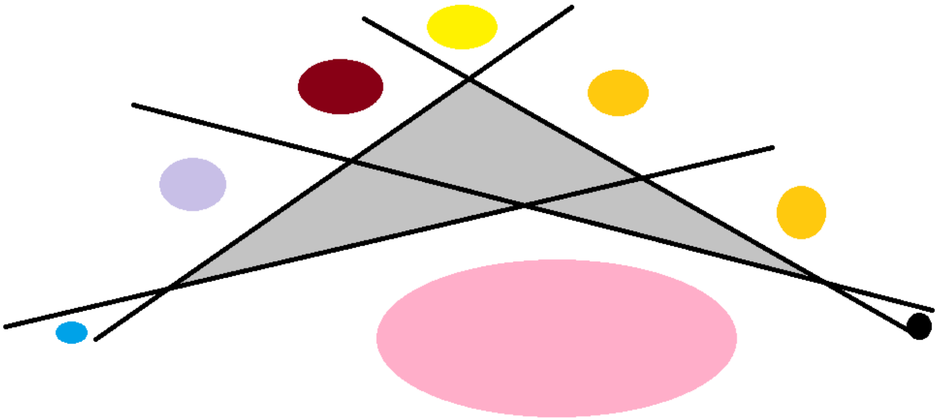
Problem 4:

No, it can be non-convex.



Problem 8:

Show that if the transfer function of the hidden units is linear, a three-layer network is equivalent to a two-layer one. Explain why, therefore, that a three-layer network with linear hidden units cannot solve a non-linearly separable problem such as XOR or n-bit parity.

A three-layer network can be represented as follow:

y1 = $f_{i\text{-}h}$(input)

y2 = tf_h(y1)

y3 = $f_{h\text{-}o}$(y2)

output = tf_o(y3)

where $f_{i\text{-}h}$ is the linear function from input layer to hidden layer, tf_h is the transfer function in hidden layer, $f_{h\text{-}o}$ is the linear function from hidden layer to output layer, tf_h is the transfer function in output layer

So output = $tf_o(f_{h\text{-}o}(tf_h(f_{i\text{-}h}(\text{input}))))$

If the tf_h() is linear, we can also find f'(output) = $f_{h\text{-}o}(tf_h(f_{i\text{-}h}(\text{input})))$, and f'() is also a linear function.(because the combination of linear functions will always be a linear one). And we can have a new network N', who has only two layers and $f_{i\text{-}o}$ = f'

In conclusion, a three-layer network is equivalent to a two-layer one.