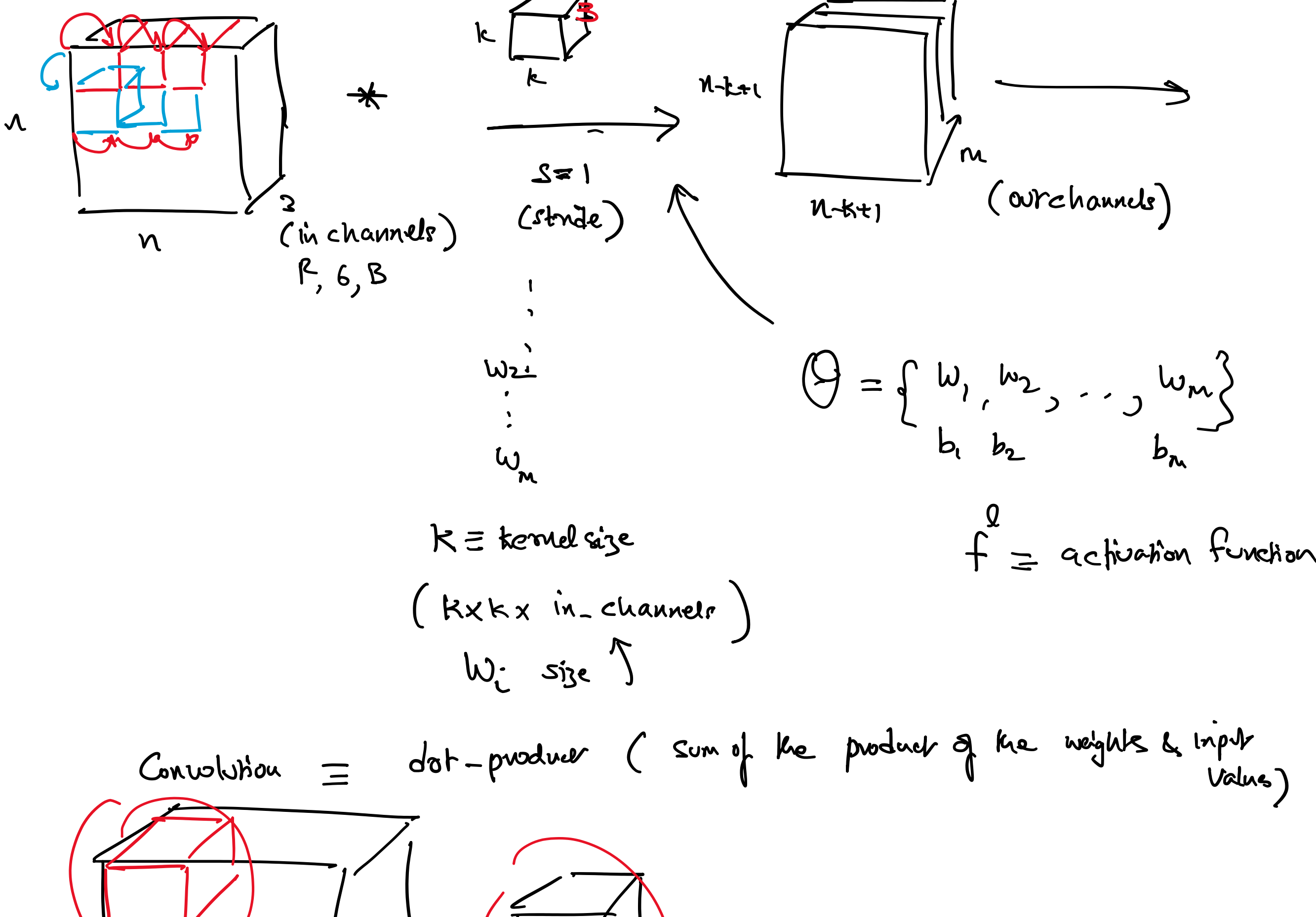
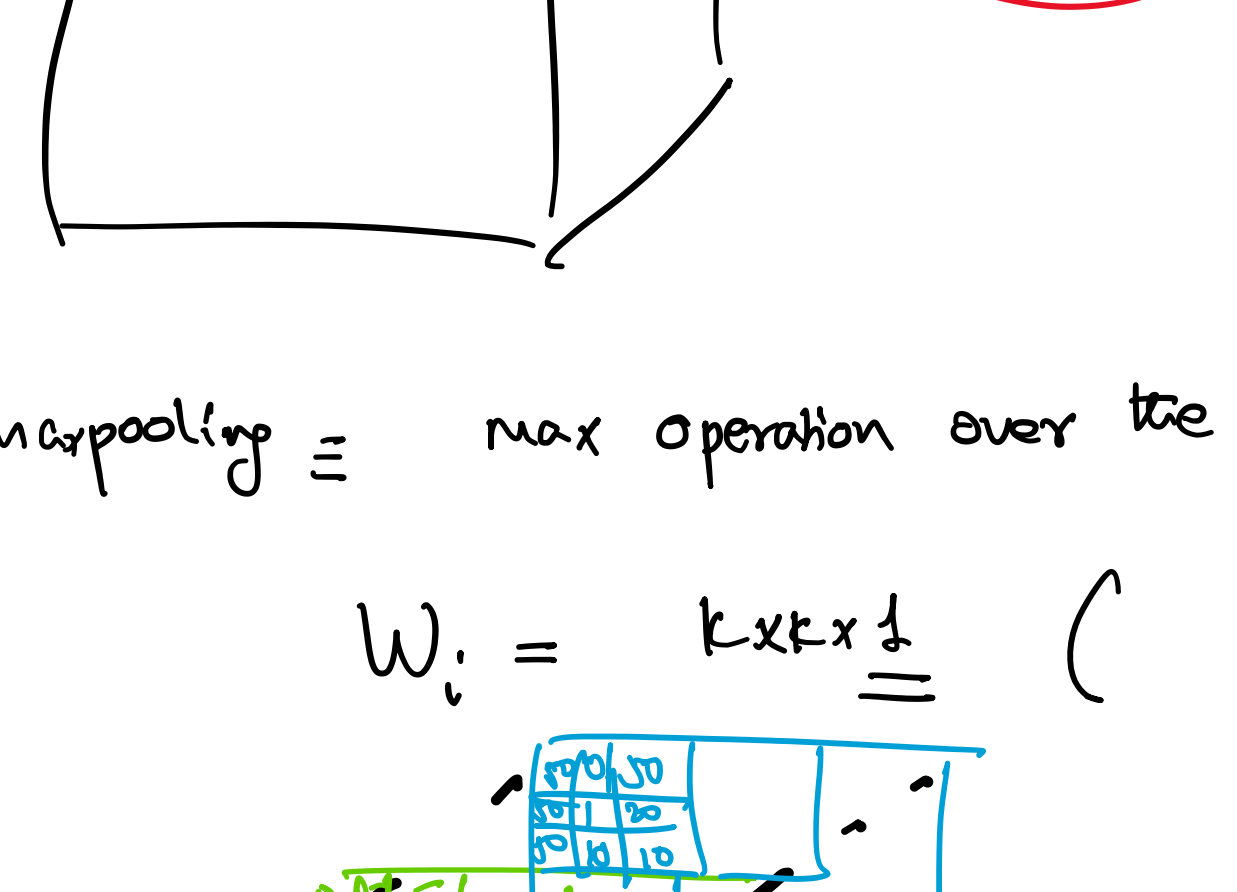


CNNs: Convolutional NN

feed forward, sparse, MLPs



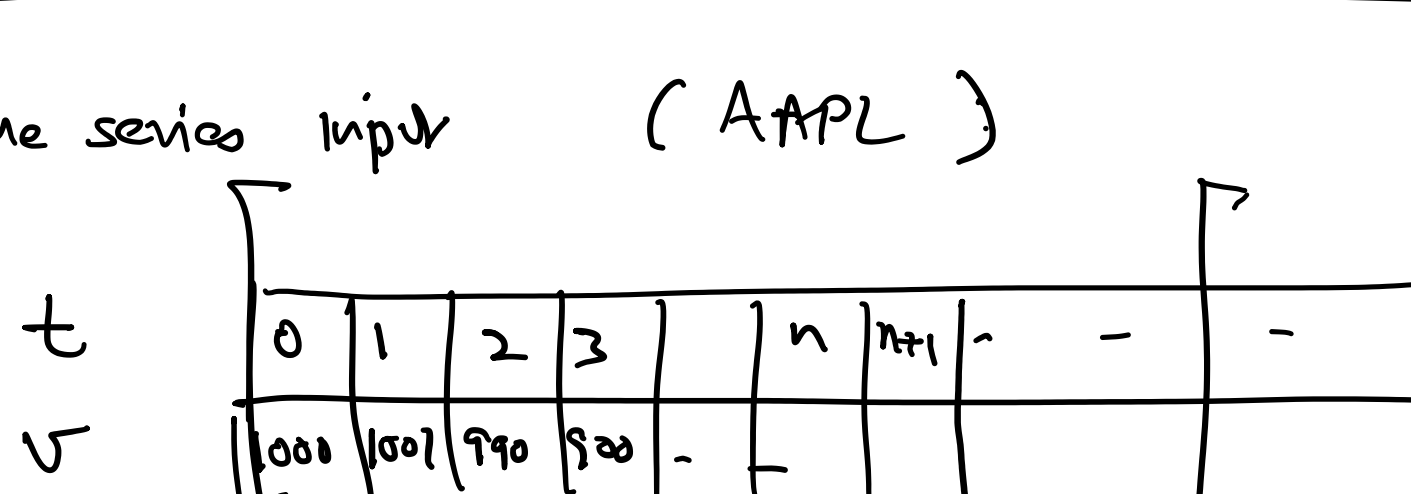
Convolution \equiv dot-product (sum of the product of the weights & input values)



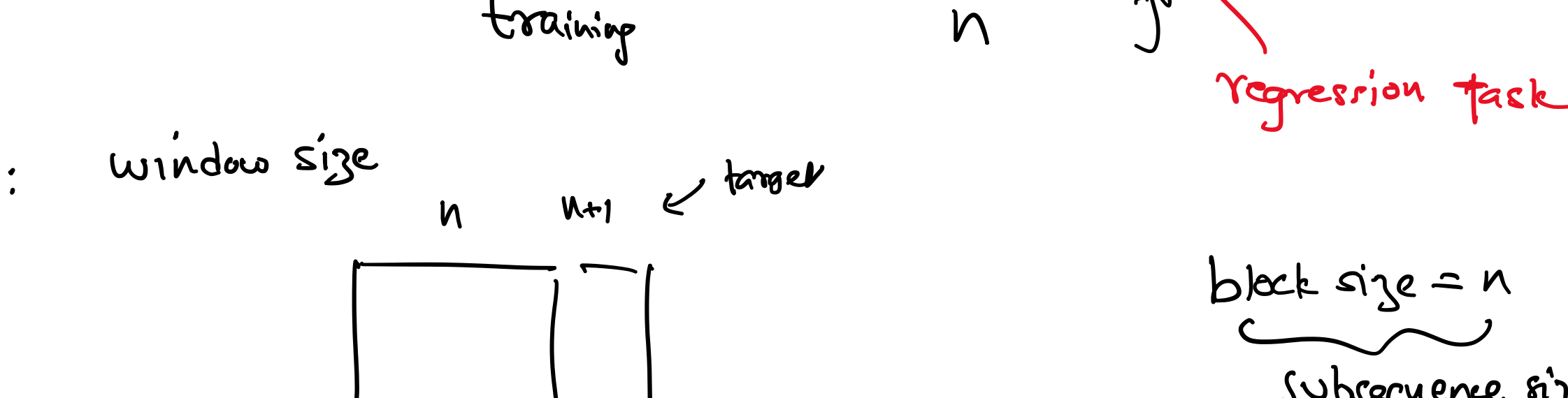
maxpooling \equiv max operation over the input value

$$W_i = k \times k \times 1 \quad (w_{ij} = 1)$$

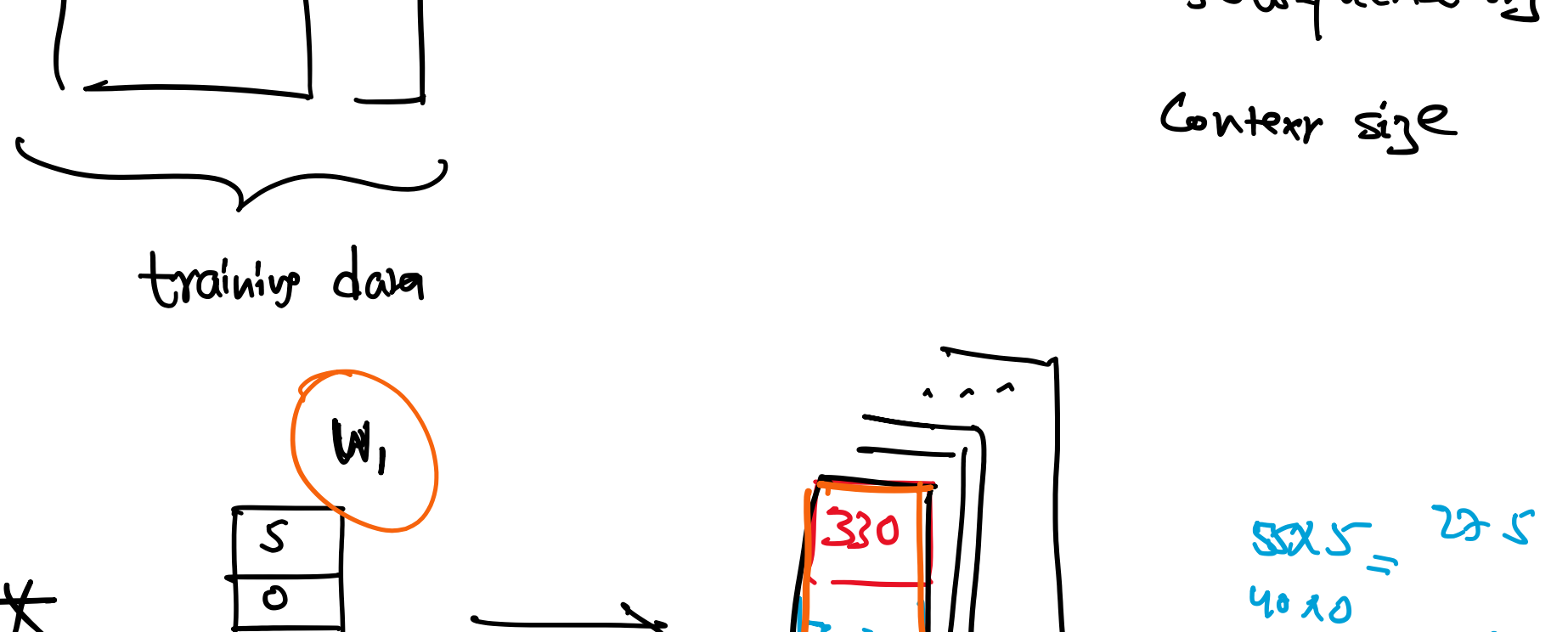
all entries of w_i are 1



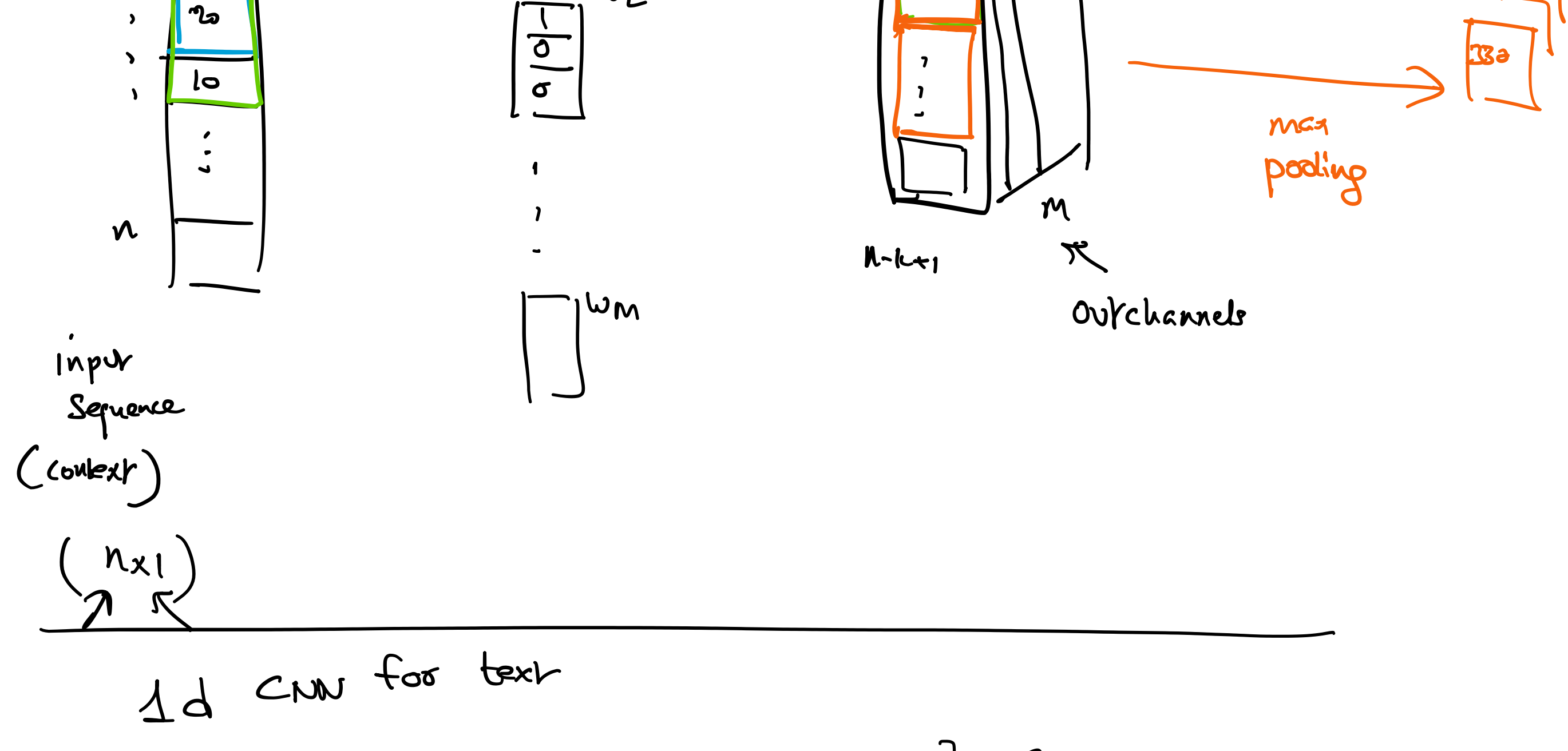
time series input (AAPL)



u: window size



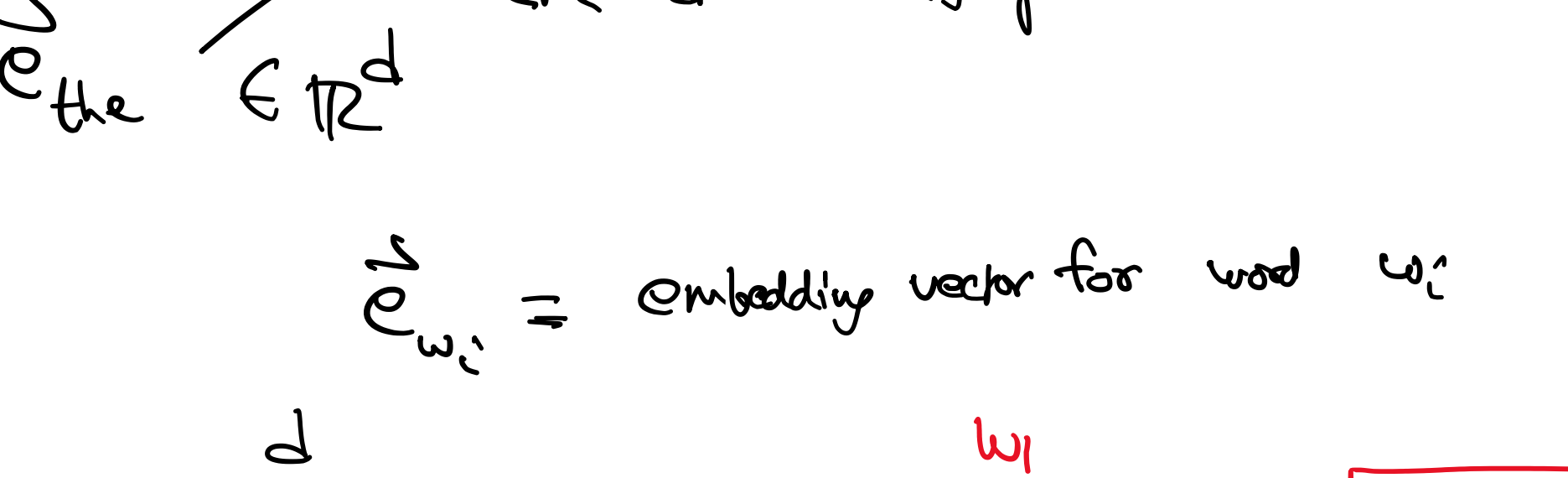
input channels = 1



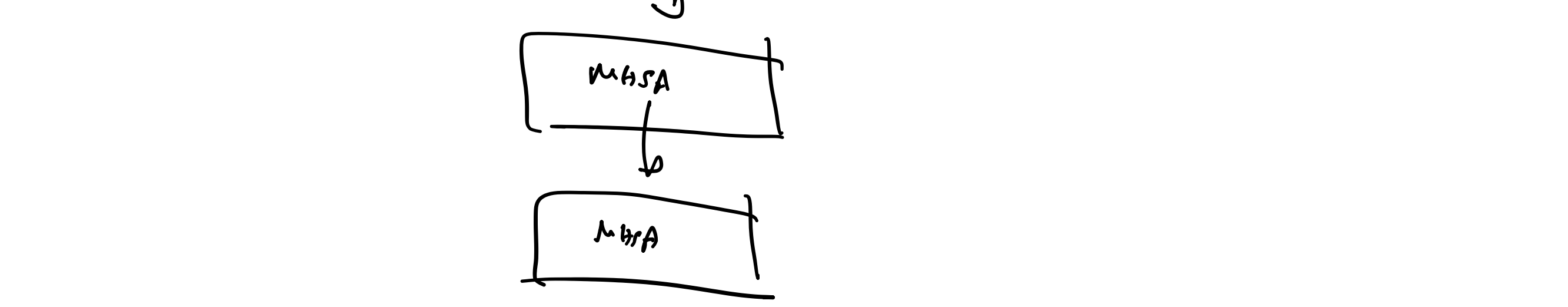
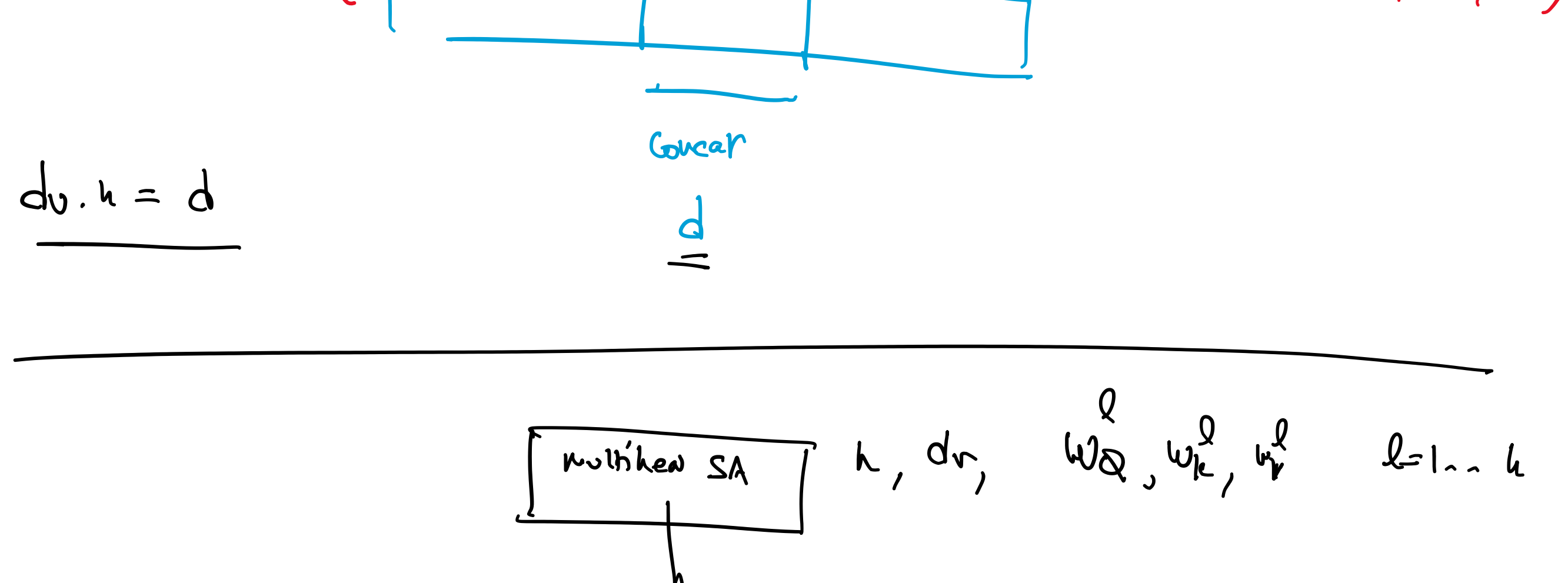
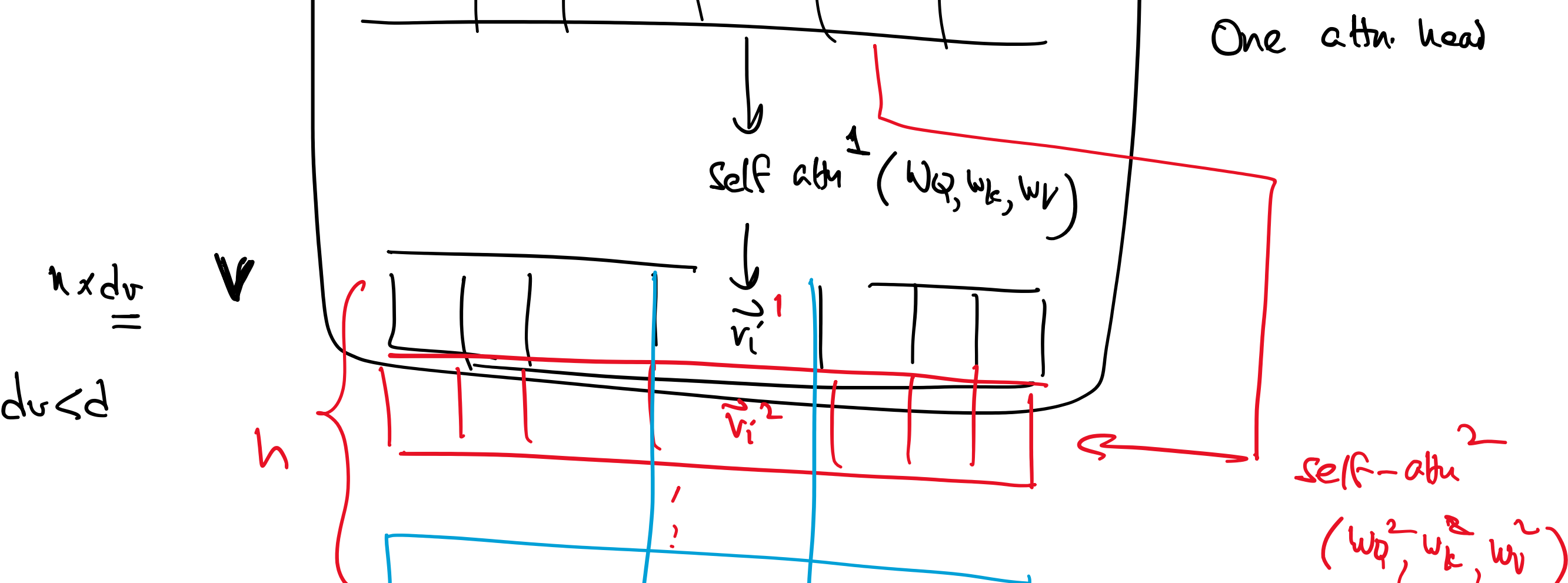
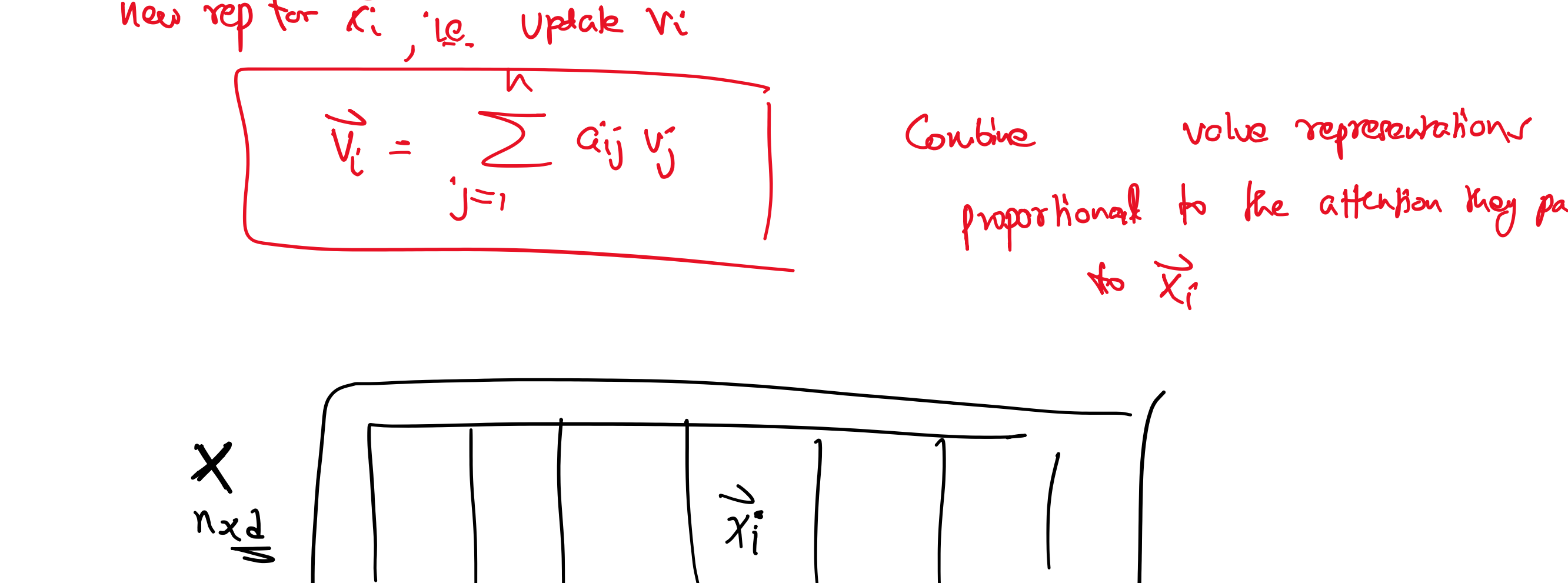
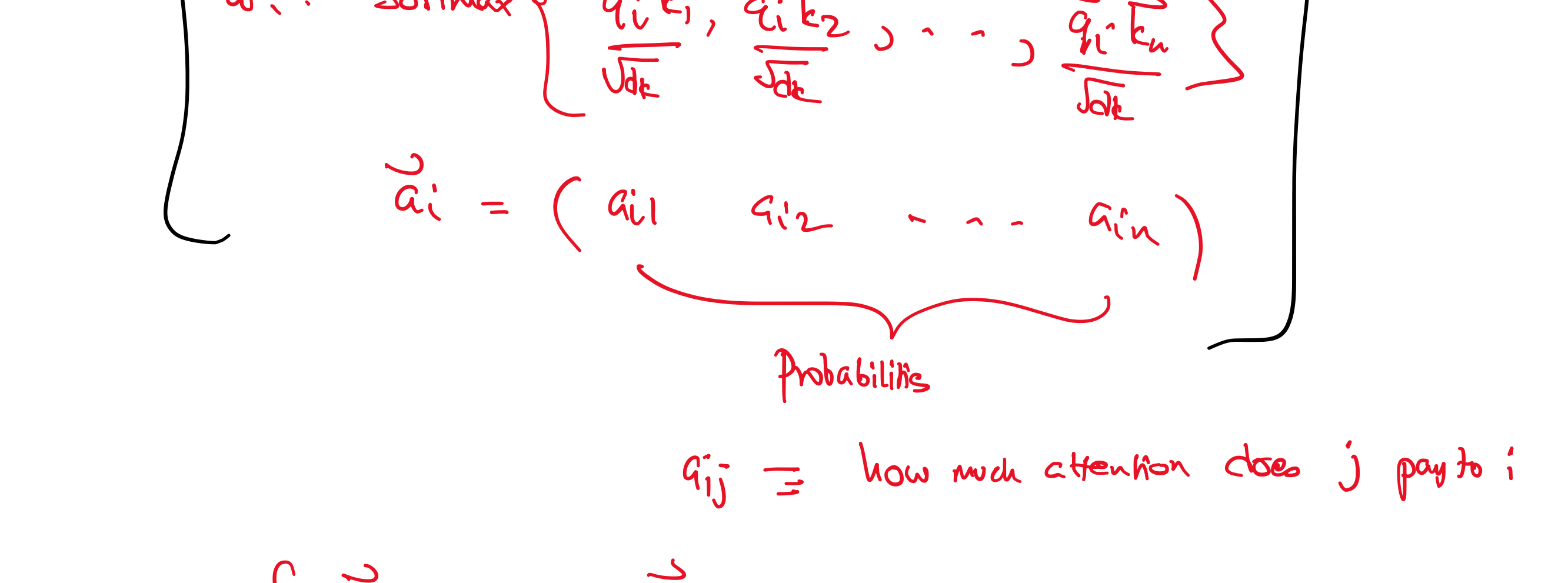
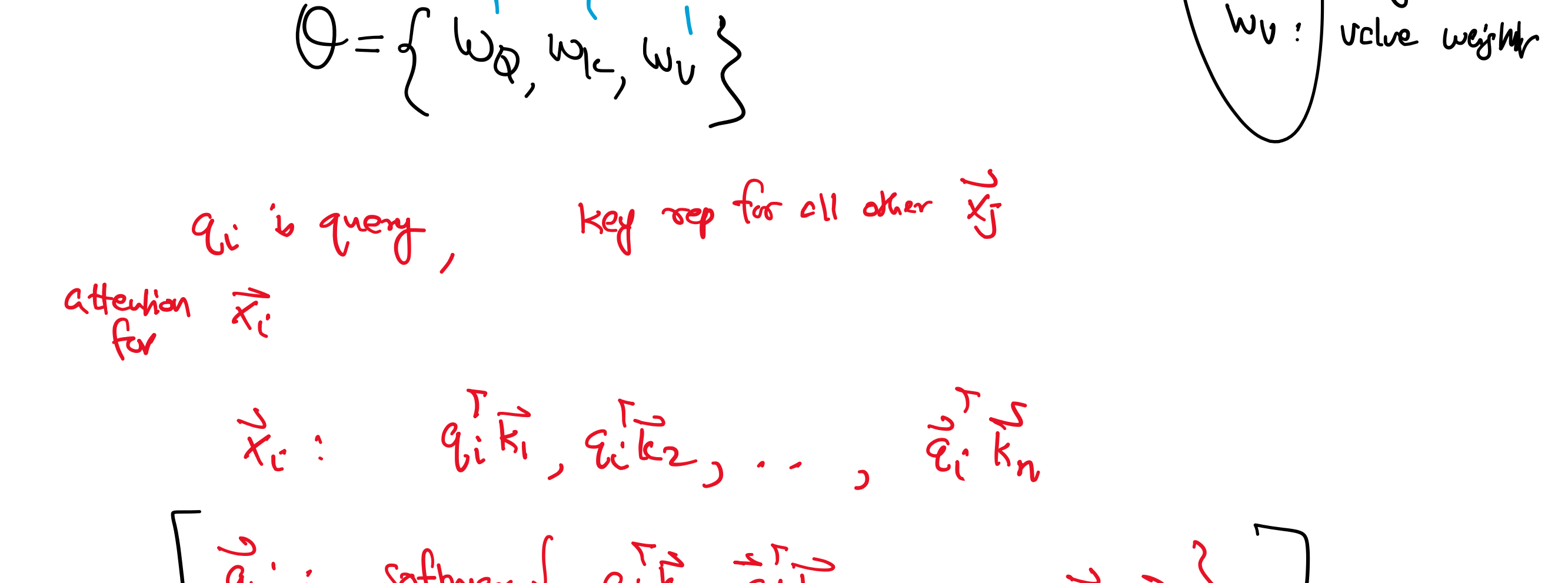
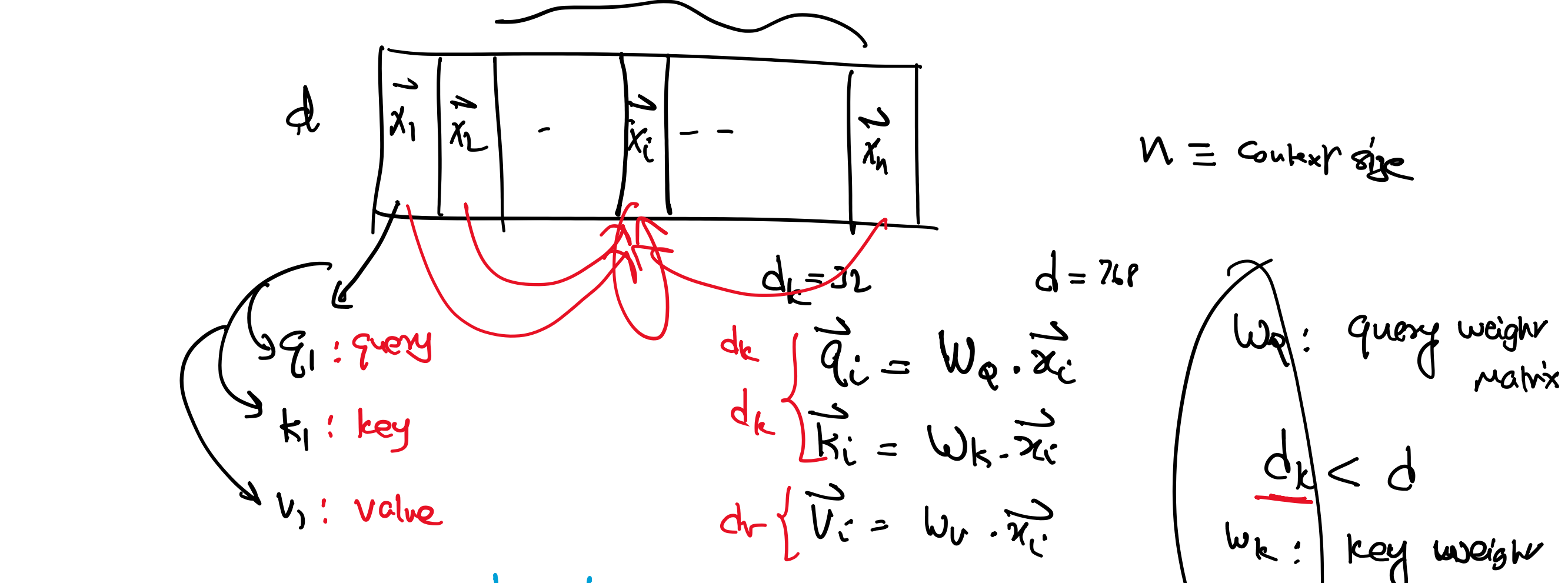
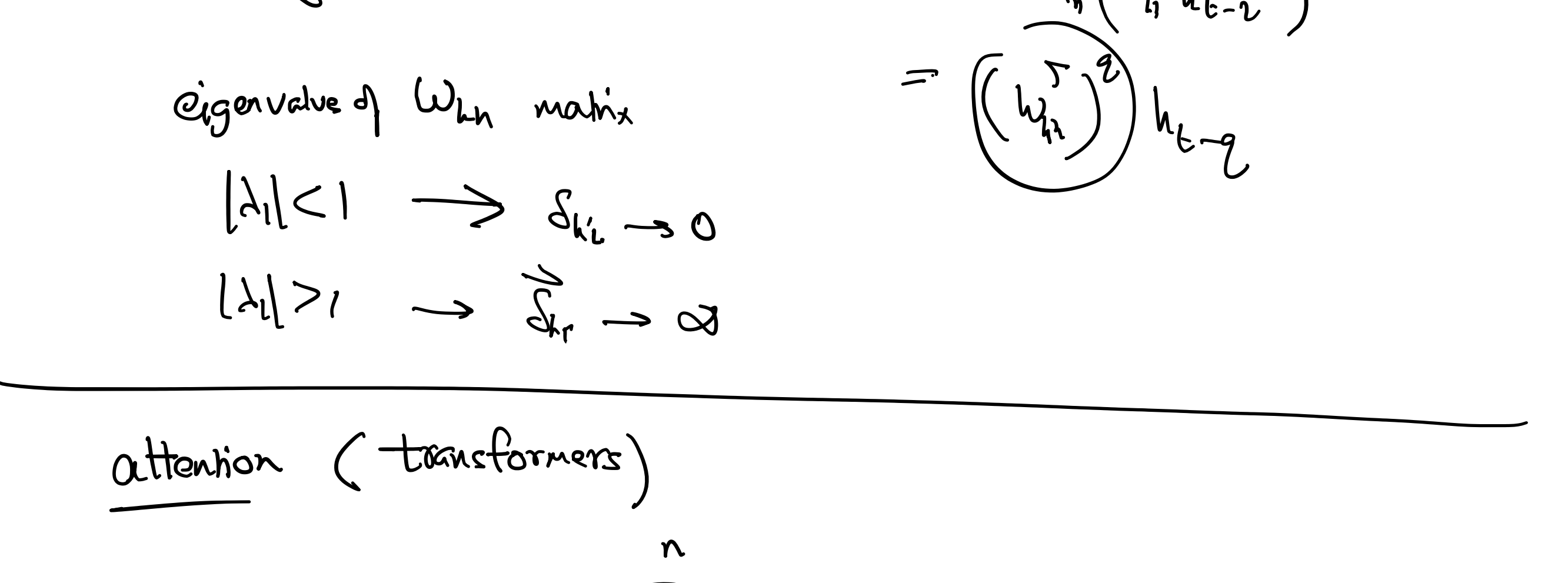
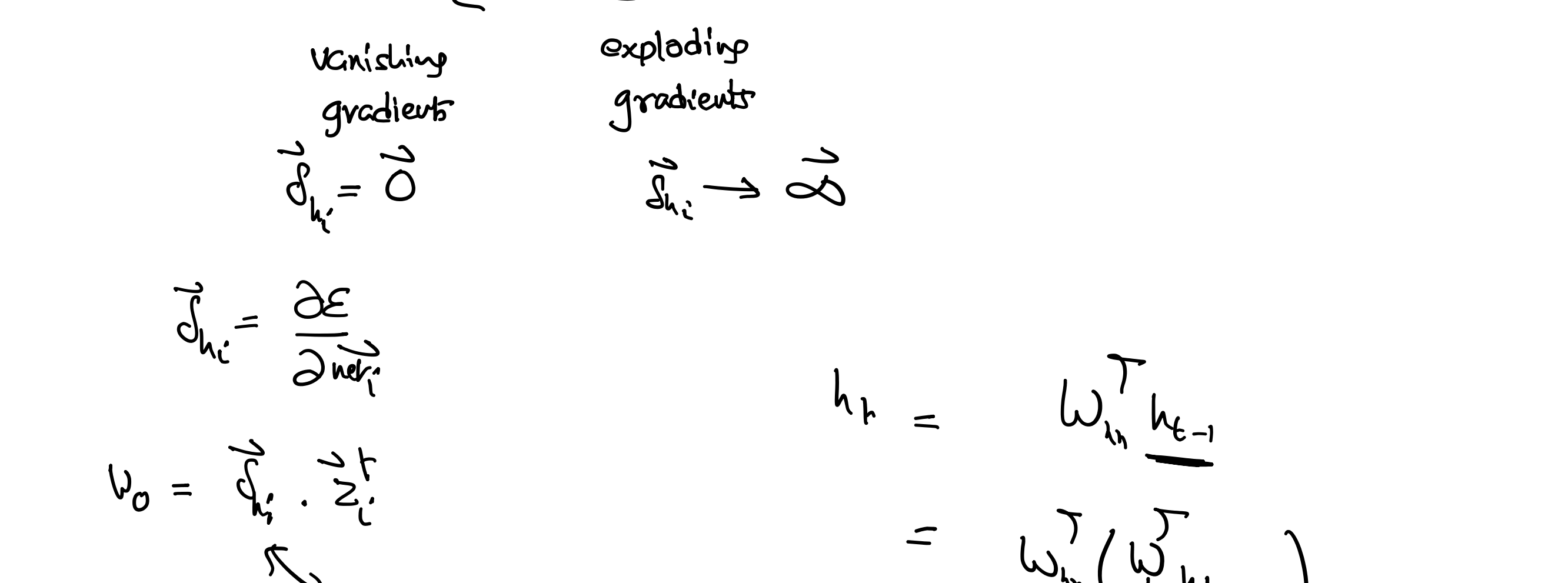
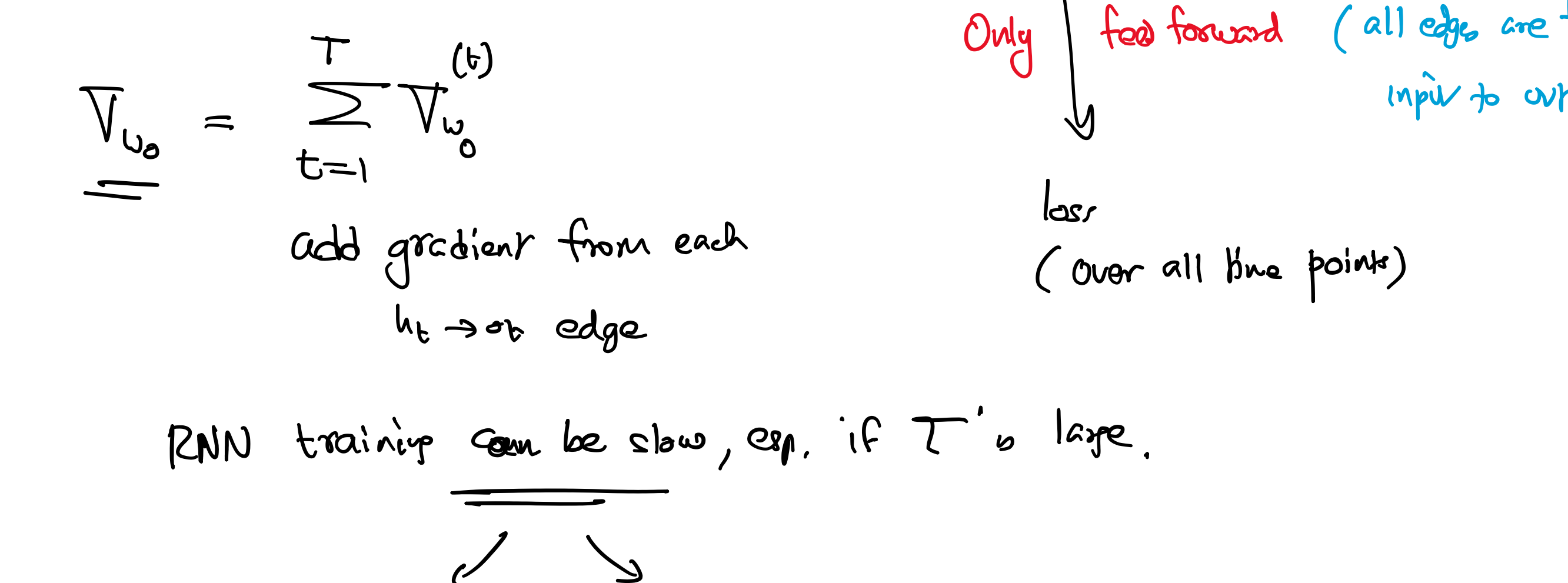
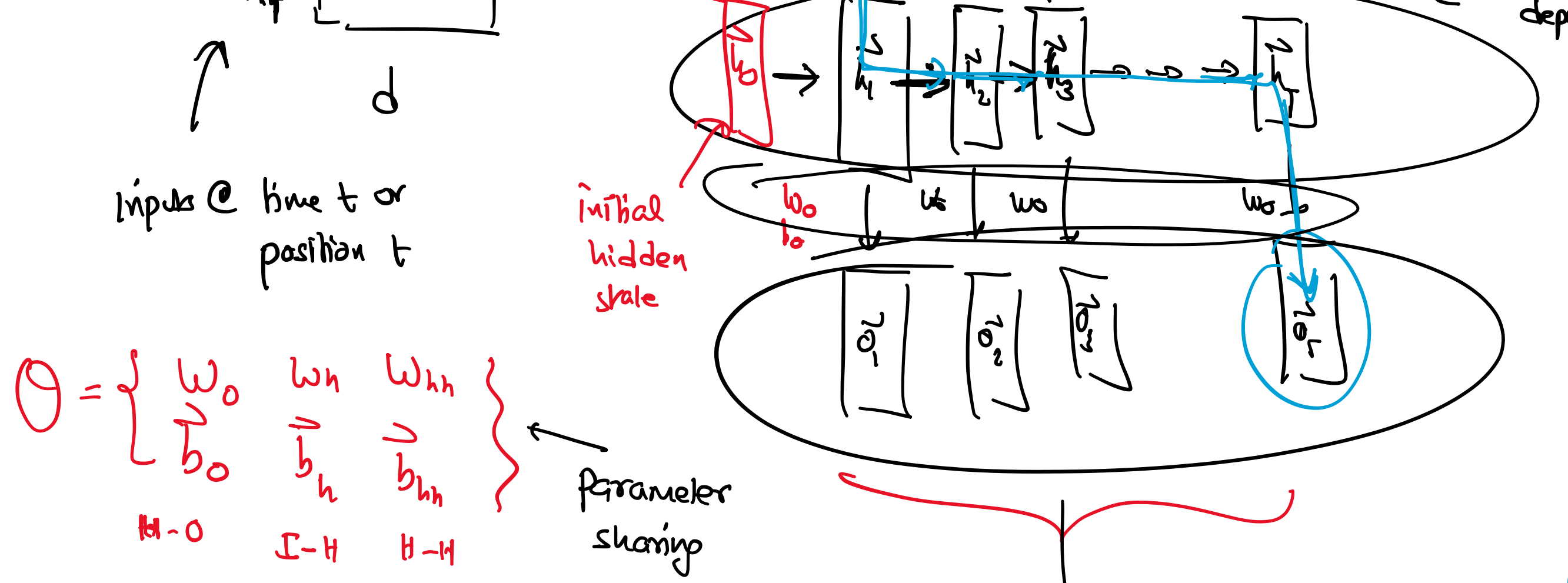
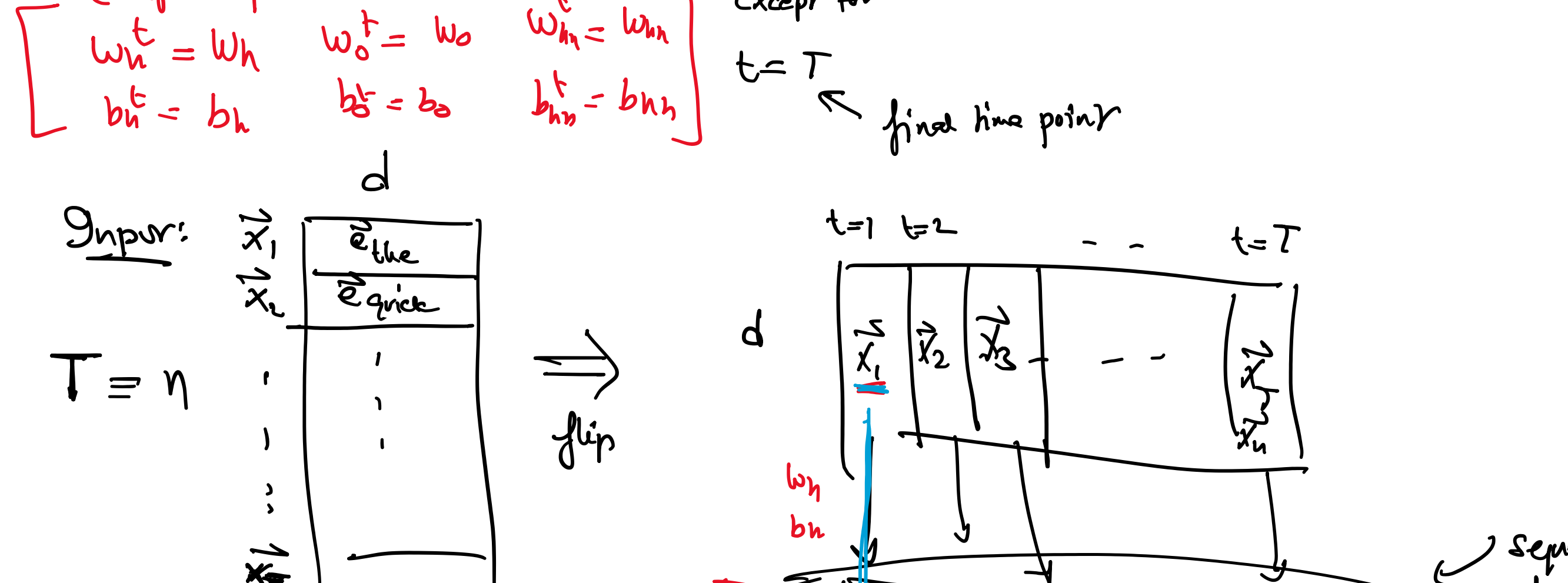
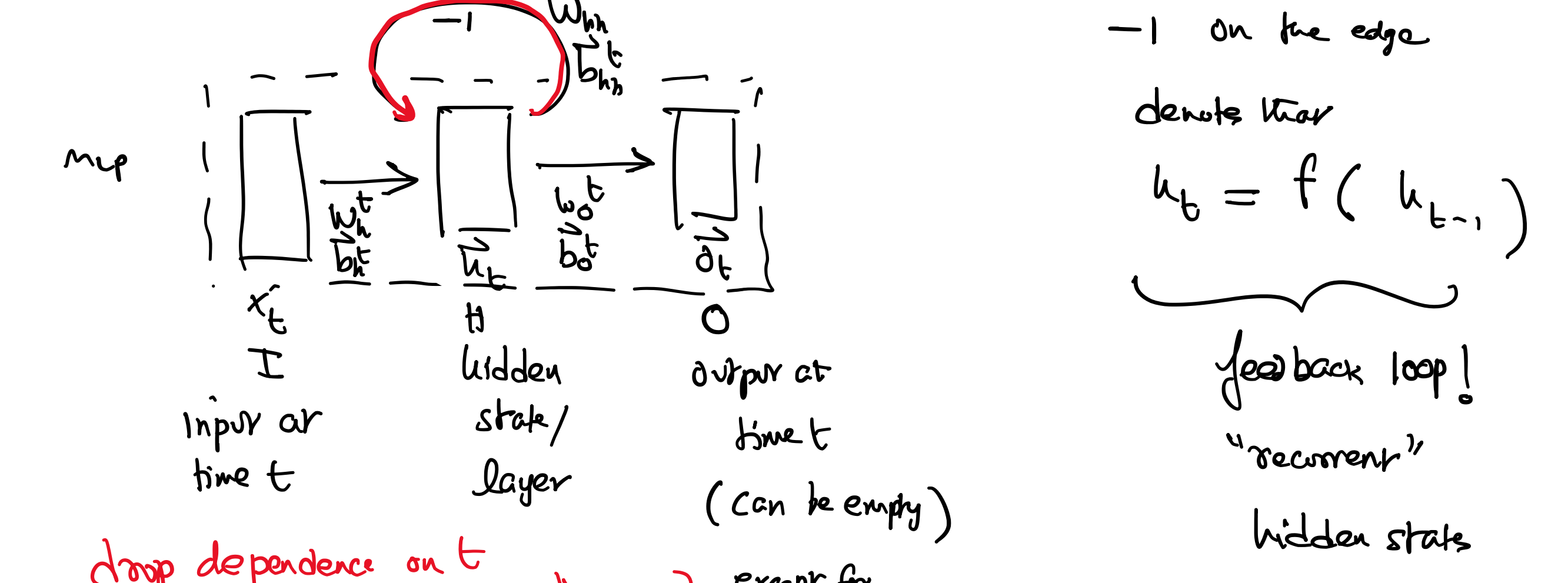
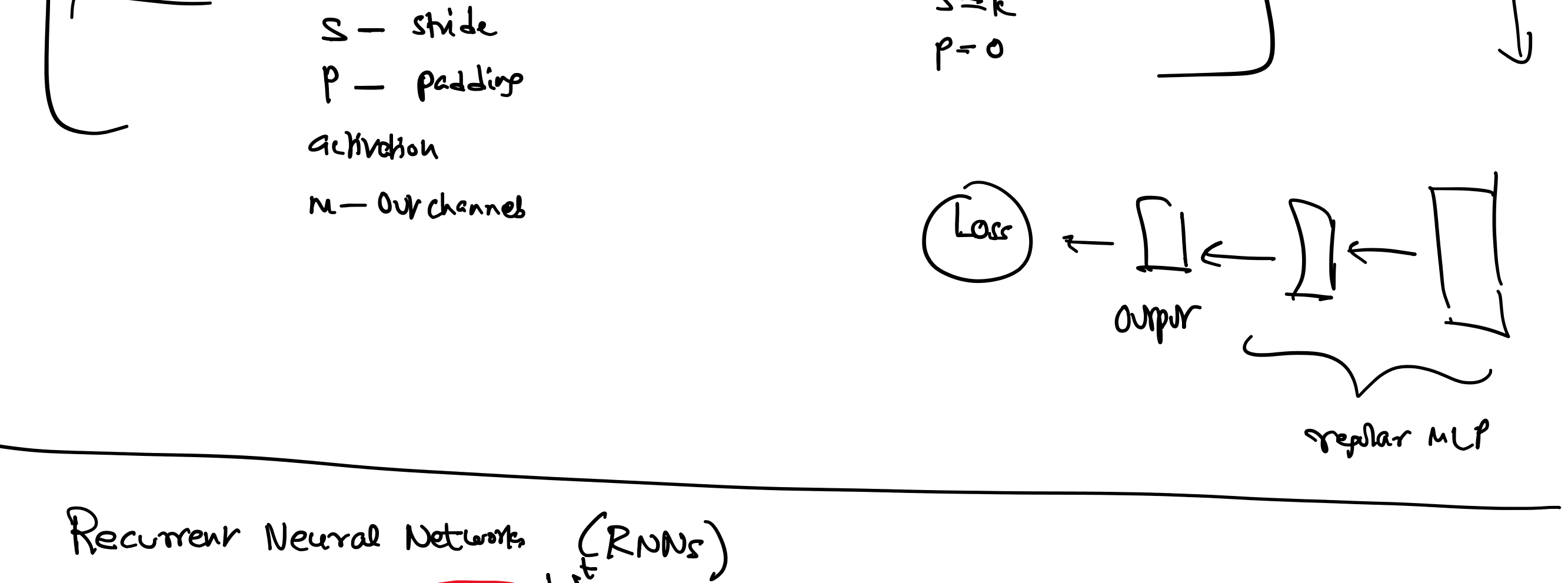
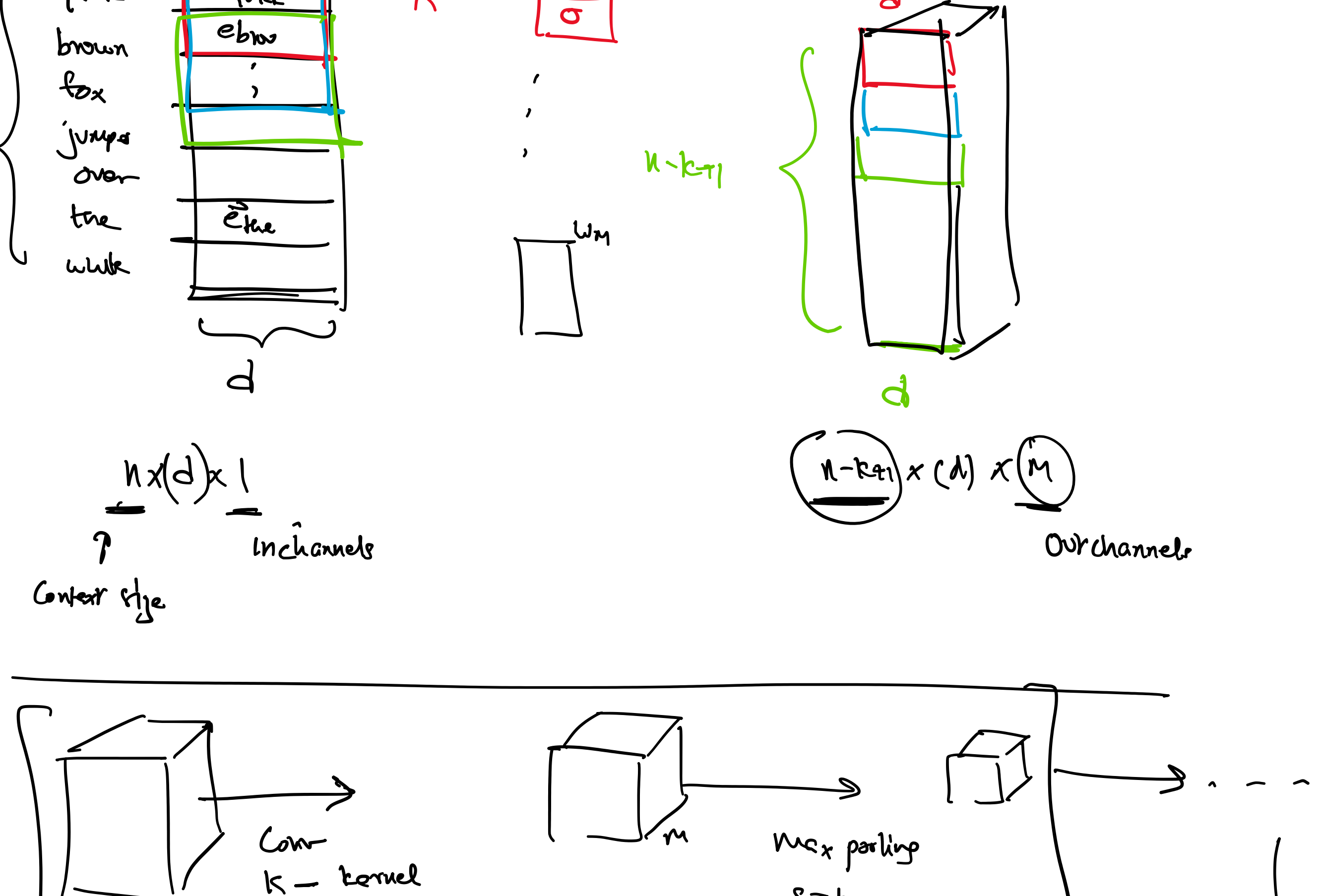
1d CNN for text

[the quick brown fox jumped over the wire] ?
what is the next word

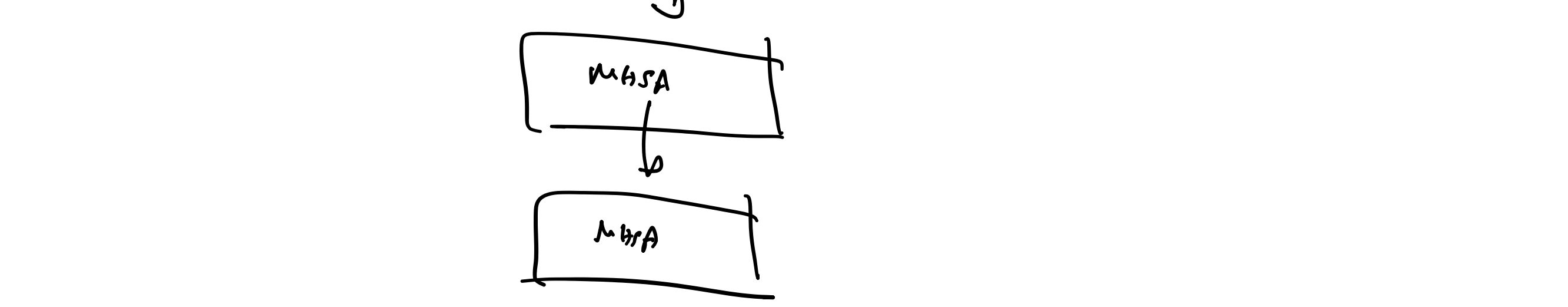
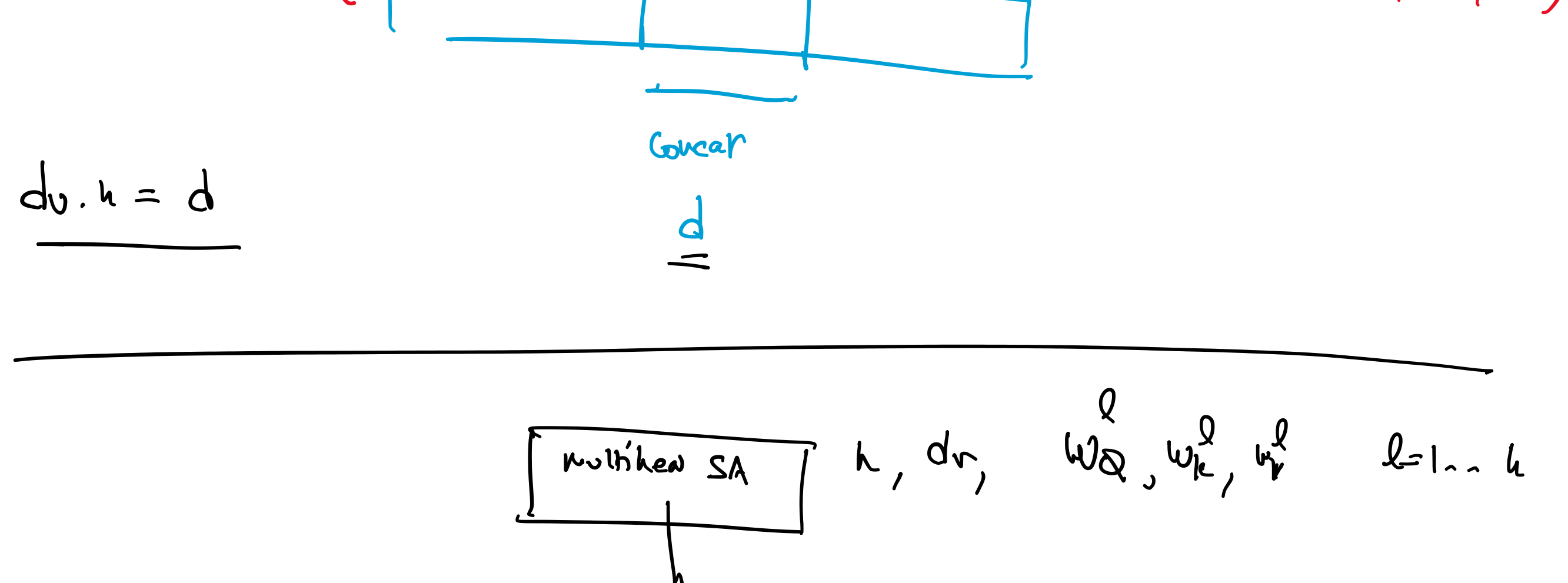
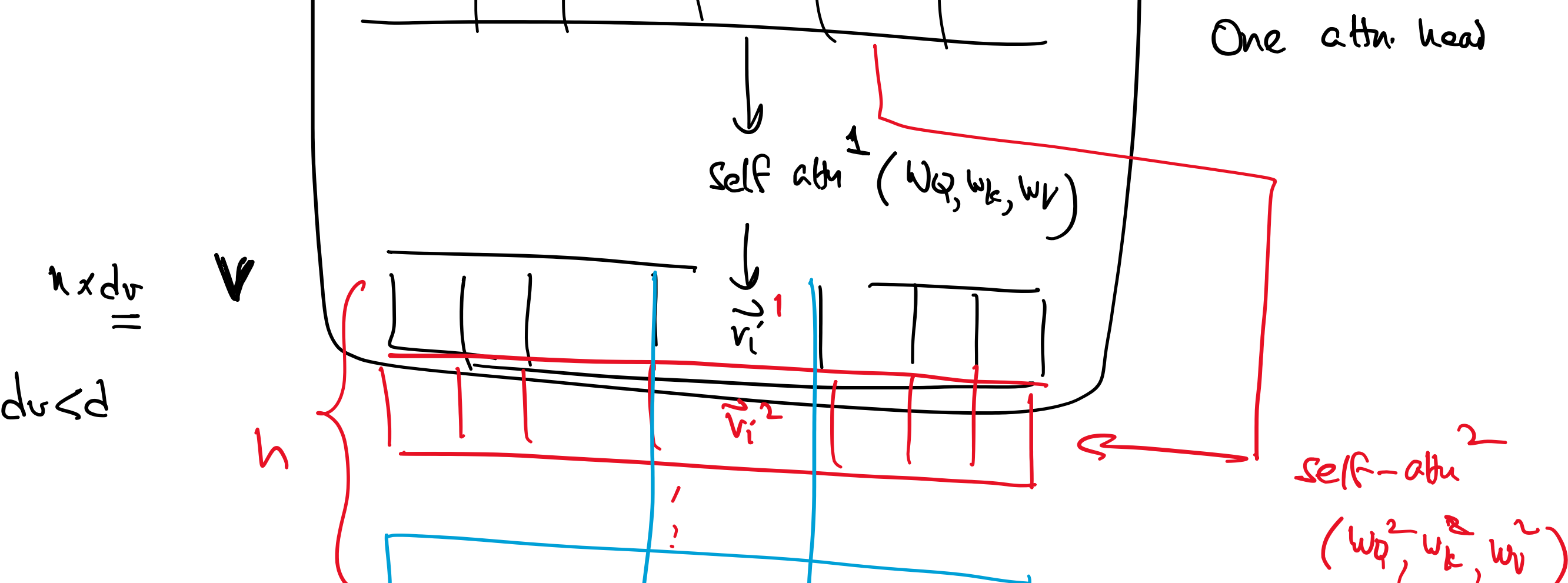
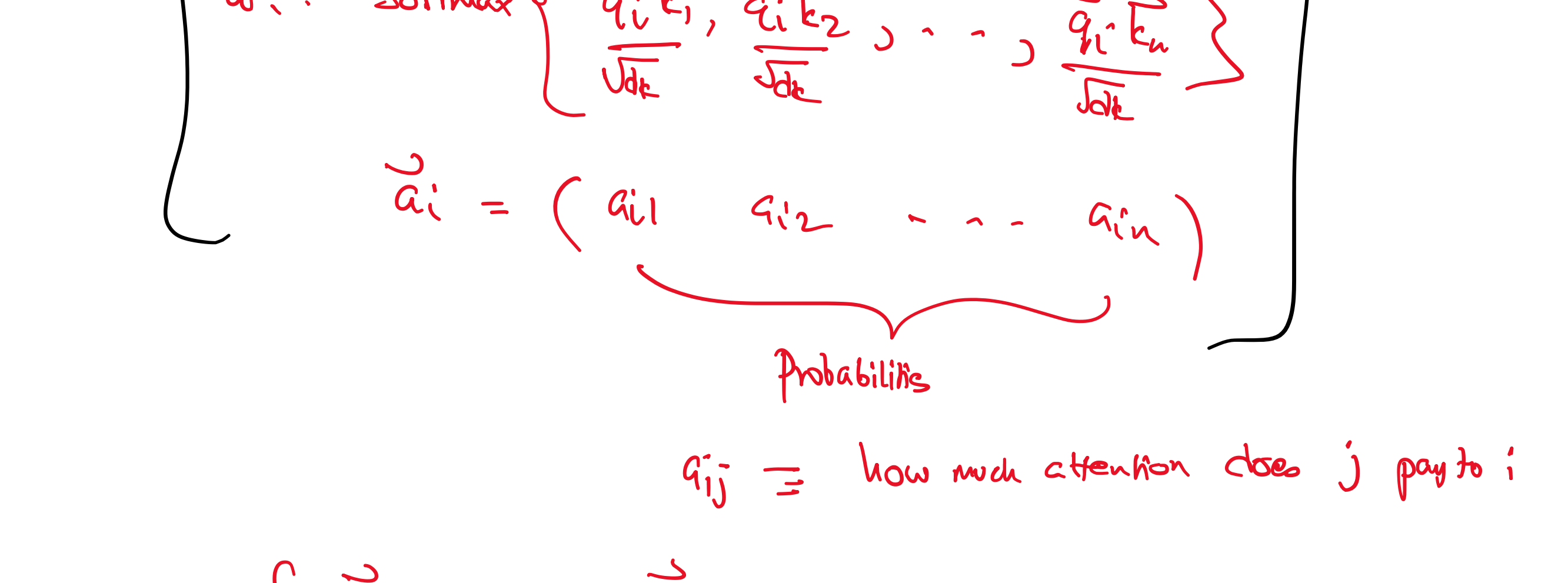
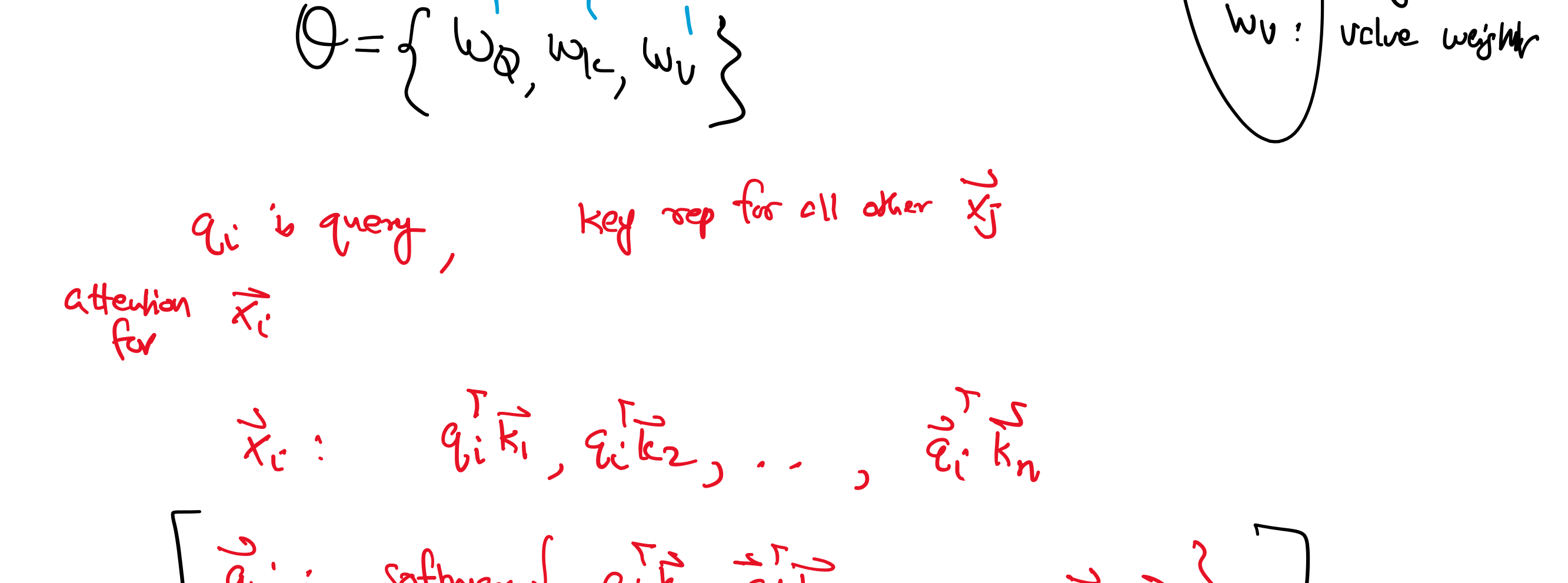
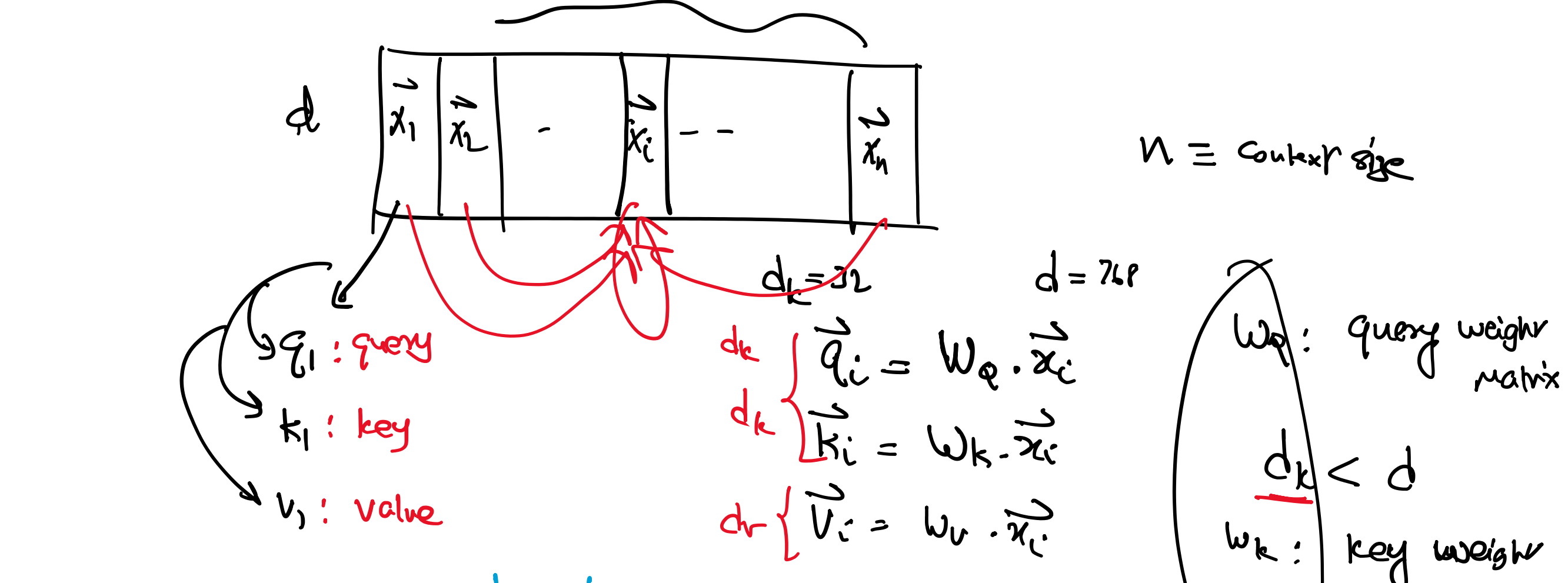
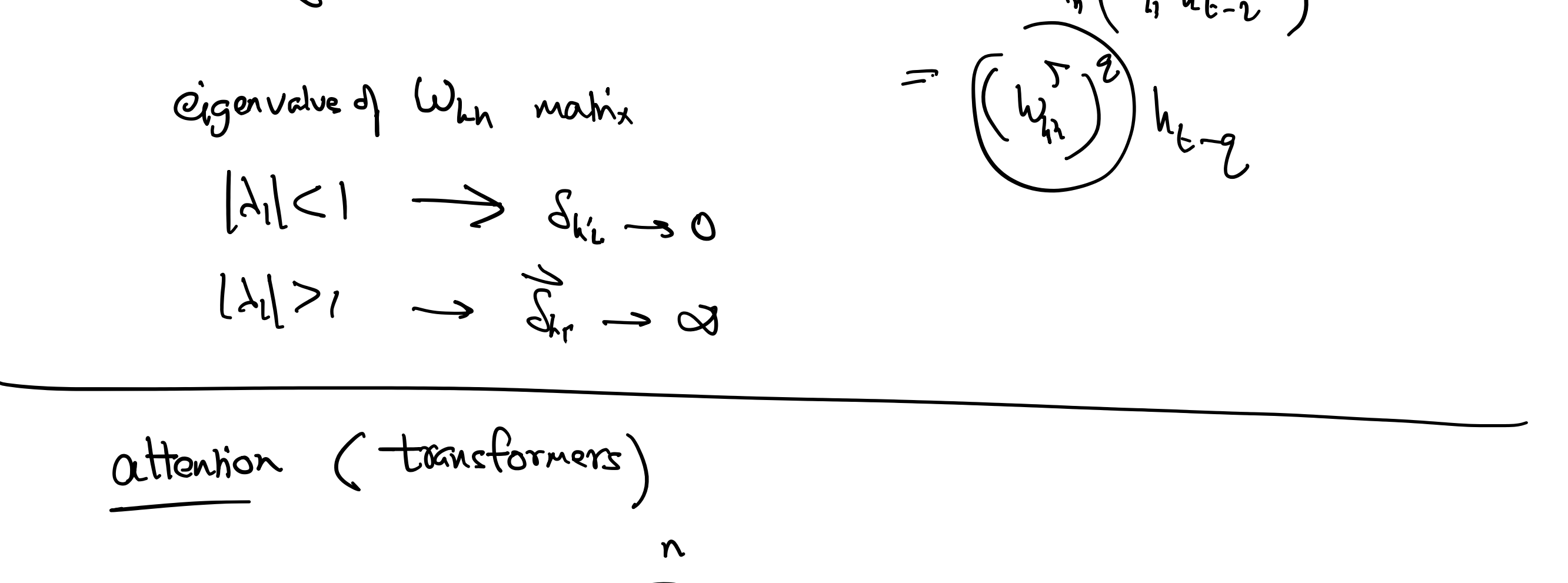
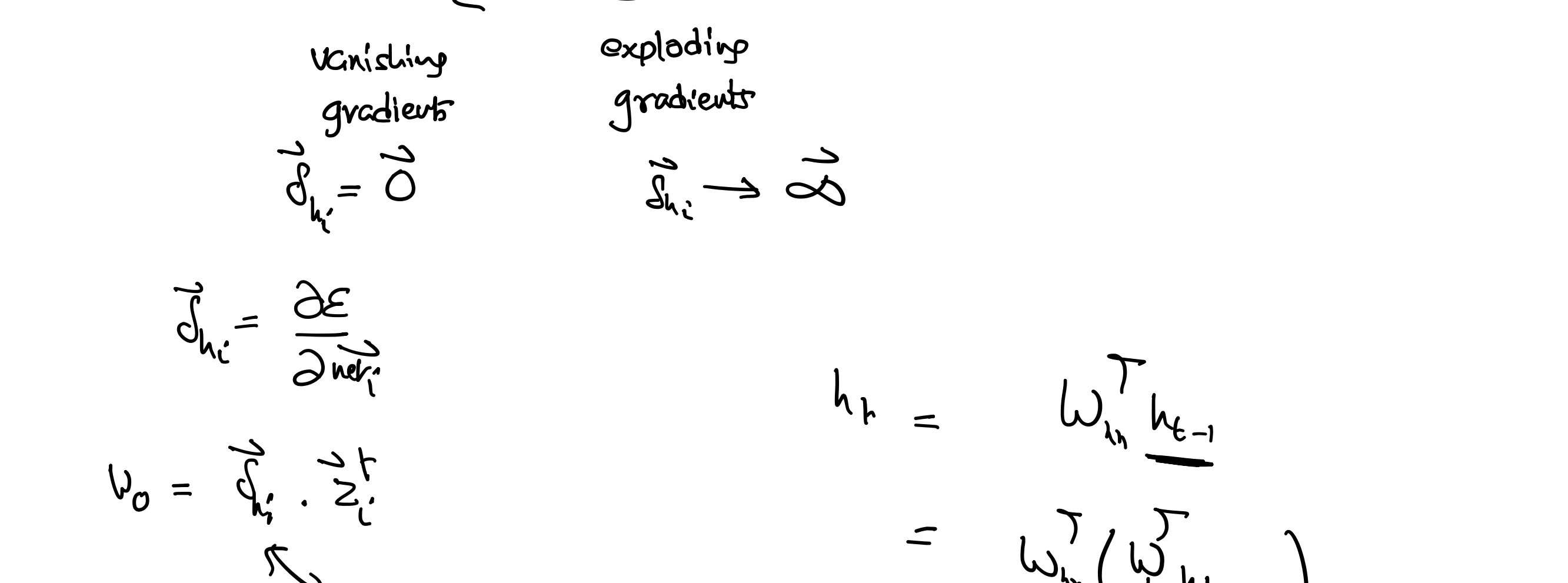
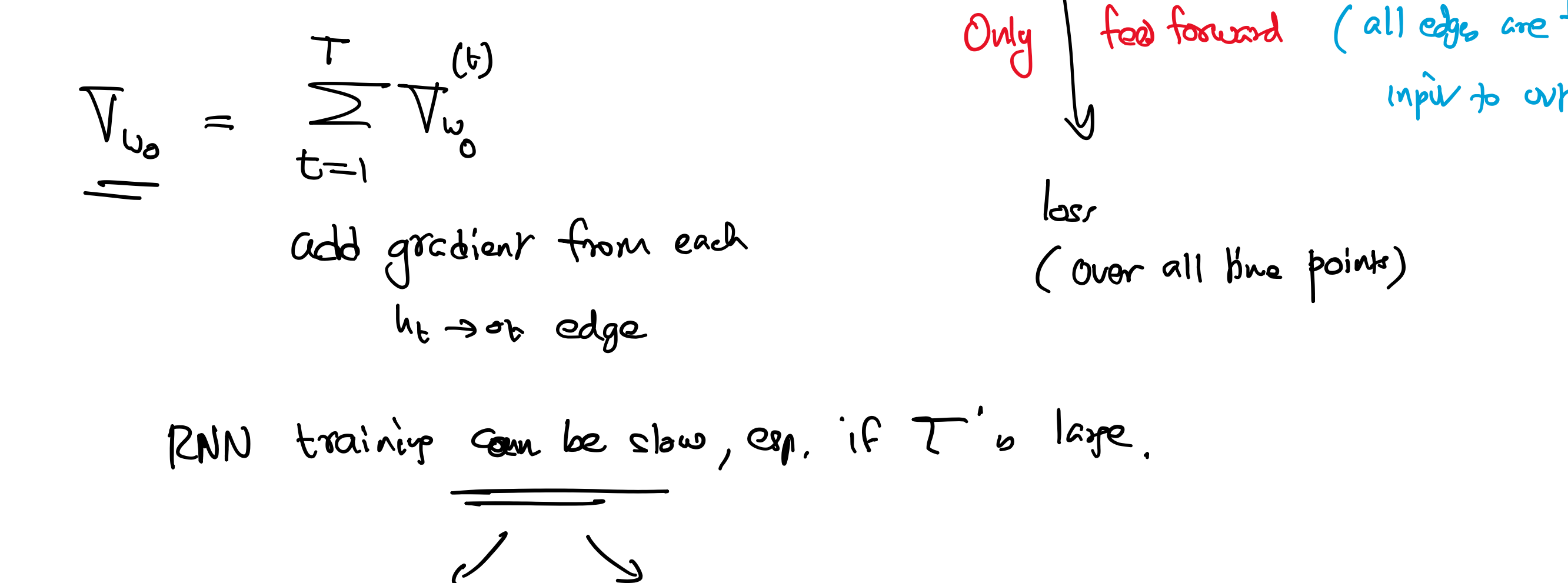
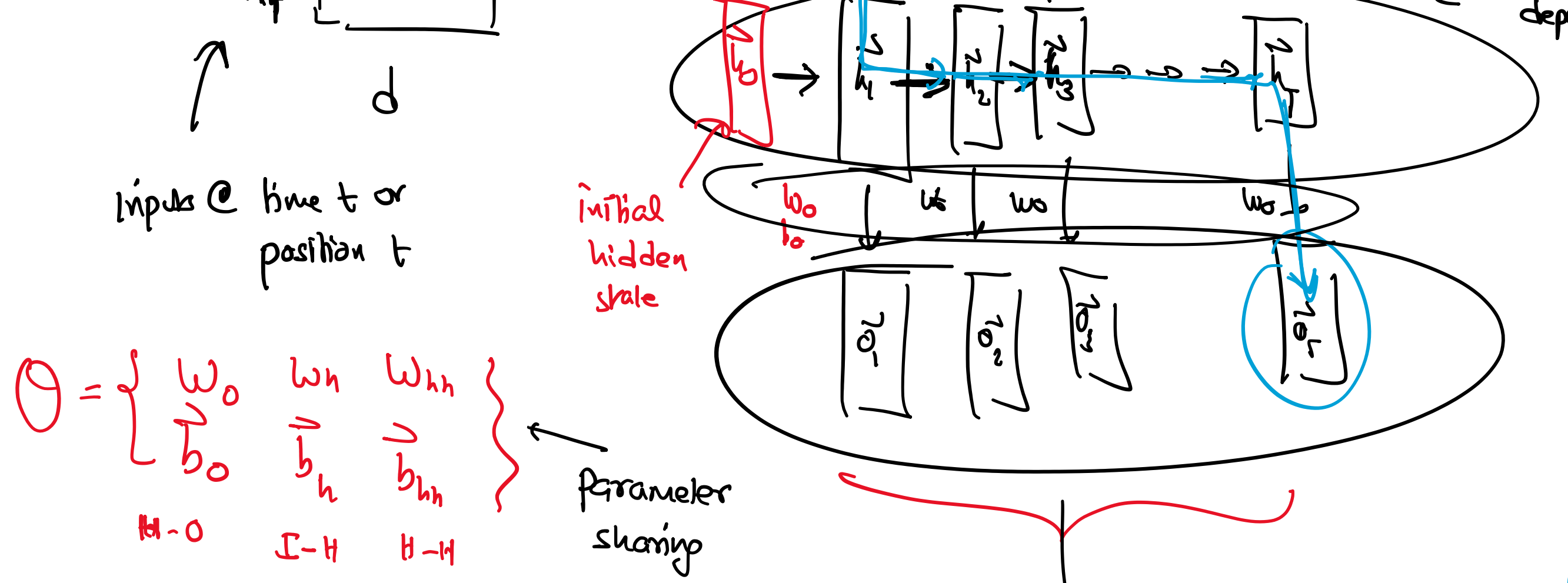
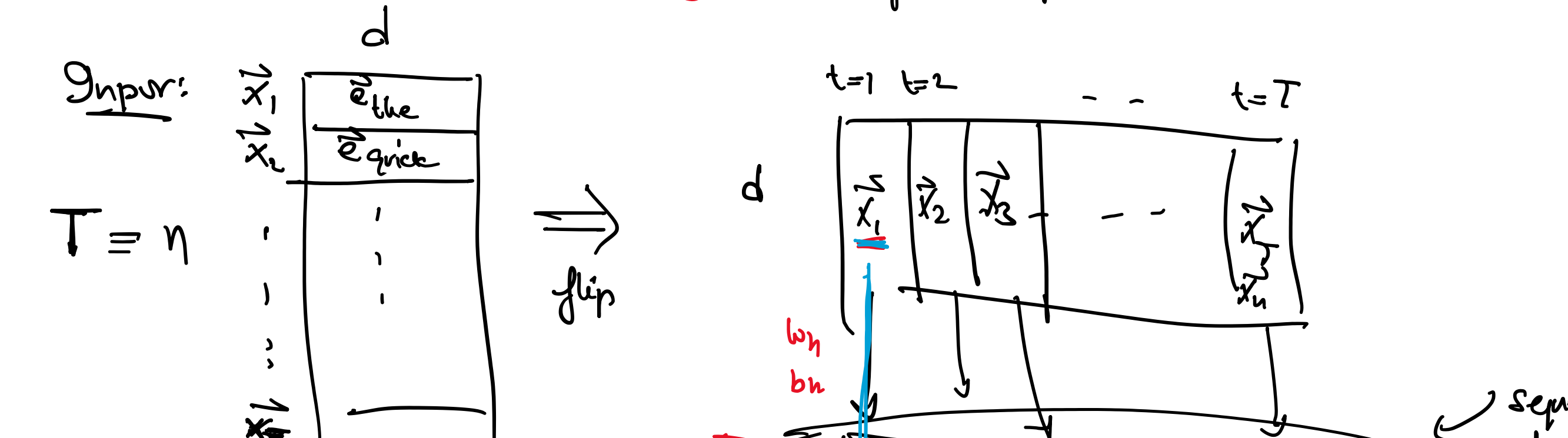
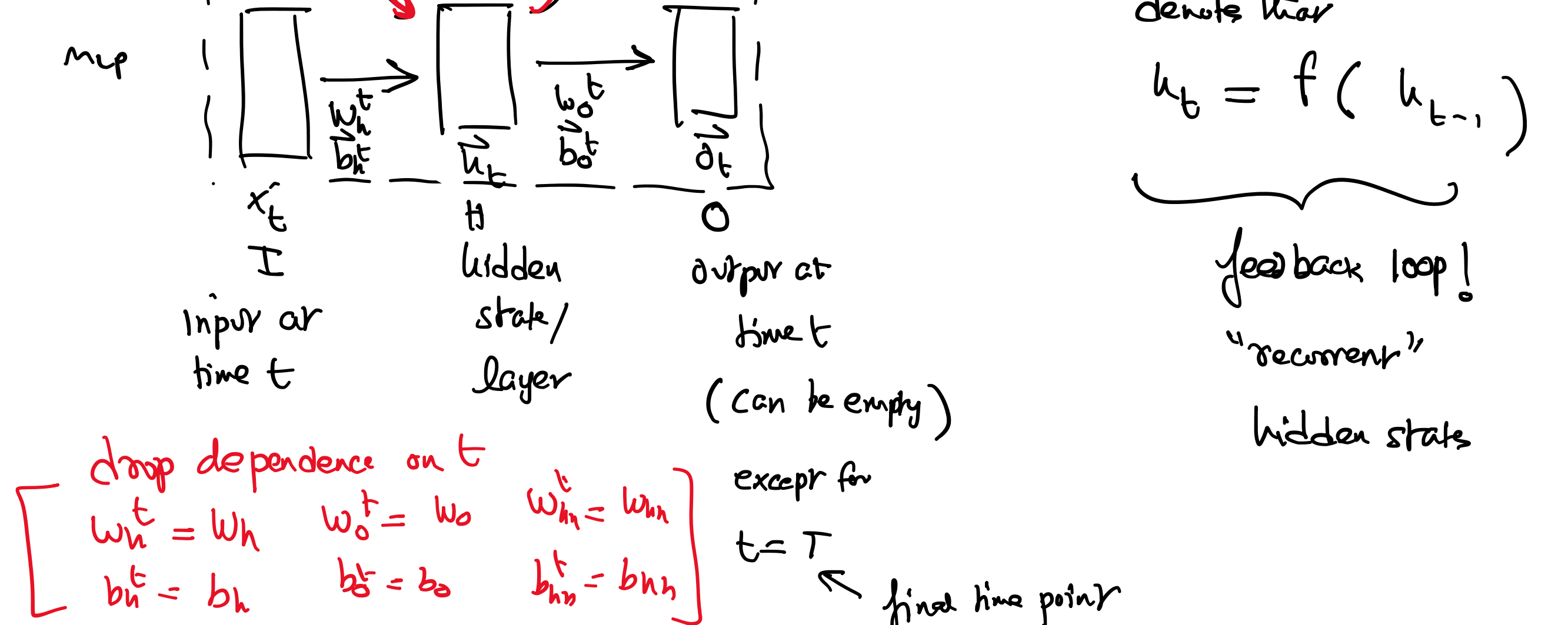
embedding layer \leftarrow learnable layer of weights \rightarrow embedding vector



$\vec{e}_{w_i} = \text{embedding vector for word } w_i$



Recurrent Neural Networks (RNNs)



attention (transformers)

