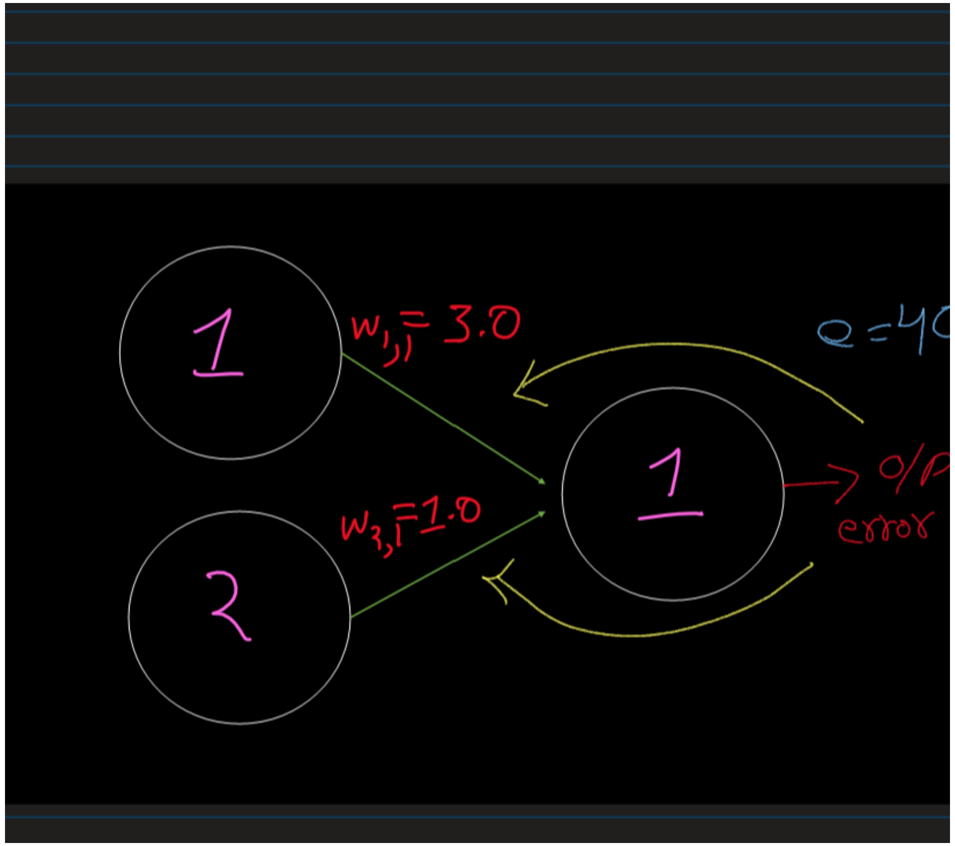
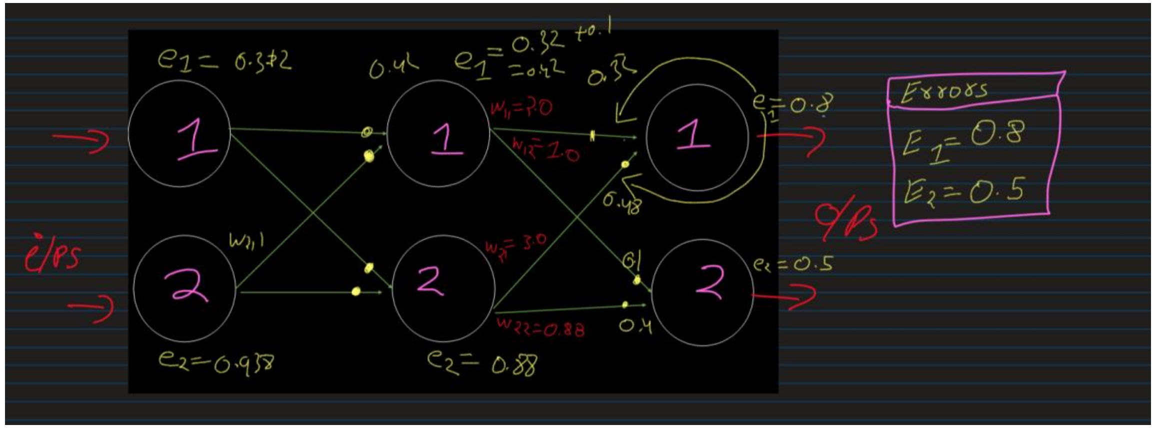
Deep Voyage 2 – Sliding with Newton and Gottfried

By

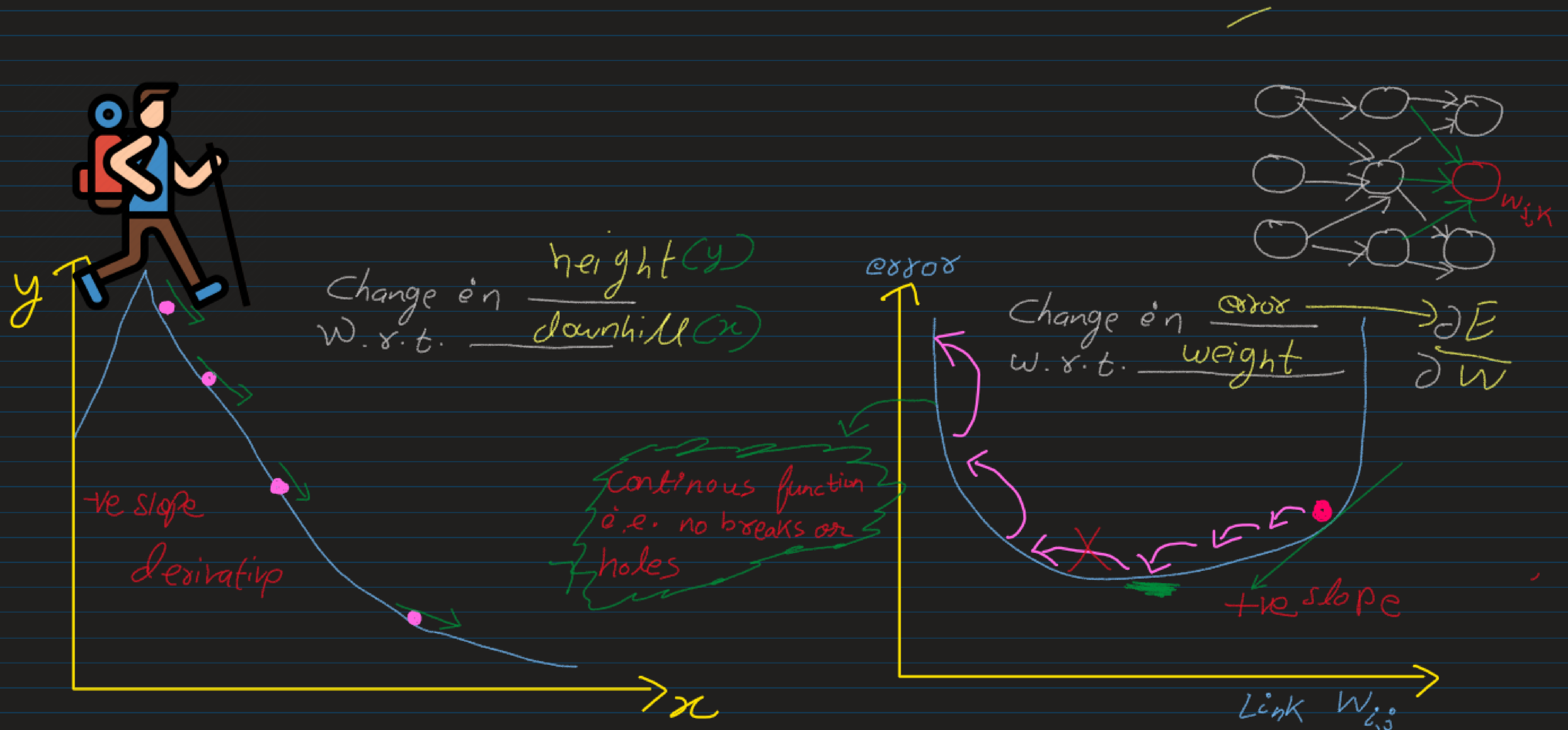
Jaspreet Singh

Recap of 1st session and calculation of error





Gradient Descent & chain Rule



w.r.t "O_k"

$$\frac{\partial E}{\partial W_{jk}} = \frac{\partial}{\partial O_k}(t_k - O_k)^2 \times \frac{\partial O_k}{\partial w_{jk}}$$

$$= 2(t_k - O_k) \cdot \frac{\partial}{\partial O_k}(t_k - O_k)$$

$$= 2(t_k - O_k) \cdot \left[\frac{\partial}{\partial O_k} t_k - \frac{\partial}{\partial O_k} O_k\right]$$

$$\frac{\partial f(x)}{\partial x} = f'(x)$$

$$\frac{\partial}{\partial x} f[g(x)] = f'[g(x)] \cdot \frac{\partial}{\partial x} g(x)$$

$\forall O_k$ = Sigmoid Activation Function

Slope

- Determination of rate of change of error, w.r.t. rate of change of link weights

$$= 2(t_k - O_k) \cdot \left[0 - \frac{\partial}{\partial O_k} \text{sigmoid}\left(\sum_j W_{jk} \cdot O_j\right)\right]$$

$$\frac{\partial}{\partial x} \sin x = \cos x$$

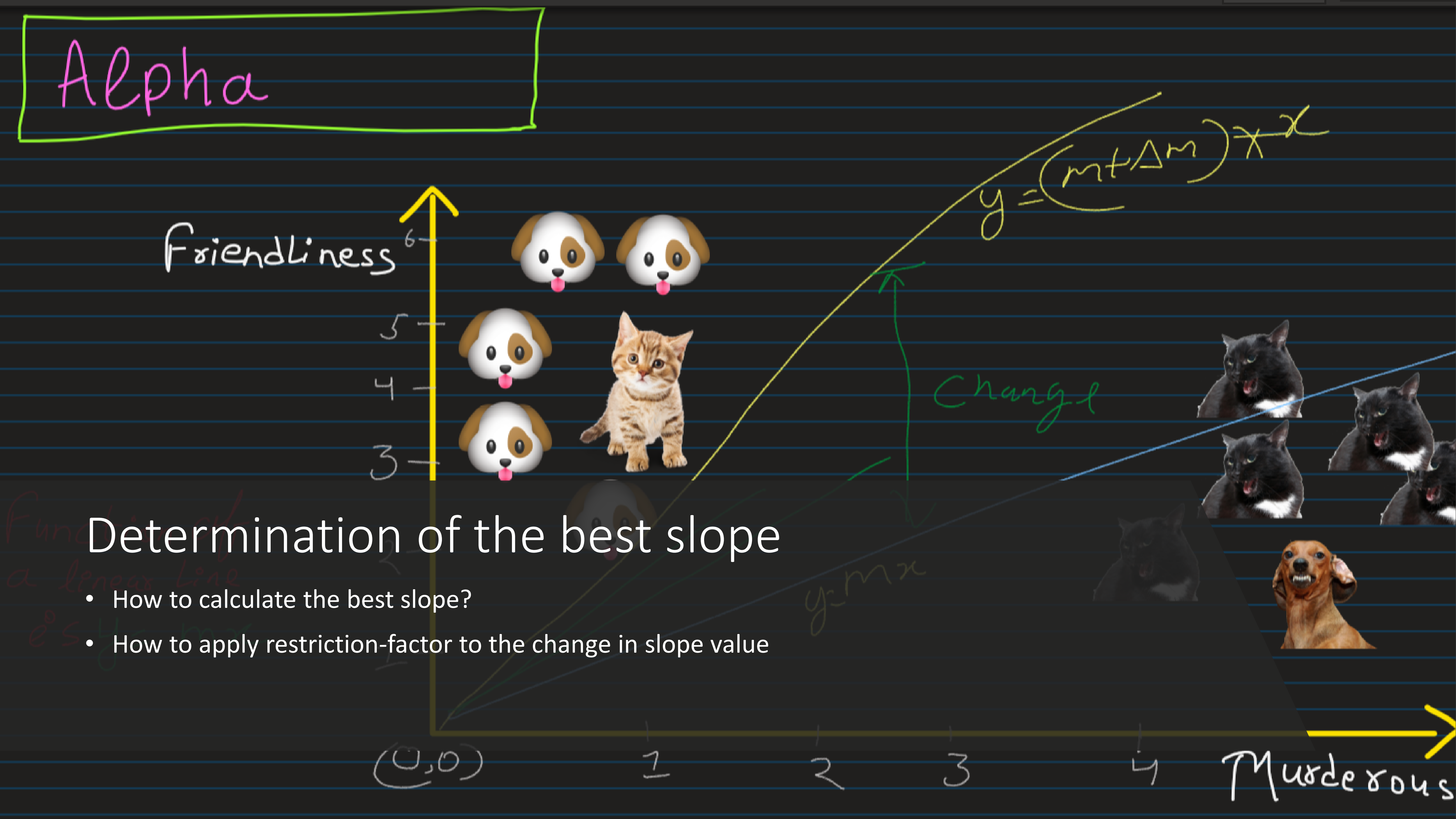
$$\frac{\partial}{\partial x} \text{sigmoid}(x) = \text{sigmoid}(x) \cdot (1 - \text{sig}(x))$$

$$= -2(t_k - O_k) \cdot \text{sigmoid}\left(\sum_j W_{jk} \cdot O_j\right) \cdot \left(1 - \text{sigmoid}\left(\sum_j W_{jk} \cdot O_j\right)\right) \cdot \frac{\partial}{\partial O_k}\left[\sum_j W_{jk} \cdot O_j\right]$$

$$\frac{\partial E}{\partial W_{j,k}} = -(t_k - O_k) \cdot \text{sigmoid}\left(\sum_j W_{jk} \cdot O_j\right) \cdot \left[1 - \text{sigmoid}\left(\sum_j W_{jk} \cdot O_j\right)\right] \cdot O_j$$

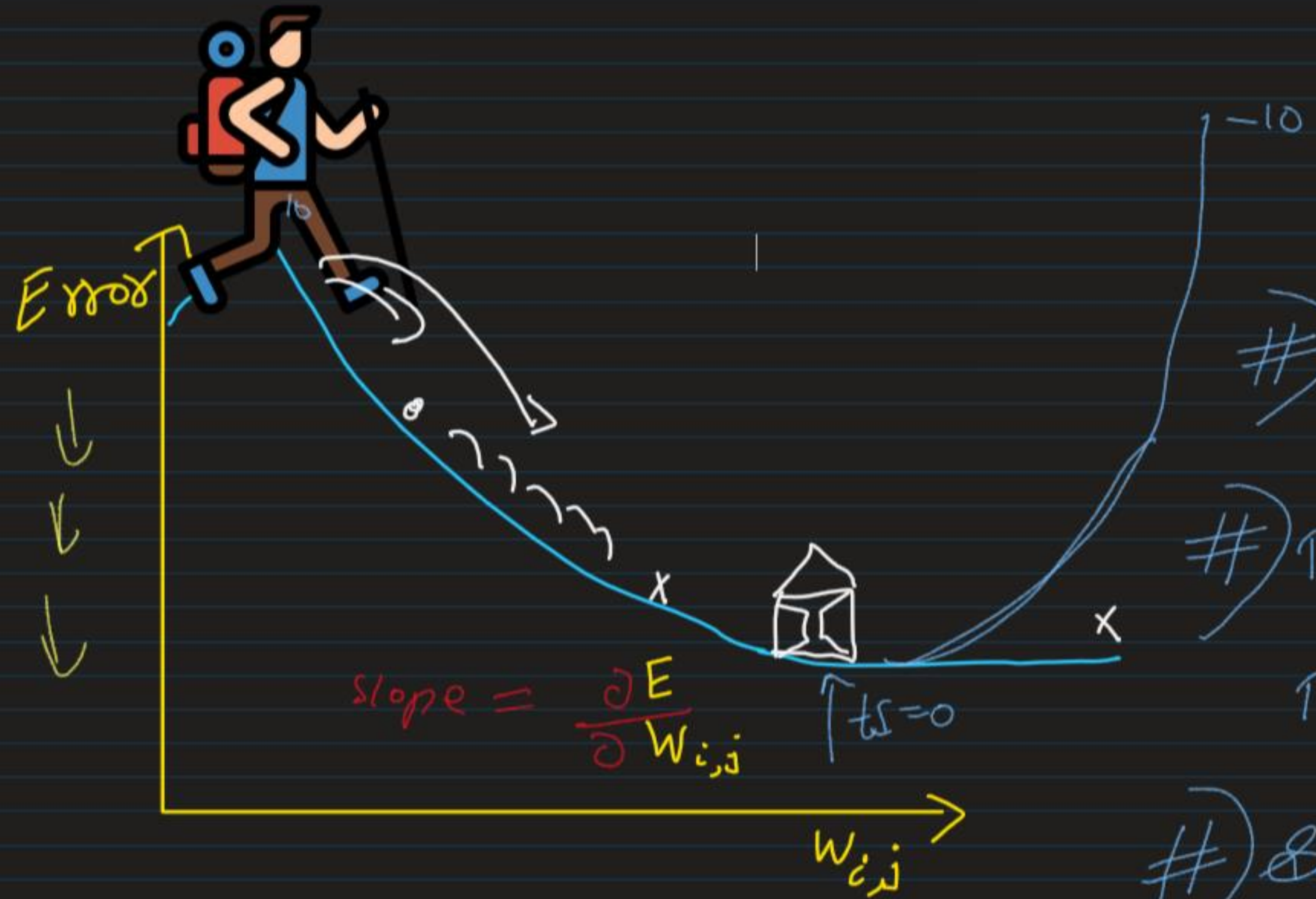
Slope

- Determination of rate of change of error, w.r.t. rate of change of link weights



Determination of the best slope

- How to calculate the best slope?
- How to apply restriction-factor to the change in slope value



Optimization Problem
↳ o/p needs to be maximized or minimized

Maximized value ⇒ When function stops increasing, and starts decreasing
Minimized value ⇒ When function stops decreasing, and starts increasing

So Optimization Problem = finding rate of change of a function at any point

?) But how to determine this [rate of change of a function at any point] ?

THANK YOU