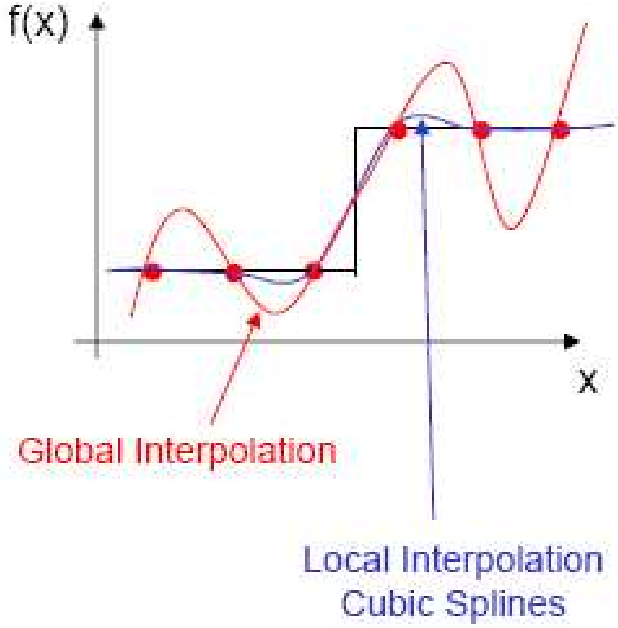


Figure 1: Runge Phenomenon in a nutshell (Runge function-red, 5th-order polynomial-blue, 9th-order polynomial-green)

Splines

- Local approach dividing into sub-intervals and fit to a low-order polynomial while preserving the following properties:
  - Continuity at the boundary
  - Slope continuity at the boundary
  - Curvature continuity at the boundary ...
- Spline candidates:
  - Linear
  - Quadratic
  - Cubic
- Useful for functions with local abrupt changes

Linear Splines



Linear Splines

$$
\begin{aligned}
f(x) &= f(x_0) + m_0(x - x_0)\,, \quad x_0 \le x \le x_1 \\
f(x) &= f(x_1) + m_1(x - x_1)\,, \quad x_1 \le x \le x_2 \\
&\vdots \\
f(x) &= f(x_{n-1}) + m_{n-1}(x - x_{n-1})\,, \quad x_{n-1} \le x \le x_n
\end{aligned}
$$

$$
m_i = \frac{f(x_{i+1}) - f(x_i)}{x_{i+1}) - x_i}
$$

Quadratic Splines

Quadratic Splines

$$f_i(x) = a_i x^2 + b_i x + c_i \;, \;\; x_{i-1} \leq x \leq x_i$$

3n coefficients

Continuity

$$\begin{aligned} a_{i-1}x_{i-1}^2 + b_{i-1}x_{i-1} + c_{i-1} &= f(x_{i-1}) \\ a_i x_{i-1}^2 + b_i x_{i-1} + c_i &= f(x_{i-1}) \end{aligned}$$

2(n-1) conditions

End Conditions

$$\begin{aligned} a_1 x_0^2 + b_1 x_0 + c_1 &= f(x_0) \\ a_n x_n^2 + b_n x_n + c_n &= f(x_n) \end{aligned}$$

2 conditions – 2n total

Derivatives

$$f_i'(x) = 2a_i x + b_i$$

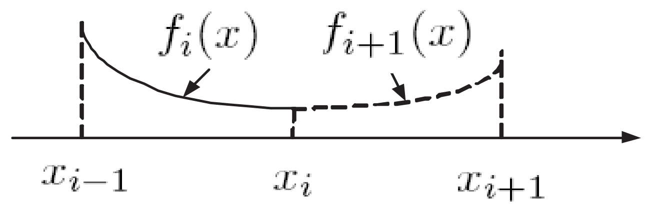
$$2a_{i-1}x_{i-1} + b_{i-1} = 2a_i x_{i-1} + b_i$$

n-1 conditions – 3n-1 total

$$a_1 = 0$$

1 condition – 3n total

Cubic Splines



- Cubic spline of the form:

$$f_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i \qquad x_{i-1} \le x \le x_i$$

- Find $4n$ unknowns from the following conditions:
  1. Continuity: $2(n-1)$ conditions
  2. End: 2 conditions
  3. Slope continuity: $(n-1)$ conditions
  4. Curvature continuity: $(n-1)$ conditions
  5. Curvature at end points: 2 conditions

Linear Regression

- Assumptions:
  - We look for a general trend of the data set
  - Noisy data are available in large scale
- Fitted model:

$$f(x) = a_0 + a_1 x.$$

- Objective function: Sum of error-squared:

$$J(a_0, a_1) = \sum_{i=0}^{n} (y_i - a_0 - a_1 x_i)^2$$

Linear Regression (Cont'd)

- To find the unknown coefficients, set the partial derivatives to be zero:

$$\frac{\partial J}{\partial a_0} = -2\sum_i (y_i - a_0 - a_1 x_i) = 0$$
$$\frac{\partial J}{\partial a_1} = -2\sum_i [y_i - a_0 - a_1 x_i] x_i = 0$$

- Rearrange the above to get

$$\begin{pmatrix} n+1 & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix}$$

- For the above case, the coefficient matrix $A = BB^T \Rightarrow BB^T[a] = By$. Solve the above system using the SVD or Cholesky.

References

- C. F. Gerald & P. O. Wheatley, *Applied numerical analysis*, Pearson, 2004.
- S. Chapra & R. Cannale, *Numerical methods for engineers*, 2006.

Thank you!

Q2. Discuss notion of statistical independence for pair of events, and corresponding situation for multivariate distribution

Events and Probabilities

- The sample space, Ω of an experiment consists of all possible mutually exclusive outcomes .
- An event is a set of outcomes.
- The probability of an event A:

$$P(A) \quad = \quad \lim_{N\to\infty} \frac{\text{No. outcomes of A}, N_A}{\text{No. trials}, N}$$

- Ex. Tossing a fair coin.

$$\begin{aligned} \Omega &= \{H, T\} \\ A &= \{H\} \end{aligned}$$

$$\text{Hence, } P(A) = \frac{1}{2}$$

Independent Events

- Definition: Two events A and B are independent if

$$P(A \cap B) = P(A)P(B).$$

- Interpretation:
  - On the LHS, $A \cap B \Rightarrow$ the event that joint/both events A and B occur.
  - On the RHS, we have the product of the probabilities of the individual events/marginals.
  - Intuitively it means that the occurrence of one event does not alter the occurrence probability of the other!

More Insight from the Conditional Probability

- Definition: The conditional probability $P(B|A)$ is the the probability of event B given that A has occurred:

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

provided $P(A) \neq 0$.

- If A and B are independent $\Rightarrow$

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B).$$

- Interpretation: The event A does not improve our knowledge about the occurrence of B. It makes no difference to B if A has occurred or not.

Toy Example

- Experiment: Toss a coin twice.
- Let A and B be the events of getting head in the first and the second trial, respectively.
- Are the two events independent? Our intuition says Yes !
- Verify from the definition:

$$\begin{aligned} P(A) &= P(HH) + P(HT) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \\ P(B) &= P(HH) + P(TH) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \\ \text{But, } P(A \cap B) &= P(HH) = \frac{1}{4} = P(A)P(B) \end{aligned}$$

- Hence they are *independent* as expected!

Mutual Exclusiveness and Independence

- Not synonyms!
- If the events A and B are mutually exclusive $\Rightarrow$
  - From the set theory: $A \cap B = \phi$.
  - From the probabilistic point of view:

$$\begin{aligned} P(A \cap B) &= 0 \\ \text{or } P(B|A) &= \frac{P(B \cap A)}{P(A)} = 0 \end{aligned}$$

- Occurrence of A $\Rightarrow$ B has definitely not occurred $\Rightarrow$ A nice piece of information!
- Two mutually exclusive events are dependent except they are zero-probabilistic.

Mutual Exclusiveness (Cont'd)

- How we benefit from these two notions?
  - Mutually exclusive $\Rightarrow$ add probabilities to get joint probability
  - Independent $\Rightarrow$ multiply probabilities to get joint pdf.

Extending the notion to three events

- Conditions for 3 events to be independent:
  1. They should be pairwise independent. i.e.,
     - A and B are independent
     - B and C are independent
     - C and A are independent
  2. Knowledge of the joint occurrence of any two events is independent of the third event:

$$\begin{aligned} P(A \cap B|C) &= P(A \cap B) \\ P(B \cap C|A) &= P(B \cap C) \\ P(C \cap A|B) &= P(C \cap A). \end{aligned}$$

Or equivalently, we write $P(A \cap B \cap C) = P(A)P(B)P(C)$ in this case.

An Example

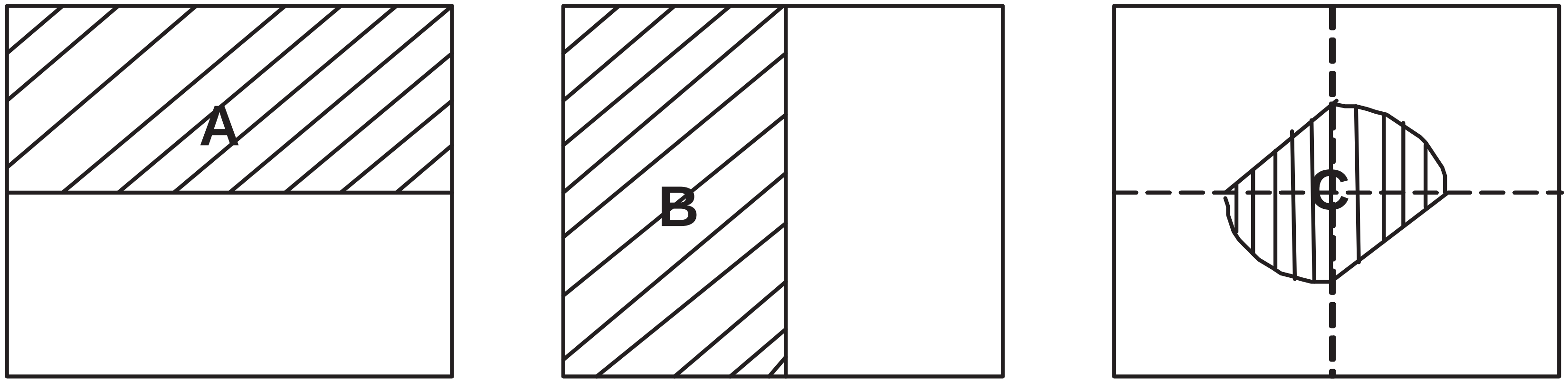


Figure 1: Simple Events

- Consider 3 events A, B and C in the Venn Diagram (Fig. 1).
- Q: Are these 3 events independent?

Example (Cont'd)

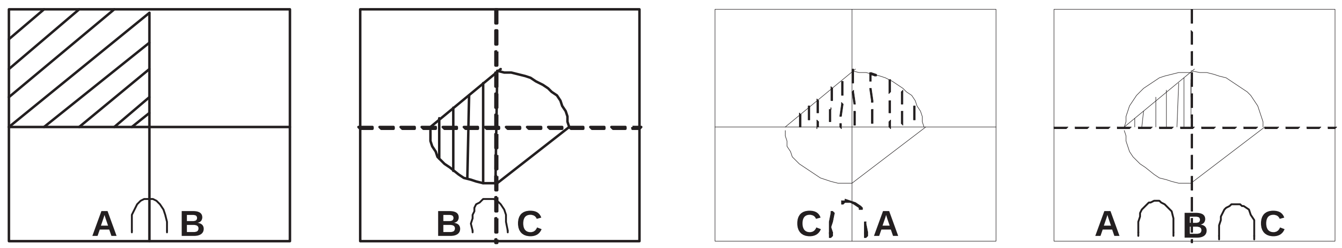


Figure 2: Joint Events

- A: They are pair-wise independent. Since

$$
\begin{aligned}
P(A|B) &= P(A) = \frac{1}{2} \\
P(B|C) &= P(B) = \frac{1}{2} \\
P(C|A) &= P(C) = \frac{1}{2}
\end{aligned}
$$

- However the 2nd condition does not hold:

$$P(A \cap B|C) < P(A \cap B) = \frac{1}{4}$$

Generalizing the notion to n events

- Multiplication rule. A set of n events $A_1, A_2, \ldots A_n$ are independent, if the probability of any subset of joint events is equal to the product of their marginal probabilities.

$$P(\cap A_i) = \prod P(A_i)$$

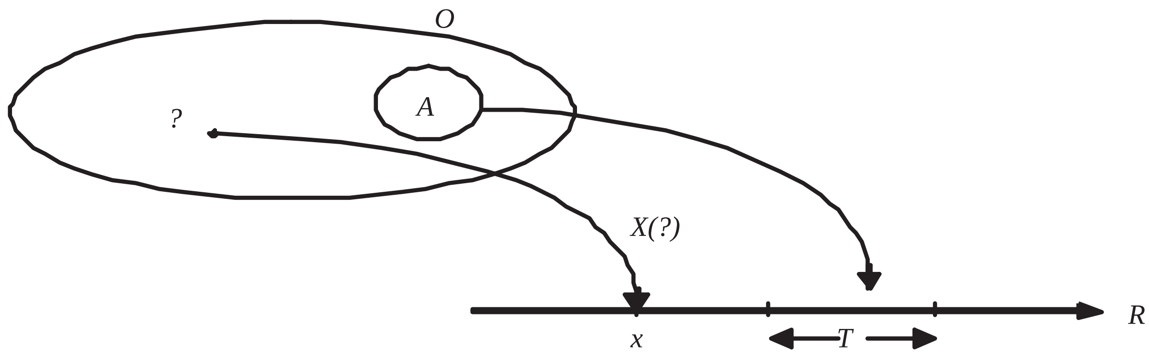
- Equivalently,

$$P(\cap_{i_1}^{i_m} A_i | \cap_{i_{m+1}}^{i_n} A_i) = \prod P(\cap_{i_1}^{i_m} A_i) = \prod P(A_i)$$

- At the heart of the independence, everything is independent of everything else.

- In practice, we assume that the outcomes of separate experiments are all independent.

Random Variables



- A real-valued random variable (rv) is function $X : \Omega \to \mathbb{R}$ that assigns a value to each outcome $\omega \in \Omega$.
- In the coin toss, suppose we receive \$1 if head appears and pay \$1 otherwise. In this case, we set the rv X to be the amount after first toss:

$$X = \begin{cases} 1, & \text{if H;} \\ -1, & \text{if T.} \end{cases}$$

Joint CDFs

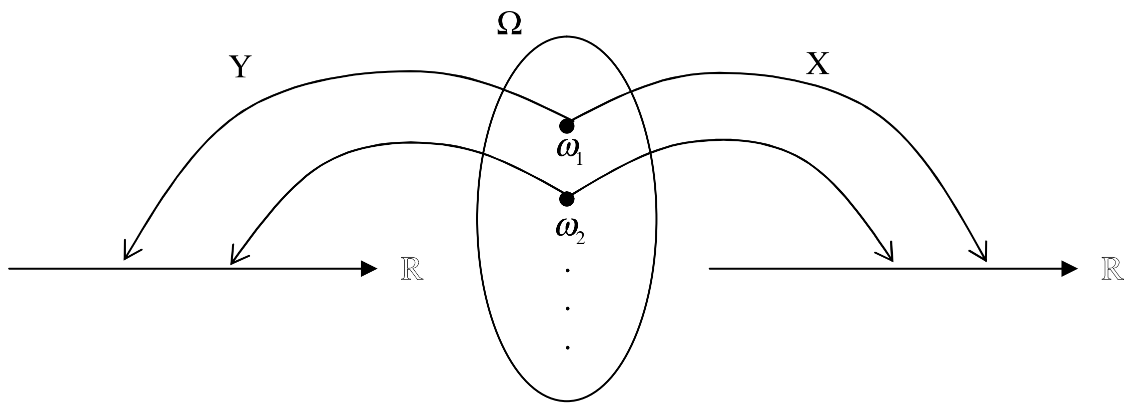


Figure 3: Multivariate RVs

- The joint (cumulative) distribution function of two RVs X and Y is the function $F : \mathbb{R} \rightarrow [0, 1]$ such that

$$\begin{aligned} F_{X,Y}(x, y) &= P(X \leq x, Y \leq y) \\ &= P(\omega \in \Omega | \{X(\omega) \leq x\} \cap \{Y(\omega) \leq y\}) \end{aligned}$$

Independence of Multivariate RVs

- Definition: The two random variables X and Y are independent $\Leftrightarrow$ For any number x and y, the event $A = \{X \leq x\}$ is independent of event $B = \{Y \leq y\}$.

- Recall the joint distribution function of X and Y:

$$F_{X,Y}(x, y) = P[\underbrace{\{X(\omega) \leq x\}}_{A} \cap \underbrace{\{Y(\omega) \leq y\}}_{B}]$$

- But events A and B are independent $\Rightarrow$

$$P[A \cap B] = P(A)P(B)$$
$$\text{hence } F_{X,Y}(x, y) = F_X(x)F_Y(y) \quad \forall x, y \in \mathbb{R}$$

- Differentiating the distribution functions, we get the joint pdf

$$f_{X,Y}(x, y) = f_X(x)f_Y(y).$$

Toy Example Revisited

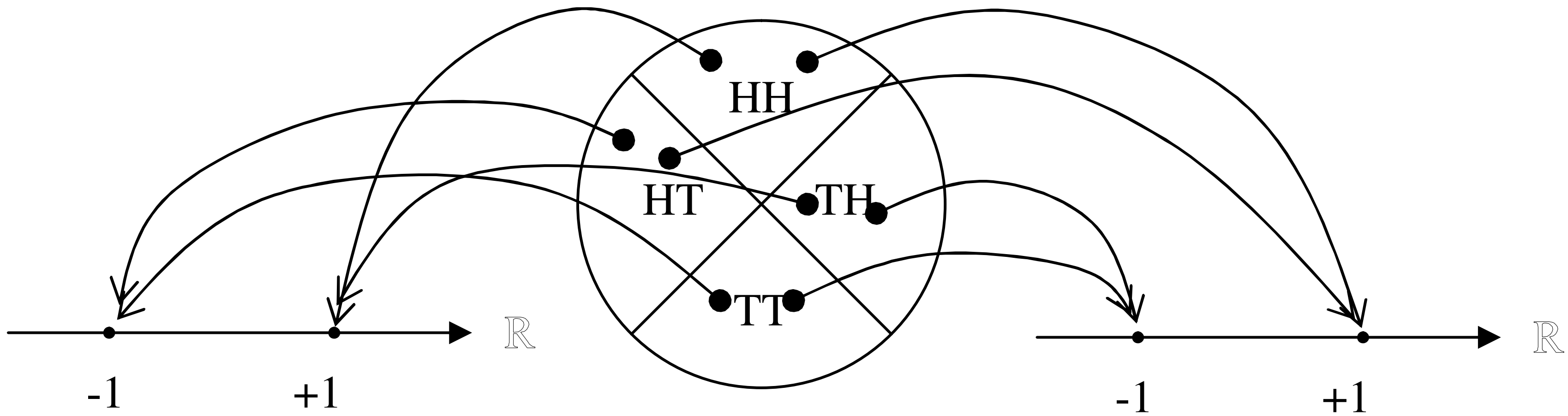


Figure 4: Tossing Coin Twice

- Experiment. Toss a coin twice
- let X and Y be the RVs denoting the outcome of first and second trials, respectively.
- Question: Are X and Y independent RVs ?

Toy Example (Cont'd)

- Solution:

$$
\begin{aligned}
P_X(x) &= \frac{1}{2} \quad x \in \{-1, 1\} \\
P_Y(y) &= \frac{1}{2} \quad y \in \{-1, 1\} \\
P_{X,Y}(x, y) &= \frac{1}{4} \\
&= P_X(x)P_Y(y) \quad \forall\{(x, y)\}.
\end{aligned}
$$

Concluding Remarks

- If RVs X and Y are independent, then
  - $E(XY) = E(X)E(Y)$
  - $\text{var}(X, Y) = 0 \Rightarrow \text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$
  - $f_{Y|X}(y|x) = f_Y(y)$.
- From the 2nd statement, independent $\Rightarrow$ uncorrelated, but not always conversely!
- Generalization. A set of n RVs are independent, if for any finite set of numbers $\{x_1, x_2 \ldots x_n\}$, the events $\{X_1 \leq x_1, X_1 \leq x_1 \ldots X_1 \leq x_1\}$ are independent.
- Equivalently, the joint pdf

$$f_{X_1,\ldots X_n}(x_1 \ldots x_n) = \prod_{i=1}^{n} f_{X_i}(x_i).$$

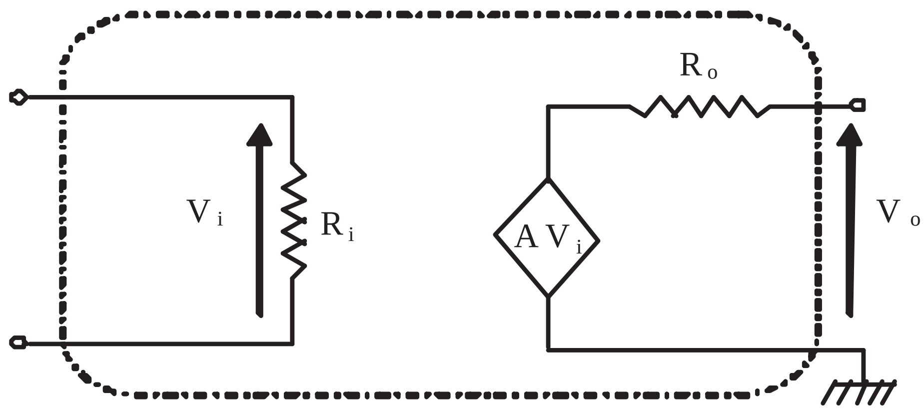
References

- A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, Hew York, NY, 1991.
- A. Leon-Garcia, *Probability and Random Processes for Electrical Engineering*, Addison-Wesley, 1989.
- R. Yates and D. Goodman *Probability and stochastic processes*, Wiley, 2004.

Thank you!

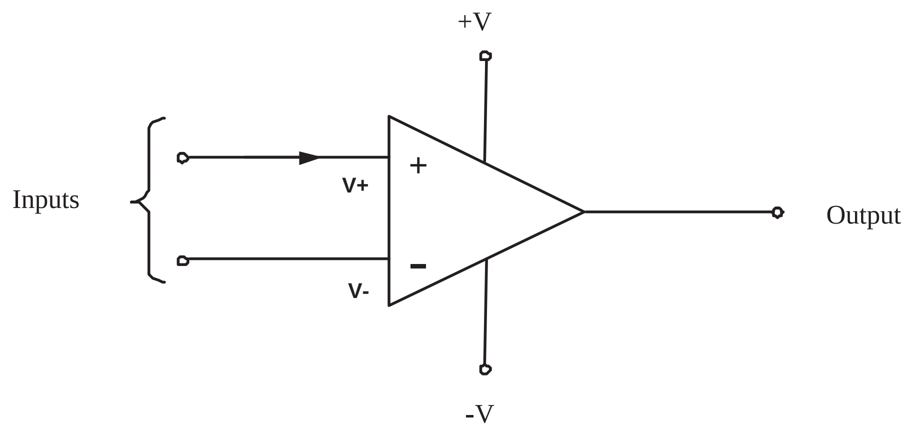
Q3. Calculate the transfer function of the following op-amp circuit and discuss applications

Real Vs. Ideal Op-amp



Parameter	Ideal	Real
R_{in}	∞	$10^6 - 10^{12}\Omega$
R_{out}	0	$100 - 1000\Omega$
A_d(OL)	∞	$10^5 - 10^9$
A_c(OL)	0	10^{-5}
Slew rate	∞	0.5V/microsecond
Gain-BW product	∞	$1 - 20$MHz

Golden Rules



- Voltage Rule: $v^+ = v^-$
- Rationale: $v_o = A_d v_i$ is limited; but $A_d \uparrow \infty \Rightarrow v_i \downarrow 0$.
- Current Rule: $i_{\text{in}} = 0$
- Rationale: $R_i = \infty$.

Why Negative Feedback?

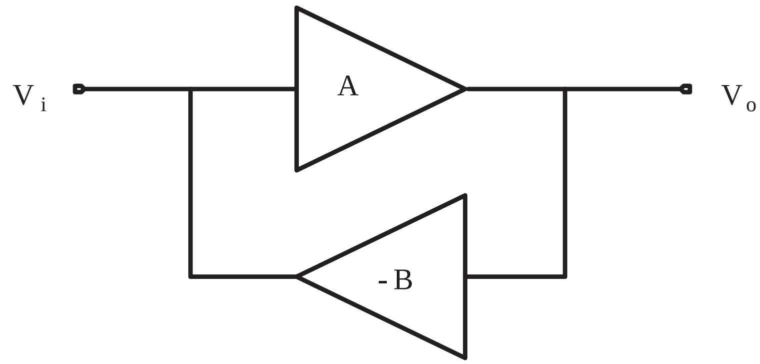


Figure 1: Typical negative feedback

- An op-amp with negative feedback provides the following benefits:
  - Allows to control the voltage gain. For the above circuit, the gain is $\frac{1}{B}$ when $A \approx \infty$.
  - No need to know about the internal characteristics.
  - Extends the useful frequency range.
  - Improves stability (against temperature variations)

The Differential Op-amp: Analysis

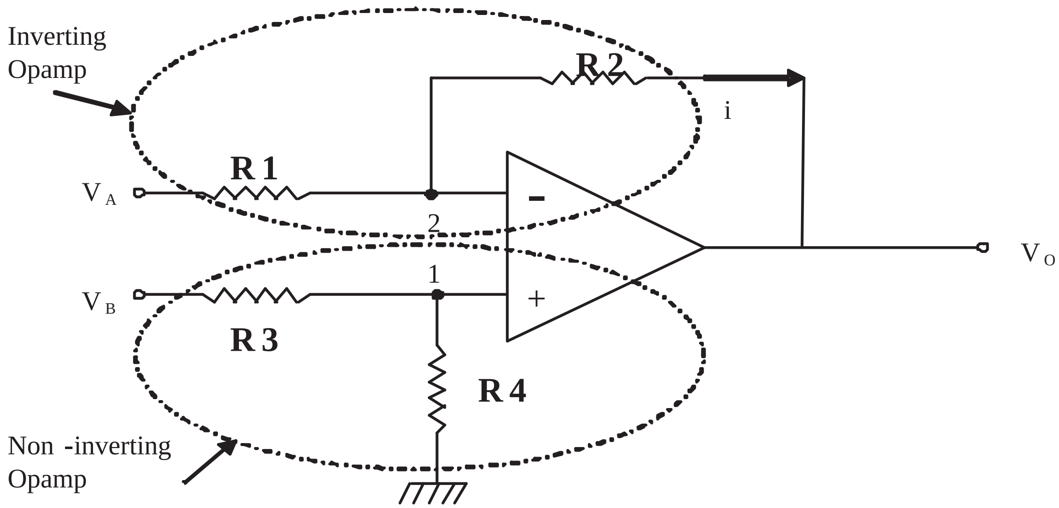


Figure 2: Example

- Compute voltage at the non-inverting terminal

$$v_1 = \frac{R_4}{R_3 + R_4} v_B \tag{1}$$

- From the voltage rule: $v_2 = v_1$.

- Recall the current rule; apply the KCL at node 2 to get

$$\frac{v_A - v_2}{R_1} = \frac{v_2 - v_o}{R_2} \qquad (2)$$

- Substituting (1) into (2) yields

$$v_o = \frac{(R_1 + R_2)R_4}{(R_3 + R_4)R_1} v_B - \frac{R_2}{R_1} v_A$$

Key Features

- Set all resistors to be equal ⇒ difference op-amp:

$$v_o = (v_B - v_A)$$

- Set $R_1 = R_3$ and $R_2 = R_4$ ⇒ amplified difference:

$$v_o = \frac{R_2}{R_1}(v_B - v_A)$$

Differential signaling

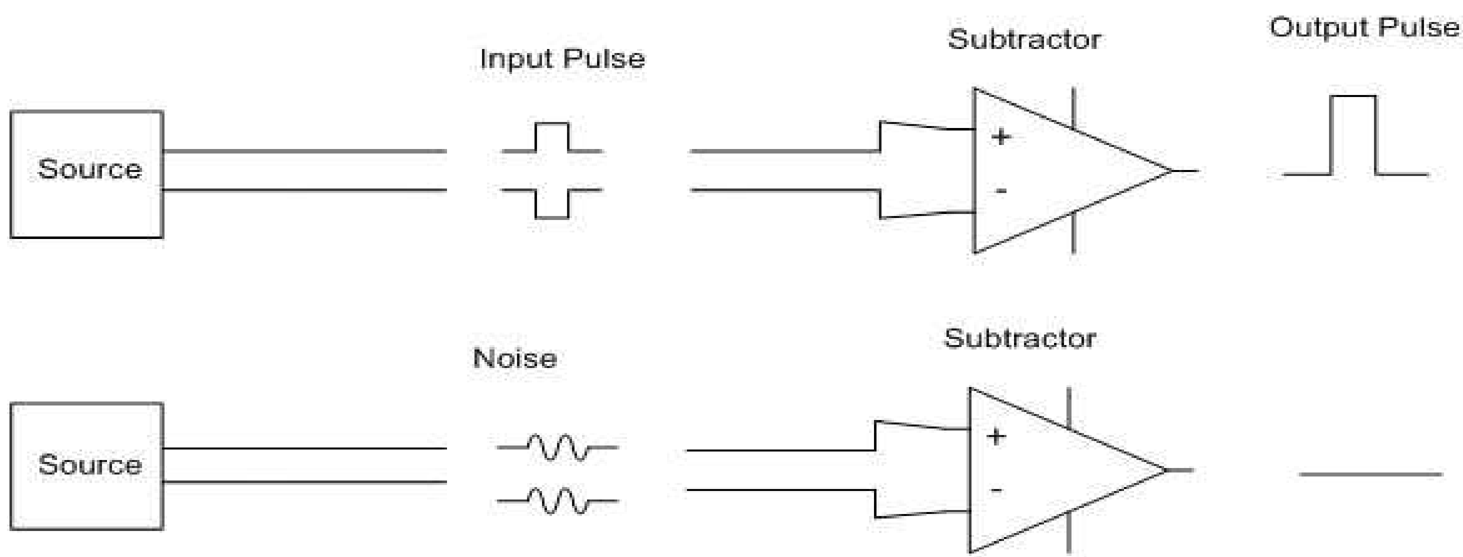


Figure 3: A differential receiver setup

- Input signals can be either analog or digital.
- – Desired input: differential-mode signal
  – Noise: common-mode signal

Differential signaling (cont'd)

- Two basic operations:
  - Amplifying desired small-signal
  - Filtering out noise
- Benefits:
  - Tolerance of ground offsets
  - Suitability for use with low-voltage (<5 volts) electronics
  - Resistance to noise interference(e.g., AC power line, circuit noise).
- Applications:
  - Data transmission (e.g., USB)
  - ECG
  - Thermocouple
  - as a stable comparator module

Comparator

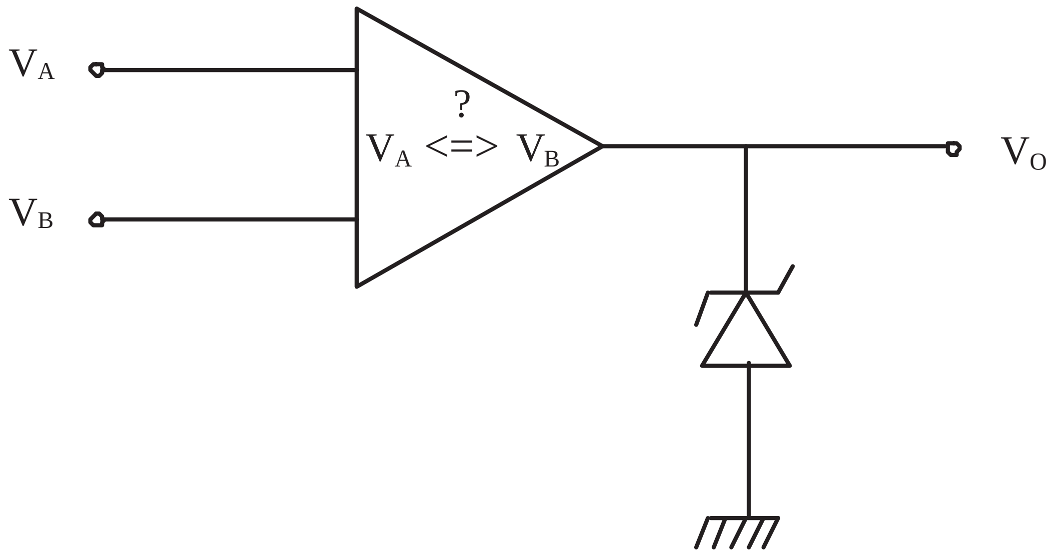


Figure 4: Differential Op-Amp as a Comparator

References

- T. Floyd, *Electronic Devices*, 6th ed., 2002.
- G. Rizzoni, *Principles and applications of electrical engineering*, McGraw Hill, 2004.

Thank you!

Q4. Discuss the linear convolution of the 2-finite length sequences using the DFT

DTFT: A Quick Recap

- Extends the FT for non-periodic discrete-time signals
- Forward DTFT:

$$X[\Omega] = \sum_{n=-\infty}^{\infty} x[n]e^{-j\Omega n}$$

- Periodic spectrum of period 2π.
- Abandon to use the DTFT in a digital signal processer for the following reasons:
  - DTFT Spectrum $X[\Omega]$ is continuous
  - Real signals have finite length

Discrete Fourier Transform

- Extends the DTFT for non-periodic discrete-time signals (finite duration) with discrete frequencies.
- Samples the DTFT spectrum on the interval $[0, 2\pi]$ using N points.
- N-point DFT-pairs:
  - Forward

$$X[k] = \sum_{n=0}^{N-1} x[n] W^{kn}, \quad k = 0, \ldots (N-1)$$

where

$$W = \exp(-j\frac{2\pi}{N}).$$

  - Inverse

$$x[n] = \sum_{k=0}^{N-1} X[k] W^{-kn}, \quad n = 0, \ldots (N-1).$$

DFT-pairs in Block-Matrix Form

- Let

$$
\begin{aligned}
\mathbf{x} &= [x(0), x(1), \ldots x(N-1)]^T \\
\mathbf{X} &= [X(0), X(1), \ldots X(N-1)]^T \\
\mathbf{W} &= \begin{pmatrix} W^0 & W^0 & \ldots & W^0 \\ W^0 & W^1 & \ldots & W^{N-1} \\ \ldots & \ldots & \ldots & \ldots \\ W^0 & W^{N-1} & \ldots & W^{(N-1)^2} \end{pmatrix}
\end{aligned}
$$

- DFT-pairs in a matrix form:

$$\mathbf{X} = \mathbf{W}\mathbf{x} \quad (1)$$

$$\mathbf{x} = \frac{1}{N}\mathbf{W}^{\mathbf{H}}\mathbf{x} \quad (2)$$

- Requires N^2 complex multiplications and $N(N-1)$ complex additions.

DFT-pairs (Cont'd)

- Taking complex-conjugate of (2) twice replaces the IDFT with DFT:

$$\mathbf{x} = \frac{1}{N}\big(\mathrm{DFT}(\mathbf{X}^*)\big)^*.$$

- Can be implemented using lightening-speed algorithms!

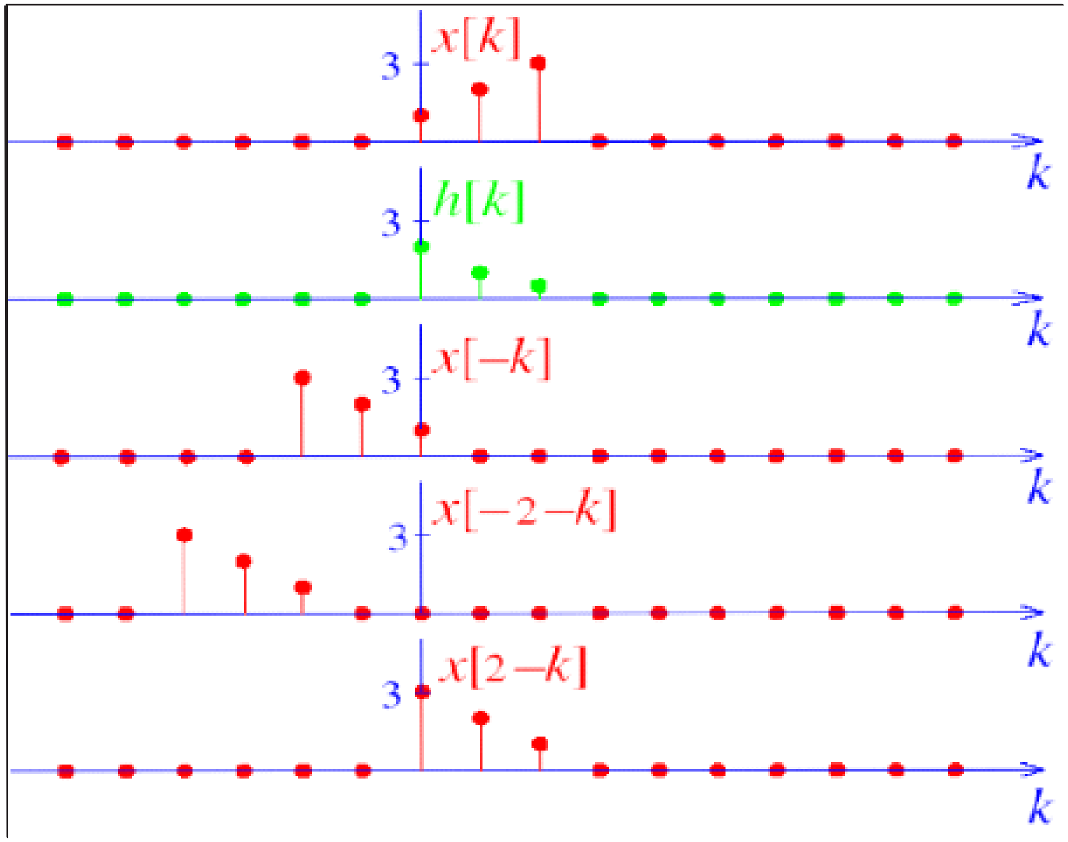
Linear Convolution

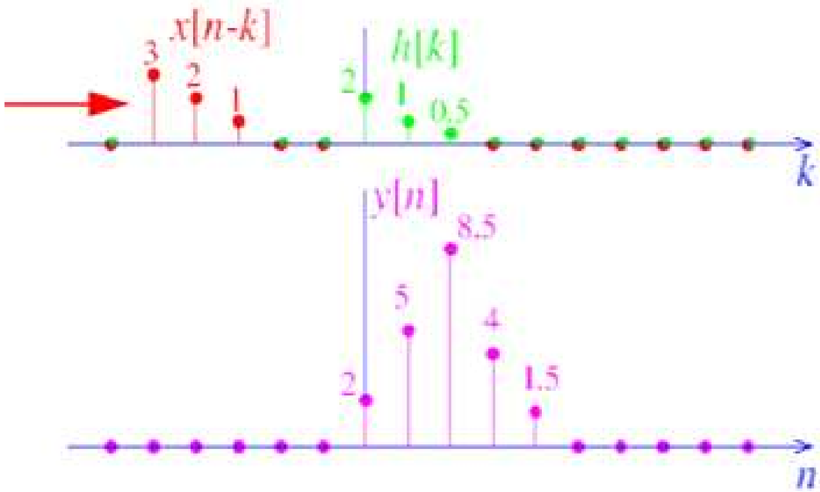
- Definition: Suppose two sequences $h[n]$ and $x[n]$ of length L and P, respectively.

$$y[n] = \big(h * x\big)[n] = \sum_{k=-\infty}^{\infty} h[k]x[n-k] \qquad (3)$$

- Basic operations:
  - Time invert one of the sequences
  - Slide it from $-\infty$ to ∞
  - When sequences intersect, sum their products
- $y[n]$ is a sequence of length $(L + P - 1)$
- Analogous to computing coefficients of the product of two polynomials.

Example: Linear Convolution





Circular Shift

- Define the circular shift of sequence $x[n]$ of length N as

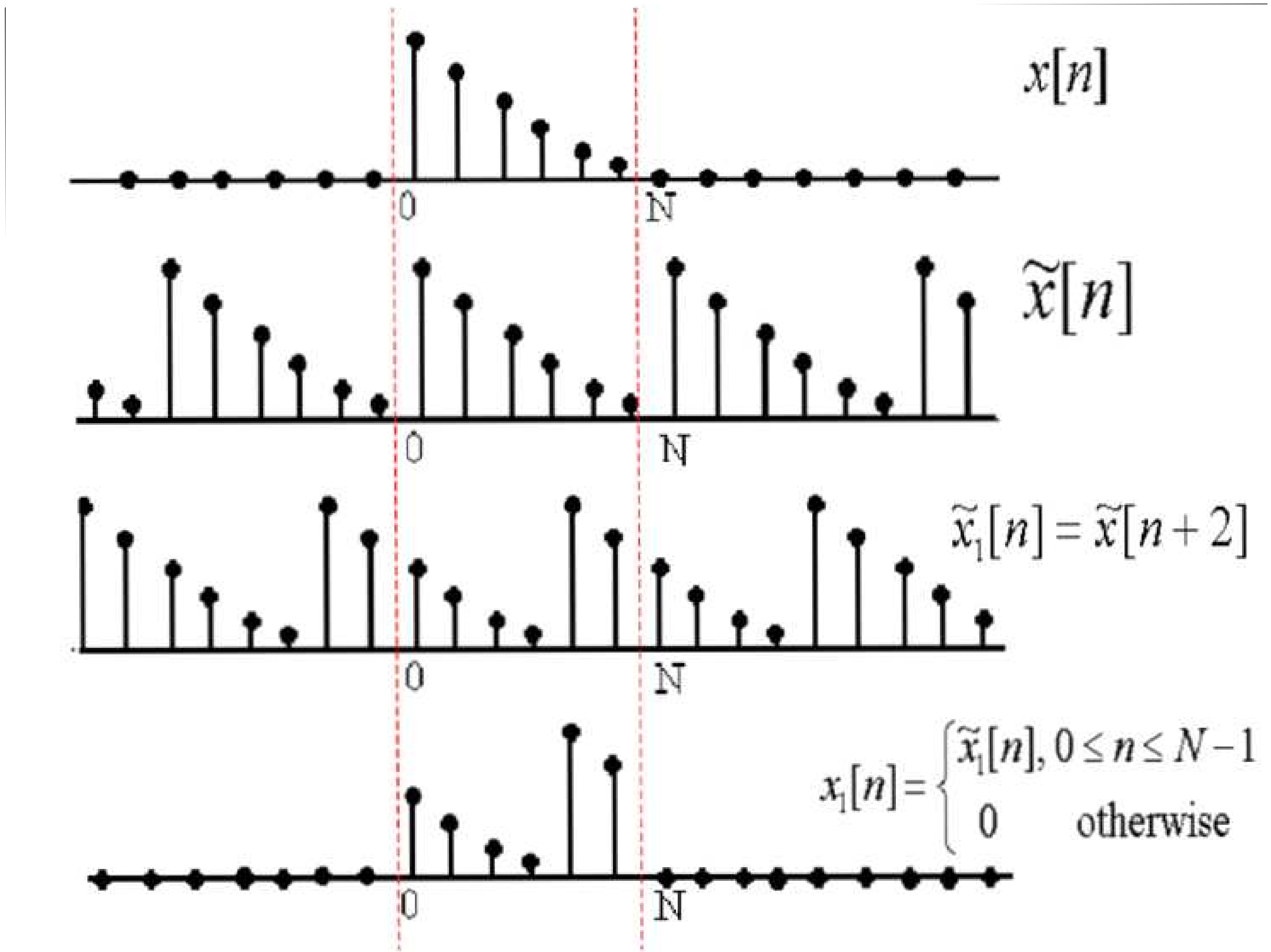
$$x_1[n] \quad = \quad \big(\tilde{x}[m-n]\big)\Pi_N(n)$$

where

  - $\tilde{x}[n]$ is the periodic extension of $x[n]$
  - $\Pi_N(n)$ the rectangular window in the interval $[0,(N-1)]$.

- 3 basic operations:
  - Periodic extension
  - Normal shift
  - Extraction of the sequence over one period $[0,(N-1)]$

Example (i): Circular Shift



Example (ii): Circular Shift

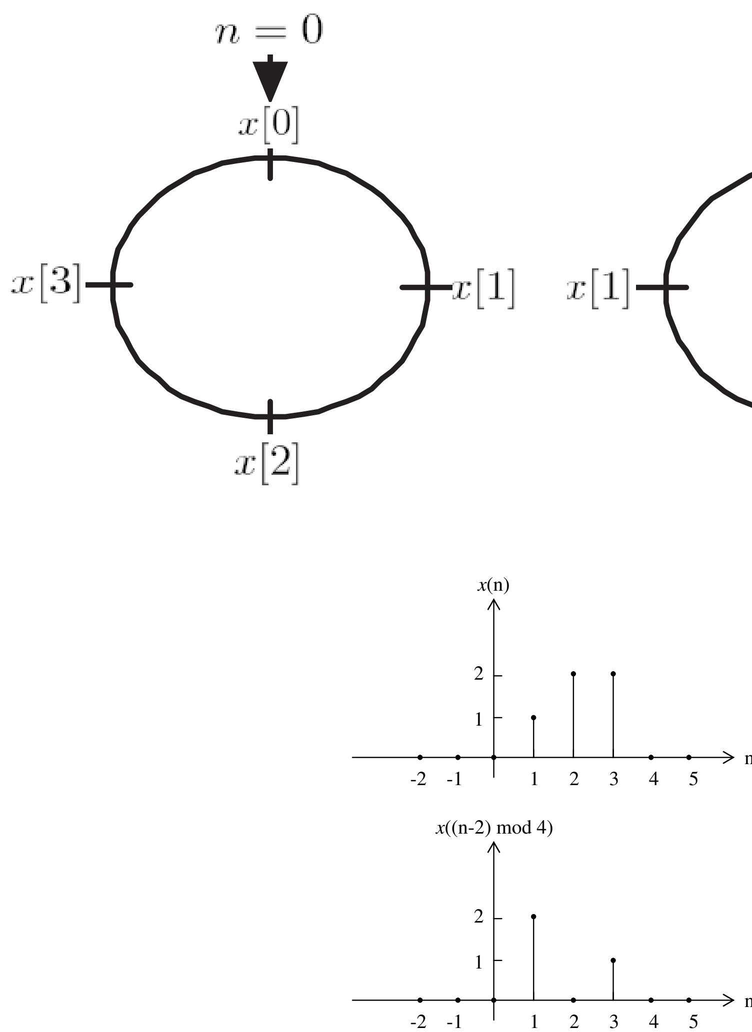


Figure 1: Right Circular Shift on $x[n] = [0, 1, 2, 2]$ by 2 points

Circular Convolution

- Definition: Suppose two sequences $h[n]$ and $x[n]$ of length N each.

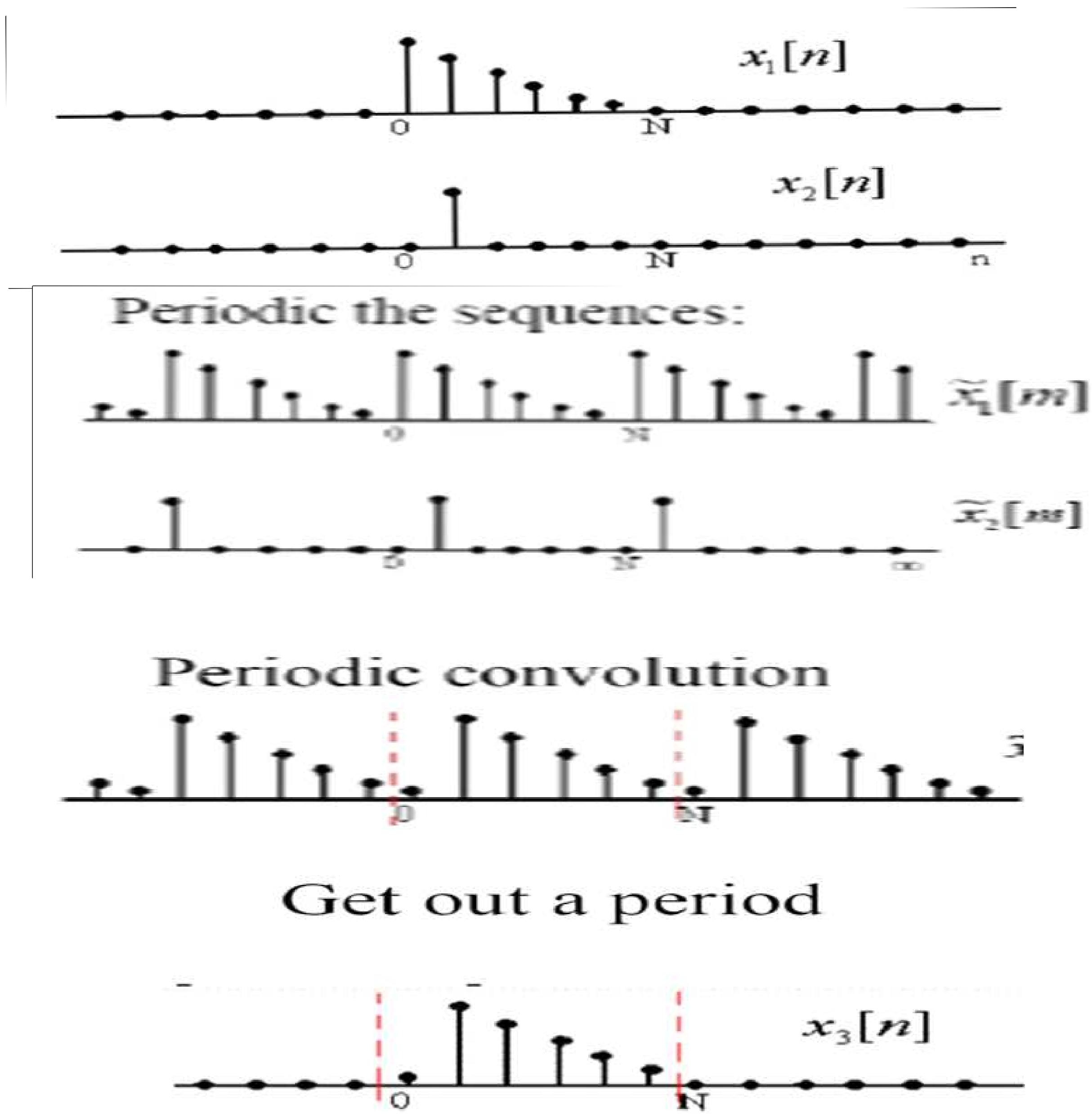
$$y[n] = h[n] \otimes x[n] \quad = \quad \Big(\sum_{m=0}^{N-1} \tilde{h}[m]\tilde{x}[n-m]\Big)\Pi_N(n).$$

- $y[n]$ is a sequence of length N.
- Key Property:

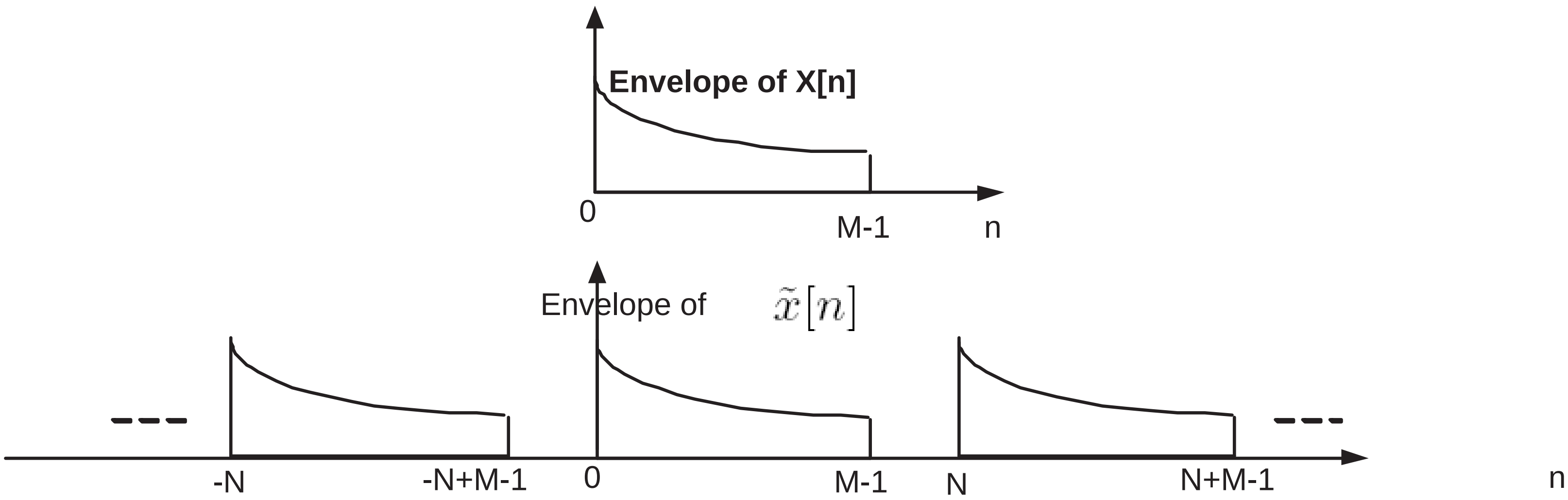
$$h[n] \otimes x[n] \quad \stackrel{\text{DFT}}{\Rightarrow} \quad H[k]X[k]$$

- 3 major differences from the linear convolution:
  - Periodic extension
  - Convolution is confined to one period
  - Truncation of one period at the end

Example: Circular Convolution



Zero Padding



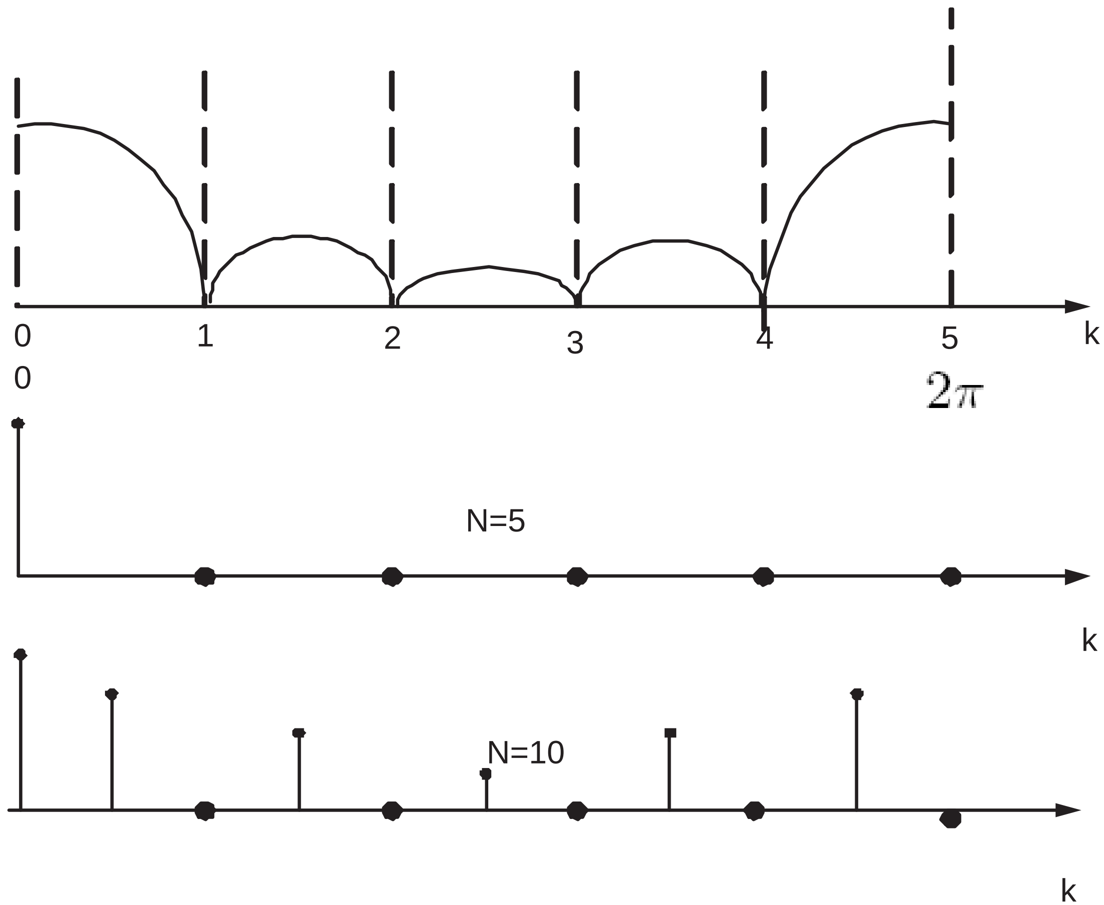
- Can we perform linear convolution using the DFT? If yes, how?
- Extend the length of each sequence such that

$$N \quad \geq \quad M = (L + P - 1),$$

then

$$h[n] \otimes x[n] \quad = \quad h[n] * x[n].$$

Zero Padding (Cont'd)



- Remarks on zero-padding:
  - improves the picture of the DTFT
  - does not increase spectral resolution or reduce the leakage.

Steps: Linear Convolution via DFT

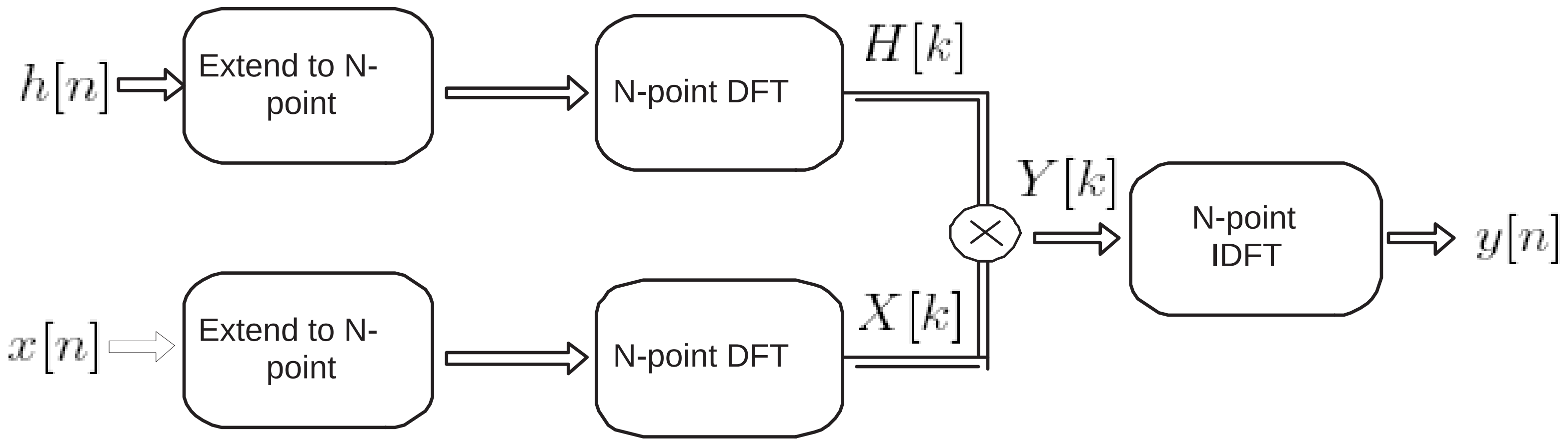


Figure 2: Flow Diagram

- Choose N to be at least $(L + P - 1)$.
- Pad the two original sequences with zeros to length N.
- Compute the N-point DFT to obtain $H[k]$ and $X[k]$.
- Compute the point-wise product:

$$Y[k] \quad = \quad H[k]X[k] \quad k = 0, \ldots (N-1).$$

Linear Convolution (Cont'd)

- Compute $y[n]$ by taking the N-point IDFT of $Y[k]$ as follows:
  - compute the DFT of $Y^*[k]$
  - take the complex conjugate
  - divide by $\frac{1}{N}$
- Save the first $(L+P-1)$ values of $y[n]$.

Final Remarks

- To speed up the process, do the followings:
  1. Use FFT in place of DFT with N being some power of 2.
  2. Suppose $h[n]$ is fixed. So pre-compute and save its DFT in advance.
- Linear convolution via DFT is faster than the 'normal' linear convolution when

$$\underbrace{O(N\log(N))}_{\text{FFT}} \quad < \quad \underbrace{O(LP)}_{\text{normal}}$$

References

- J. K. Zhang, *CoE 4TL4: Digital Signal Processing*, Course notes.
- S. Hayes, *Digital Signal Processing*, Schaum's Outline, 1999.
- A. Oppenheim & R. Schafer *Discrete-time signal processing*, 2nd ed.

Thank you!