Question 2.2

Given: Network 1's weights and biases are $W^{(1)}$, $W^{(2)}$, $W^{(3)}$, $b^{(1)}$, $b^{(2)}$, and $b^{(3)}$.

Let input be $\vec{a}^{(0)}$, we'll find the output of Network 1 :

$$\text{Output of Network 1} = \vec{a}^{(3)} = W^{(3)}\vec{a}^{(2)} + \vec{b}^{(3)}$$

$$= W^{(3)}\left(W^{(2)}\vec{a}^{(1)} + \vec{b}^{(2)}\right) + \vec{b}^{(3)}$$

$$= W^{(3)}W^{(2)}\vec{a}^{(1)} + W^{(3)}\vec{b}^{(2)} + \vec{b}^{(3)}$$

$$= W^{(3)}W^{(2)}\left(W^{(1)}\vec{a}^{(0)} + \vec{b}^{(1)}\right) + W^{(3)}\vec{b}^{(2)} + \vec{b}^{(3)}$$

$$= W^{(3)}W^{(2)}W^{(1)}\vec{a}^{(0)} + W^{(2)}\vec{b}^{(1)} + W^{(3)}\vec{b}^{(2)} + \vec{b}^{(3)}$$

$$\text{Output of Network 2} = \tilde{W}\vec{a}^{(0)} + \vec{\tilde{b}}$$

Let Output of Network 1 = Output of Network 2

Hence, $\underline{W^{(3)}W^{(2)}W^{(1)}}\vec{a}^{(0)} + \underline{\underline{W^{(2)}\vec{b}^{(1)} + W^{(3)}\vec{b}^{(2)} + \vec{b}^{(3)}}} = \underline{\tilde{W}}\vec{a}^{(0)} + \underline{\underline{\vec{\tilde{b}}}}$

$$\boxed{\begin{aligned} \therefore\ \tilde{W} &= W^{(3)}W^{(2)}W^{(1)} \\ \vec{\tilde{b}} &= W^{(2)}\vec{b}^{(1)} + W^{(3)}\vec{b}^{(2)} + \vec{b}^{(3)} \end{aligned}}$$