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*

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

這個lemma本質上是用一個函數（控制h來調整誤差）去逼近另一個函數（原函數），並且確保兩函數的0～k階導數誤差均小於ε

$f_p \uparrow$, p odd

(Norm in Sobolev space)

Def. $\|f-g\|_{W^{k,\infty}} = \max_{0 \le \alpha \le k} \sup |D^\alpha f - D^\alpha g|$

※ $D^\alpha f$ 是 f 的 α 階導數

這個範數是用來衡量原函數和逼近函數的誤差，他的特性是確保0～k階導數的形狀都大致吻合

接下來用以下的方式定義函數 $\hat{f}_{q,h}(y)$ 用來逼近 $f_p(y) = y^p$

3.1. Univariate polynomials

We first describe how to approximate univariate polynomials of any degree with tanh neural networks. We introduce the p-th order central finite difference operator δ_h^p for any $f \in C^{p+2}([a, b])$ for some $p \in \mathbb{N}$ by

$$\delta_h^p[f](x) = \sum_{i=0}^{p} (-1)^i \binom{p}{i} f\left(x + \left(\frac{p}{2} - i\right) h\right). \tag{15}$$

Next we define for any $p \in \mathbb{N}$, $q \in 2\mathbb{N}-1$ and $M > 0$ the monomials $f_p : [-M, M] \to \mathbb{R}$ and the tanh neural networks $\hat{f}_{q,h} : [-M, M] \to \mathbb{R}$ as

$$f_p(y) := y^p \quad \text{and} \quad \hat{f}_{q,h}(y) := \frac{\delta_{hy}^q[\sigma](0)}{\sigma^{(q)}(0) h^q}. \tag{16}$$

ex. for $p=3$, $\delta_h^3[f](x) = f\left(x+\frac{3h}{2}\right) - 3f\left(x+\frac{h}{2}\right) + 3f\left(x-\frac{h}{2}\right) - f\left(x-\frac{3h}{2}\right)$

根據一些技巧性的運算以及定理應用後，我們得到以下的不等式

$$\left\| f_p - \hat{f}_{p,h} \right\|_{W^{k,\infty}} \le \left((2(p+2)pM)^{p+3} + (2pk)^{k+1} \right) h^2 =: \epsilon.$$

對於任意給定的epsilon我們可以藉由微調h來控制誤差的大小，確保逼近函數近似於原函數

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 3.1 專注於p是奇數的情況，3.2則是不限制p的奇偶。
對於任意整數p，我們能將問題拆成奇數次方部分（使用Lemma 3.1）及偶數次方的部分分開處理

對於偶數次方的部分，Define:

$$(\psi_{s,\epsilon}(y))_{2n} = \frac{1}{2\alpha(2n+1)}\Bigg(\hat{f}_{2n+1,h}(y+\alpha) - \hat{f}_{2n+1,h}(y-\alpha) - 2\sum_{k=0}^{n-1}\binom{2n+1}{2k}\alpha^{2(n-k)+1}(\psi_{s,\epsilon}(y))_{2k}\Bigg).$$

使用 notation $E_p = \left\| f_p - (\psi_{s,\epsilon})_p \right\|_{W^{k,\infty}}$.

By Lemma 3.1 $\max\limits_{\substack{p \leq s \\ p \text{ odd}}} E_p \leq \varepsilon$ (by choosing h)

We can proof $E_p \leq E_p^* := \dfrac{2^{p/2}(1+\alpha)^{(p^2+p)/2}}{\alpha^{p/2}} \cdot \epsilon.$ by M.I.

for p=2n 在證明這個不等式時，我們會得到以下結果

$$E_{2n} \leq \frac{2}{\alpha}(1+\alpha)^{2n+1} E^*_{2(n-1)} \leq \left(\frac{2}{\alpha}(1+\alpha)^{2n+1}\right)^n \cdot \epsilon = E^*_{2n}.$$

也就是說，透過控制h及alpha的值，我們能夠控制逼近函數與原函數f_p 的0～k階導數近似