5.1 Note: The question has been worded weirdly. We don't need to find the relation between $\sigma(a)$ and $\tanh(a)$, we need to find a relationship between $\sigma(a)$ and $\tanh(b)$ such that b is a linear transformation of a.

$$\tanh(a) = \frac{1 - e^{-2a}}{1 + e^{-2a}}$$

$$= \frac{1 - (1 + e^{-2a}) + 1}{1 + e^{-2a}}$$

$$= \frac{2 - (1 + e^{-2a})}{1 + e^{-2a}}$$

$$= \left(\frac{2}{1 + e^{-2a}}\right) - \left(\frac{1 + e^{-2a}}{1 + e^{-2a}}\right)$$

$$= 2\sigma(2a) - 1$$

The parameters of the two networks are related as follows.

$$y_k(\mathbf{x}, \mathbf{w}) = \sigma\left(\sum_{j=1}^{M} w_{kj}^{(2)} h\left(\sum_{i=1} w_{ji}^{(1)} x_i + w_{j0}^{(1)}\right) + w_{k0}^{(2)}\right)$$

Using tanh for h, we get:

$$y_k(\mathbf{x}, \mathbf{w}) = \sigma\left(\sum_{j=1}^{M} w_{kj}^{(2)} \tanh\left(\sum_{i=1} w_{ji}^{(1)} x_i + w_{j0}^{(1)}\right) + w_{k0}^{(2)}\right)$$

$$y_k(\mathbf{x}, \mathbf{w}) = \sigma\left(\sum_{j=1}^{M} w_{kj}^{(2)} 2\sigma\left(2\left(\sum_{i=1} w_{ji}^{(1)} x_i + w_{j0}^{(1)}\right)\right) - 1 + w_{k0}^{(2)}\right)$$

$$y_k(\mathbf{x}, \mathbf{w}) = \sigma\left(\sum_{j=1}^{M} \left(2w_{kj}^{(2)}\right) \sigma\left(\sum_{i=1} \left(2w_{ji}^{(1)}\right) x_i + \left(2w_{j0}^{(1)}\right)\right) + \left(w_{k0}^{(2)} - 1\right)\right)$$

As we can see, the tanh network is equivalent to a σ network, with the network parameters being related to each other via a linear transformation.