Written Assignment 3

Lemma 3.1. *Let $k \in \mathbb{N}_0$ and $s \in 2\mathbb{N} - 1$. Then it holds that for all $\epsilon > 0$ there exists a shallow tanh neural network $\Psi_{s,\epsilon} : [-M, M] \to \mathbb{R}^{\frac{s+1}{2}}$ of width $\frac{s+1}{2}$ such that*

$$\max_{\substack{p \le s, \\ p \text{ odd}}} \left\| f_p - (\Psi_{s,\epsilon})_{\frac{p+1}{2}} \right\|_{W^{k,\infty}} \le \epsilon, \tag{17}$$

Lemma 3.1: Approximating Odd-Degree Polynomials with tanh Network:

Main idea: Lemma 3.1 addresses odd-degree polynomials, like x^3 and x^5. The central idea is that a tanh neural network with just one hidden layer can approximate any odd-degree polynomial to an arbitrary degree of accuracy. In other words, if the network has enough neurons (width $\frac{s+1}{2}$), its output curve can be made to align almost perfectly with the polynomial's curve.

How does it work? It use a mathematical tool called finite difference. Recall that for a small step h, the difference

$$\frac{(x+h)^p - (x-h)^p}{2h}$$

is an approximation to $x^{p-1} \cdot p$. By cleverly combining values of tanh at shifted points $x + h$ and $x - h$, one can cancel out unwanted terms and isolate x^p. In essence, the network builds x^p out of a weighted sum of tanh activations at different shifts.

Lemma 3.2. *Let $k \in \mathbb{N}_0, s \in 2\mathbb{N} - 1$ and $M > 0$. For every $\epsilon > 0$, there exists a shallow tanh neural network $\psi_{s,\epsilon} : [-M, M] \to \mathbb{R}^s$ of width $\frac{3(s+1)}{2}$ such that*

$$\max_{p \leq s} \left\| f_p - (\psi_{s,\epsilon})_p \right\|_{W^{k,\infty}} \leq \epsilon. \tag{26}$$

Lemma 3.2: Approximating Even-Degree Polynomials with tanh Network

Lemma 3.2 is an extension of Lemma 3.1, solving the approximation problem for even-degree polynomials (like x^2, x^4). Since the tanh function is inherently an odd function, directly approximating an even function with it would be challenging. Therefore, after obtaining the odd powers, one can use algebraic identities to express even powers as combinations of odd powers.

For example:

$$y^2 = \frac{1}{2a}[(y+a)^3 - (y-a)^3] - y$$

which expresses the square term in terms of cubic terms. Similarly, higher-order even powers can also be constructed from odd powers through addition and subtraction.