Lecture # 1 21st-January-2025

• Introduction to CV.

jus66wg → gcr code

Lecture # 2 23rd-January-2025

• Images Reading (1st Book Chp 1)

⇒ Model based vs learning based CV

Model based: classical approach → done manually, on the basis of defined rules

Learning based: modern approach → AI based, labelled dataset pr train & then test.

Limitation of LB → depends on dataset (needs big dataset) & environment.
→ less interpretable

Test sorting on a nueral network using fixed array size.

Principle of Continuity

Image Formation:
↓
pinhole camera → perspective projection.

As all lightwaves are being projected hence no clear image is formed → hence pinhole camera focuses image reflection.

Make a pinhole camera

Lecture # 3 28-January-2025

Limitations of Computer Vision:

- privacy due to survillence
- biases in AI model due to training on specific data.

1) = Fairness:

Two problems occur

in data in algorithm.

Solution: GAN (generates new data without much diff in real & generated data)

protected attribute: exclude some properties that can affect the model, for example skin colour, race.

↳ attributes on which we don't want to train.

2) Privacy:

medical records & voter records were used simultaneously

Differential Privacy: a method that adds randomness to prevent identification but causes issues in back tracking of model.

3) Ethics:

- allowing machines to decide in case of life & death matters (self driving cars).

Image Formation

In image formed no 'z' dimension as object is 3D & image is 2D.

Perspective Projection.

greater the z smaller the image.

$x = f\frac{X}{Z}$, $y = f\frac{Y}{Z}$, magnification $= m = \frac{f}{Z}$, $area_{img} = area_{obj} \times m^2$

Vanishing Point: point where parallel lines converge.

False Perspective: things that appear to be far but are near.
↳ creation of depth.

Ideal Pin hole size: small size of pinhole causes diffraction of rays.

Exposure Time

Orthographic Projection:

Lecture # 4 30-January-2024

Image Formation – Lens.

Pinhole main img clear ni hoti due to less light coming and also because ray dont converge to one point isi liye lens have been now used.

Gaussian Lens Law:
Jab obj infinity pr hota hai to lines parallel ati hai.

Image Magnification.

$$m = \frac{h_i}{h_o} = \frac{i}{O}$$

Two lens system

$$m = \frac{i_2}{O_2} \cdot \frac{i_1}{O_1}.$$

Aperture of lens: used to control amt of light

F-number of lens: f-stop, f-ratio.

Lens Defocus

On reducing the aperture A & B will not change position, the focus point will not change only blur circle will be small.

small blur circle: less blur points.

Depth of field: focus on face more rather than on the background.

A point ka focus wohi rhy ga but light kaam hogi matlab brightness kaam hogi.

In real world one point is fully focused, usky agy or peechy waly points defocus hoty hain.

Depth of field (DOF):

objects in depth of field are equally focused

if size of pixel = size of blur circle or blur circle is less than pixel size then $O_2 - O_1$ is DOF

Image formation – Color

Light and its color - diffraction.

Q why does sky appear blue?

↳ Blue color has less wavelength & more frequency and has more dispersion.

Light Reflecting from Surfaces.

agr sai colour dekhna hai to normal light main dekhain

Colour Perception.

Lecture # 5 — 4-February-2025

Introduction to learning

learning main hum apna kaam asan krny ki kosish krty hai

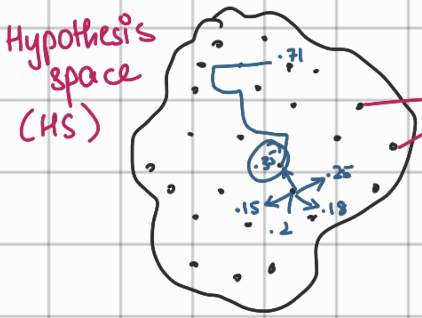
Types of learning

- Supervised: examples, labelled / classified data
- Unsupervised: clustering, objective function
  - L has also the hyperparameter value.
- Reinforcement: no training data

Reinforcement learning with human feedback.

Key Ingredients:

- Objective function: Tells whether the model is succeeding or losing.
- Hypothesis space:

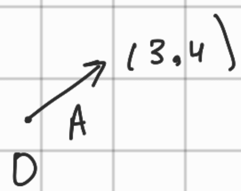


These all are models/functions which will give an output on basis of input.
All possible models that we can train.

- Optimizer: Tells how we will search HS.

Parameterization:

1) Hypothesis space
   a) All mapping from $\mathbb{R}^2 \to \mathbb{R}$
   b) $\mathbb{R} \times \mathbb{R} \to \mathbb{R}_0$ (Distance matrix)



2) Parameterized Hypothesis space

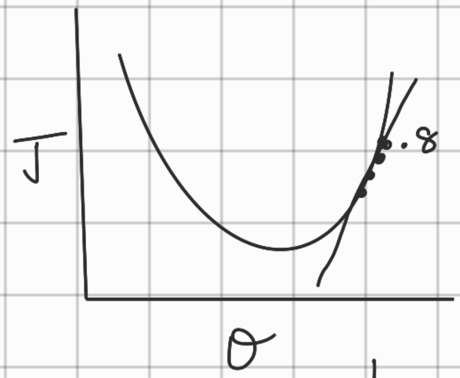
a) $\mathbb{R} \to \mathbb{R}$

i) $y = \theta_1 x + \theta_0$ ii). $y = \theta_2 \theta_1 x + \theta_0$.

iii) $y = \theta_1 x^2 + \theta_2 x + \theta_3$.

Take a simple start & start with linear, and if accuracy is decent enough we can carry on with them.

$y = \theta x$ $J(\theta) = \frac{1}{m} \sum (y - \hat{y})^2$



→ graph ka gradient nikalain gein or descent krain gein.

↳ jab gradient zero hojaiga tab hum ruk jain gein.

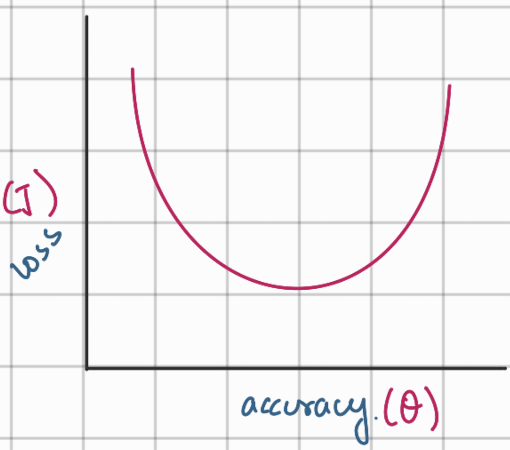
Two methods < Gradient descent / Random start.

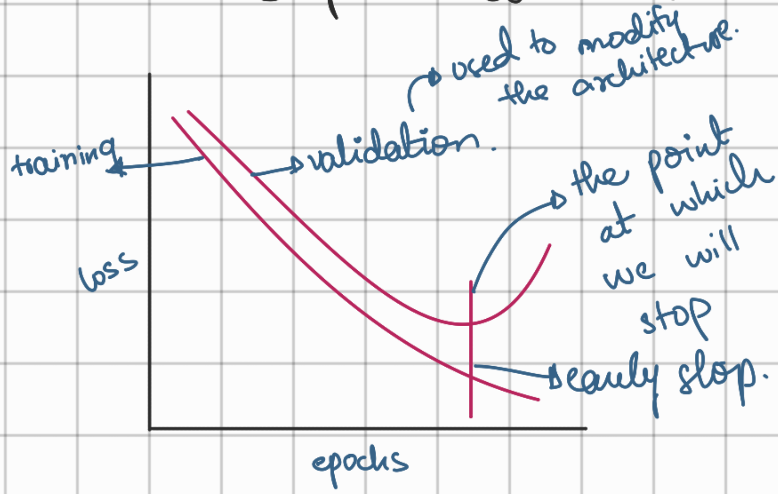
Overparametrization.

- we have to also be cautious of amount of parameters
  if # parameter = # of example then overfitting ka issue ajata hai.
- in normal circumstances have more e.g than parameters.
- deep learning.

Overparameterization:

Gradient descent looks for the simplest solution.





Empirical Risk Minimization.

$$\arg\min \frac{1}{N} \sum_{i=1}^{N} \mathcal{L}(f(x^{(i)}), y^{(i)}) \leftarrow \text{ERM.}$$

- Maximum Likelihood (we try to maximise data while minimizing loss)
  - used when we only have data.

- Maximum A Posteriori
  - $p(F) \rightarrow$ prior belief regarding what we are trying to find.
  - external source sy koi andaza & uspr model aid krain gain.

- Classification and Soft max Regression:
- 0-1 loss
  - discontinuous function hence we can't create gradient descent on this.
- Cross-Entropy loss (uncertainity)

• Cross Entropy loss.

coin	H	5	T	5	$p\log_2(p)$
	H	8	T	2	
	H	1	T	0	

General Optimization loop

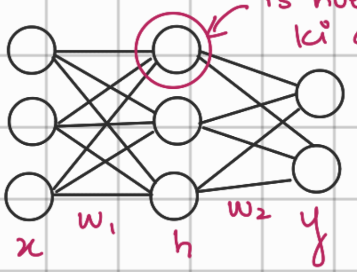
Lecture # 8 — 13-February-2025

Nueral Networks

-> **Activation vs Parameters** (both are tensors & inko learn krny ka algo backpropagation hai).
• To learn activation parameters must be kept constant.

Optimizing Activation

• jis pr sab sy zayada output ae ga us value pr us nueron ki activation sab sy ziada hy.



is nueron ki activation find krni hai

• loss function. • find krna hai maximizing nueron.

-> **Deep Nets** -> stack of linear & non-linear
-> **Deep learning**
• gradient based optimization • zeroth order as well.
• Hebbian learning (bottom up approach)
↳ more biological feasible then backpropagation.
↳ locally ird gird dekhta hai, or jo nueron sath fire holy hai unky weight strong ho jaty hai.

— Clip 12

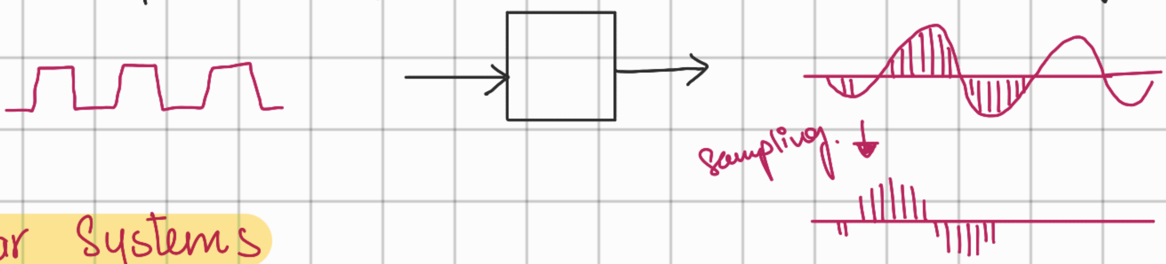
Linear Image Processing:

- Extract useful tokens from img.
- Enhance img structures (edge)
- Remove variability (noise)

Signal & Images

measurement of some physical quantity as a func of another independent quantity.

- system — process or function that transforms a signal into another.



Linear Systems

The output is the linear combination of every input.

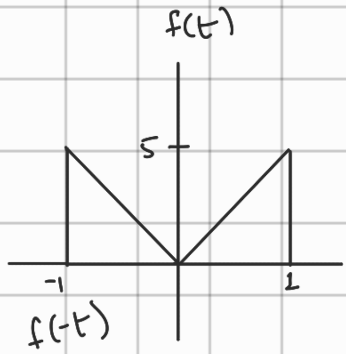
- Linear Translational Invariant Systems:

Aik system pr translation lagain gai to output bhi translate hoga
Transformation main sab pixels ki orientation same rehti hai pr rotation main ni

Rotation & Scaling main transformation in dependent on location of pixel.

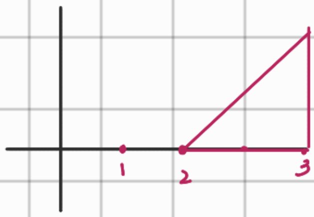
Convolution.

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau$$



$f(t-2)$ shifting

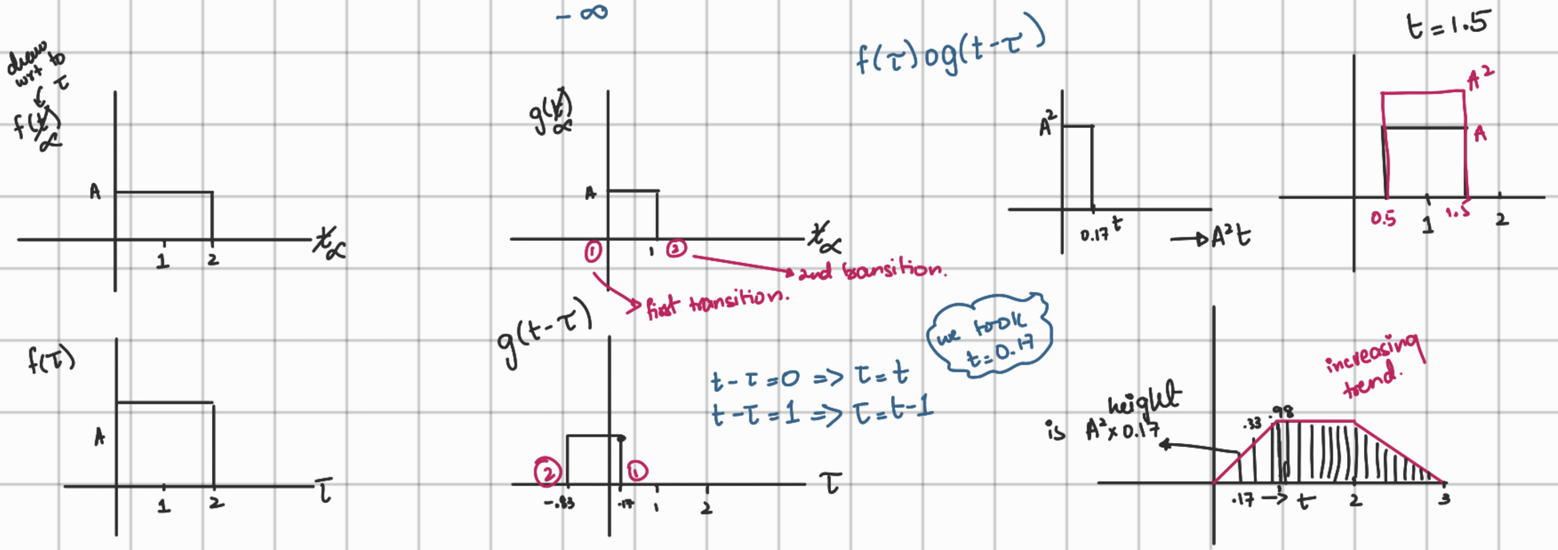
$t-2=1$
$t=1+2$
$t=3$



$f(-t) \to -t = 1$
$t = -1$

Lecture # 9   18-February-2025

Convolution: $(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)\, g(t-\tau)\, d\tau$



A^2 integral ki wajah sy ae ga. (Phly multiply krna hai or phir integrate krna hai.

→ δ function is impulse.

System Identification - Acoustics

T_1 delay

$h(t) = a_0 \delta(t) + a_1 \delta(t - T_1) + a_2 \delta(t - T_2) + a_3 \delta(t - T_3)$

$l_{out}(t) = (l_{in}(t) \circ h(t) = a_0 l_{out}(t) + a_1 l_{in}(t - T_1) + a_2 l_{in}(t - T_1) \ldots$

→ direct sound.

→ kaam amplitude reflection ki wajah sy echo.

Lecture # 10   20-February-2025

LTI solutions and Convolutions are directly related.

System Identification:

- noise cancellation
- same sound produce krwana.

↳ $h_{\Gamma}(t) = -a\,\delta(t - T_1) - a_2\,\delta(t - T_2) - a_3\,\delta(t - T_2)$

Noise Cancellation

- adaptive filter

Cross - Correlation vs Convolution.

'−' hota hai
not commutative,
not associative

'+' hota hai

Cross correlation is best for Tempelate matching.
It return flipped output but best for matching.

MIDTERM II

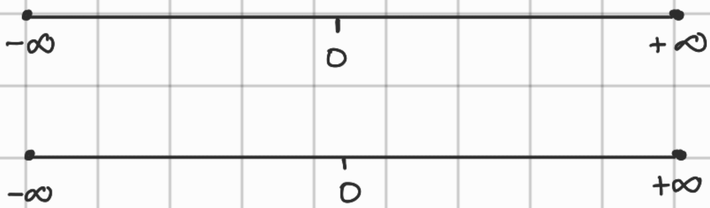
Lecture # 11 4-March-2025

Fourier Analysis.

Real

Imag

$j = i = \sqrt{-1}$



$\frac{x}{r} = \cos\theta$

$x = r\cos\theta$

$y = r\sin\theta$

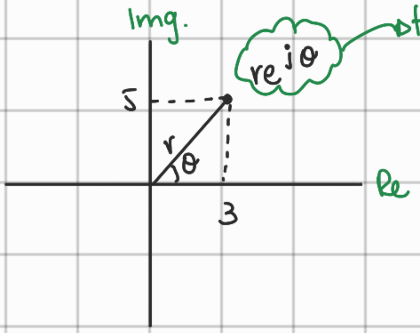
$r = \sqrt{x^2 + y^2}$

$\tan\theta = \frac{y}{x}$

$\theta = \tan^{-1}(y/x)$

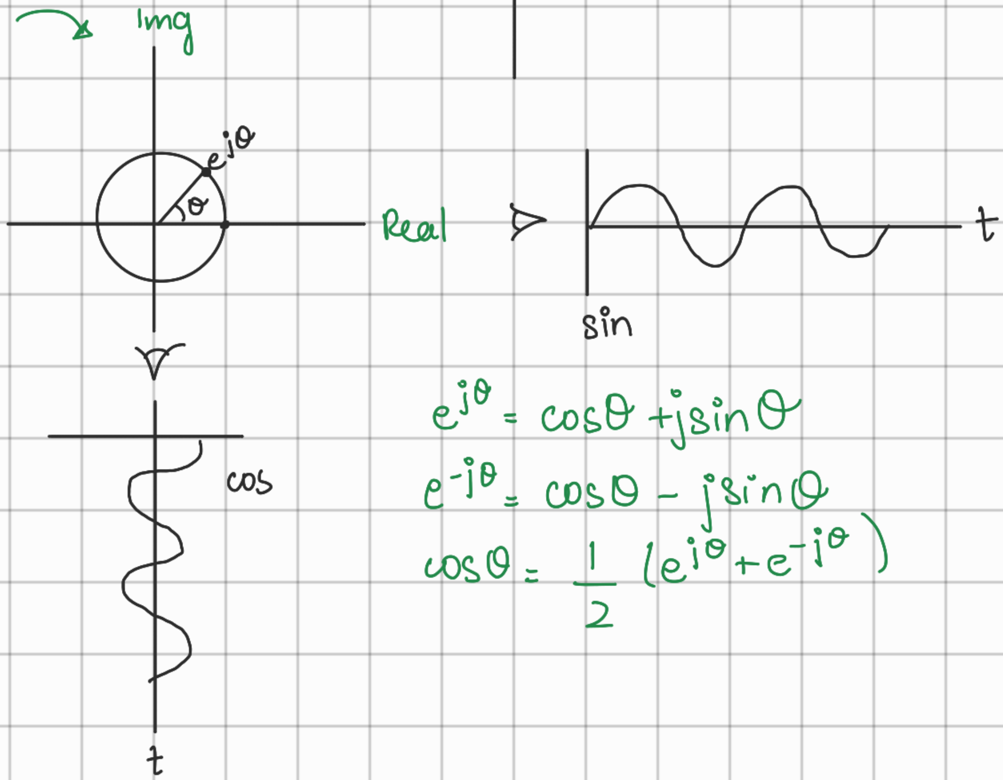
complex $(x + jy)$

$3 + 5j$



$re^{j\theta}$ → third form called exponential representation.

Euler's



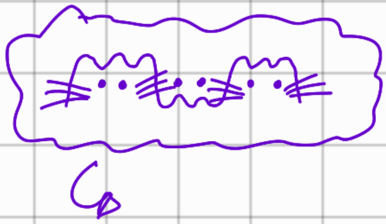
$e^{j\theta} = \cos\theta + j\sin\theta$

$e^{-j\theta} = \cos\theta - j\sin\theta$

$\cos\theta = \frac{1}{2}(e^{j\theta} + e^{-j\theta})$

- Fourier transform gives us constituent frequency. Once we find the frequencies we can manipulate them.

we can't transform frequency in spatial & time domain.



$Mv = \lambda v$

$$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$ ← eigen vector.

$$\begin{bmatrix} 3 \\ 3 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

LTI

$$[\;] = [\;][\;]$$

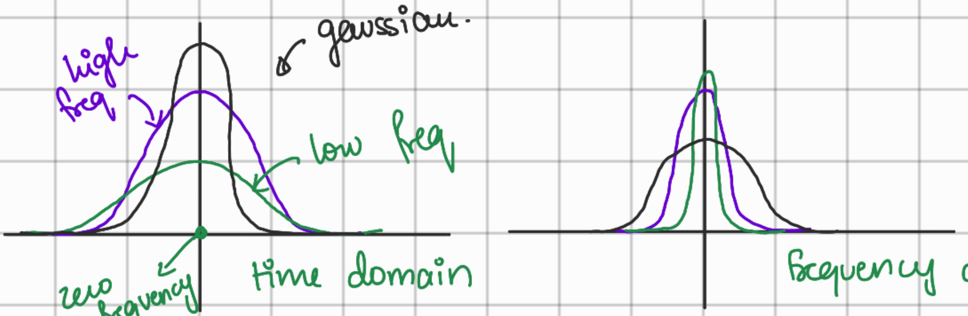
Lecture # 12 | 6-March-2025

Fourier Transform

↳ frequencies nikal kr deta hai.

graph pr y-axis → magnitude , x-axis → frequency.

• constant signals mai koi frequency ni hoti.



time domain

Frequency domain.

• zero frequency at origin.

higher frequency pr x-axis ka distance bdh jai ka.
agr gradual change hai tu frequency kam hai or agr steep ya abrupt change hai to frequency zayada hai.

$$\overbrace{F(\mu)}^{\text{frequency domain}} = \int_{-\infty}^{\infty} \underbrace{f(t)}_{\text{original func (time domain)}} \underbrace{e^{-j2\pi\mu t}}_{\text{sinosidal wave}}\, dt$$

original func main (μ) ki bht values hai.

hamara function time domain ka tha, hum ny usko exponential function sy multiply kiya or phir usko 't' pr integrate kiya, all possible frequencies sy multiply kiya or phir integrate kiya to sari frequencies bhir agei.

$$f(t) = \int_{-\infty}^{+\infty} F(\mu)\, \underbrace{e^{j2\pi\mu t}}_{\cos(2\pi\mu) + j\sin(2\pi\mu)}\, d\mu$$

$$f(t) = \left(\begin{array}{l} F(w_0) \times (\ \ t + \ \ t) \\ F(w_1) \times (\ \ t + \ \ t) \\ \vdots \\ \text{all possible values of } \mu \end{array} \right)$$

Discrete fourier transform.

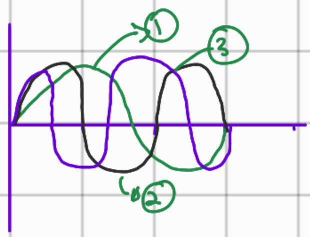
Is main multiply krky summation laity hai.

Fourier transformation is linear and hence can be written in matrix form.

fourier coefficients.

this row is summing, DC frequency low hai.

$$\begin{bmatrix} X[0] \\ X[1] \\ X[2] \\ X[3] \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & e^{-2\pi j\, 1/4} & e^{-2\pi j\, 2/4} & e^{-2\pi j\, 3/4} \\ 1 & e^{-2\pi j\, 2/4} & e^{-2\pi j\, 4/4} & e^{-2\pi j\, 6/4} \\ 1 & e^{-2\pi j\, 3/4} & e^{-2\pi j\, 6/4} & e^{-2\pi j\, 9/4} \end{bmatrix} \begin{bmatrix} x[0] \\ x[1] \\ x[2] \\ x[4] \end{bmatrix}$$



1 freq ki aik jump

2 freq ki 2 jump

3 freq ki 3 jumps.

upar wali sy bhi high.

frequency high hai.

• jab multiply krty hai to corelation nikltea hai.

Lecture # 13

11-March-2025

Fourier transform

cosine main aik hi frequency hoti hai.

$e^{j\theta} = \cos\theta + j\sin\theta$ → euler equation.

$e^{j2\pi kt} = \cos 2\pi kt + j\sin 2\pi kt$ — (i)

$e^{-j2\pi kt} = \cos 2\pi kt - j\sin 2\pi kt$ — (ii)

→ cosine θ ka fourier transform nikal rhy hai.

$\cos 2\pi kt = \frac{1}{2}(e^{j2\pi kt} + e^{-j2\pi kt})$ — eq(i) + eq(ii)

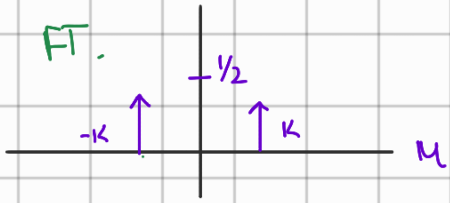
$\sin 2\pi kt = \frac{1}{2j}(e^{j2\pi kt} - e^{-j2\pi kt})$ — eq(i) - eq(ii) complex exponential main 1 frequency hai jo 'k' pr hai.

$F(e^{j2\pi kt}) = \delta(k)$, $F(e^{-j2\pi kt}) = \delta(-k)$

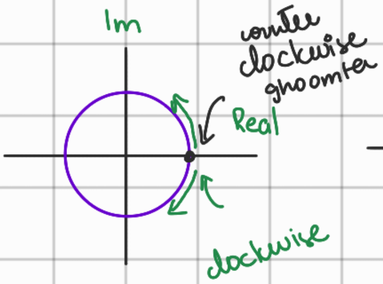
F.T of cosine:

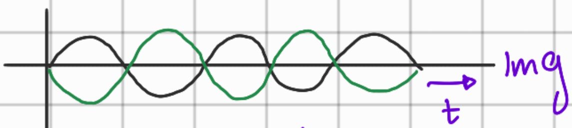
Iska FT real hai

$F(\cos 2\pi kt) = \underline{\tfrac{1}{2}(\delta(k)+\delta(-k))}$ cosine ka FT.



complex exponential ki depiction.





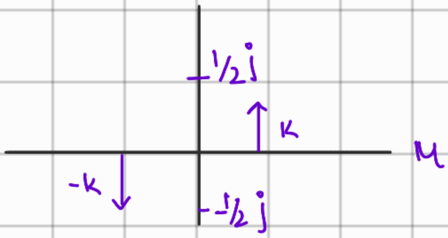
Ye waly cancel out ho jai ki

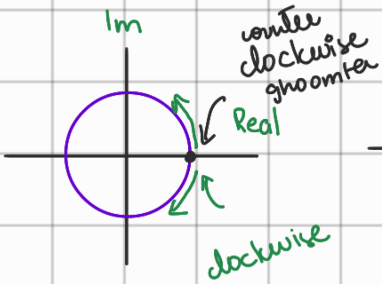
yh wali double ho jai gi.

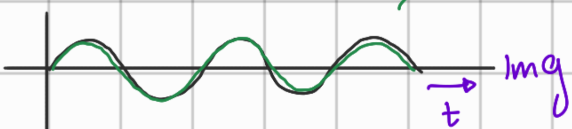
F.T of sine

Iska FT img hai.

$F(\sin 2\pi kt) = \frac{1}{2j}(\delta(k) - \delta(-k))$

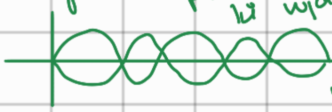




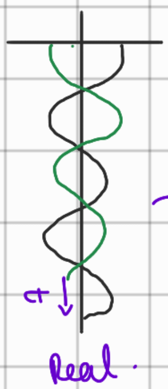


yh wali double ho jai gi.

phly yhi bny ga phir minus ki wjah sy

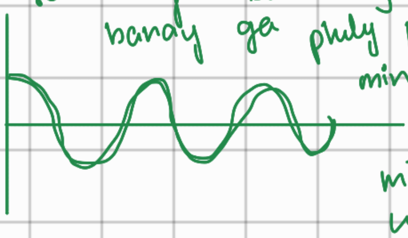


mirror hoga



yh waly cancel out ho jai ki

same yh bhi aisy banay ga phly phir minus hoke mirror hoga.



Discrete world main sampling hoti hai, hum saray points pr ni values leity, issi liye aik point ata hai jab highest or lowest ki value aik jitni ho jati hai.

• When sampling we can't differentiate b/w higher & lower frequencing. → Aaliasing.

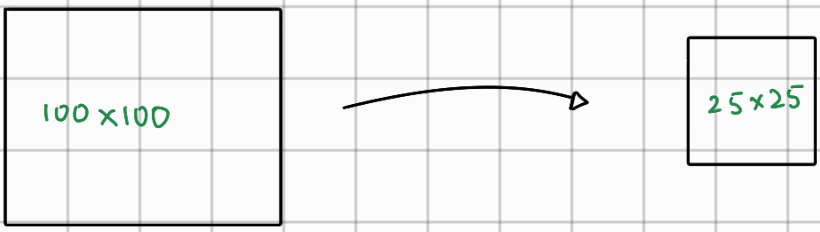
Lecture # 15 18-March-2025

Linear Filters – Blur filter

└ high frequencies are removed

Blur filter - Box filter

- reduces noises.
- different scales
- sampling & downsampling



$\frac{1}{9}\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$

└ avg pooling ka 3/3 ka kernel

max-pooling isn't an LTI

Down-Sampling (Reduced Image Size)

Horizontal filter ny vertical lines ko remove kr diya hai, kiynky vertical lines main horizontal change zioda hota hai.

Properties of Box Filter

a) Box = [1, 1, 1]

b) $[\ldots, 1, -1, 1, -1, 1, -1, \ldots]^n$

c) $[\ldots, 0.5, 0.5, -1, 0.5, 0.5, -1, \ldots]^n$

Lecture # 17 25-March-2025

IMAGE SAMPLING & ALIASING

- slow & smooth prior
- kisi bhi signal mai smooth change ae tu bhti hota hai
- information loss bhi nah ho or compute power or storage power bhi zeyada nah chaiye ho.
- sampling main wave recover krny ky liye smooth & slow prior chaiye hota hai.

Sampling frequency — 18
Aliasing frequency — 11

jab sampling adequate ni hoti tu artifacts ajaty hai.

Sampling Theorem.
Scale invariance.
Image Pyramids.
- Gaussian Pyramid
- laplacian Pyramid.
  - divides img into diff frequency bands.