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*

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

Moreover, the weights of $\Psi_{s,\epsilon}$ scale as $O\left(\epsilon^{-s/2}(2(s+2)\sqrt{2M})^{s(s+3)}\right)$ for small ϵ and large s.

Concept：先介紹有限差分的概念

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

Then define.

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

對 $\delta_{hx}^p[\sigma](0)$ 做P階導數

$$\begin{aligned}
\frac{d^m}{dx^m} \delta_{hx}^p[\sigma](0) &= \sum_{i=0}^{p} (-1)^i \binom{p}{i} \left(\frac{p}{2} - i\right)^m h^m \cdot \sigma^{(m)}\left(\left(\frac{p}{2} - i\right) hx\right) \\
&= \boxed{\sum_{i=0}^{p} (-1)^i \binom{p}{i} \left(\frac{p}{2} - i\right)^m h^m \left(\sum_{l=m}^{p+1} \frac{\sigma^{(l)}(0)}{(l-m)!} \left(\frac{p}{2} - i\right)^{l-m} (hx)^{l-m}\right)} \\
&\quad + \sum_{i=0}^{p} (-1)^i \binom{p}{i} \left(\frac{p}{2} - i\right)^m h^m \frac{\sigma^{(p+2)}(\xi_{x,i})}{(p+2-m)!} \left(\frac{p}{2} - i\right)^{p+2-m} (hx)^{p+2-m}.
\end{aligned}$$

From Katsuura (2009, Theorem 1) it follows that

$$\sum_{i=0}^{p} (-1)^i \binom{p}{i} \left(\frac{p}{2} - i\right)^l = p!\,\delta(l-p) = \begin{cases} p!, & (l = p), \\ 0, & (l \neq p). \end{cases}$$

$$\begin{aligned}
&\sum_{i=0}^{p} (-1)^i \binom{p}{i} \left(\frac{p}{2} - i\right)^m h^m \left(\sum_{l=m}^{p+1} \frac{\sigma^{(l)}(0)}{(l-m)!} \left(\frac{p}{2} - i\right)^{l-m} (hx)^{l-m}\right) \\
&= h^m \sum_{l=m}^{p+1} \frac{\sigma^{(l)}(0)}{(l-m)!} (hx)^{l-m} \sum_{i=0}^{p} (-1)^i \binom{p}{i} \left(\frac{p}{2} - i\right)^l \\
&= \begin{cases} h^m \frac{\sigma^{(p)}(0)}{(p-m)!} (hx)^{p-m} p!, & 0 \leq m \leq p \\ 0, & m = p+1 \end{cases} = h^p \sigma^{(p)}(0) f_p^{(m)}(x).
\end{aligned}$$

$\because \dfrac{h^p \sigma^{(p)}(0) f_p^m(x)}{h^p \sigma^{(p)}(0)} = f_p^m(x)$

且 $\dfrac{\frac{d^m}{dx^m} \delta_{hx}^p[\sigma](0)}{h^p \sigma^{(p)}(0)} = \hat{f}_{p,x}$

$\therefore$

$$\hat{f}_{p,h}^{(m)}(x) - f_p^{(m)}(x) = \sum_{i=0}^{p} (-1)^i \binom{p}{i} \frac{1}{(p+2-m)!} \frac{\sigma^{(p+2)}(\xi_{x,i})}{\sigma^{(p)}(0)} \left(\frac{p}{2} - i\right)^{p+2} h^2 x^{p+2-m}.$$

define 上界 m

$$\begin{aligned}
\left| f_p - \hat{f}_{p,h} \right|_{W^{m,\infty}} &\leq \sum_{i=0}^{p} \binom{p}{i} \frac{\left|\sigma^{(p+2)}(\xi_{x,i})\right|}{\left|\sigma^{(p)}(0)\right|} \left|\frac{p}{2} - i\right|^{p+2} h^2 M^{p+2} \\
&\leq 2^p \frac{(2(p+2))^{p+3}}{1} \left(\frac{p}{2}\right)^{p+2} h^2 M^{p+2} \\
&\leq (2(p+2)pM)^{p+3} h^2.
\end{aligned}$$

補上PP階導數後的誤差項

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

所以控制好上界跟夠小的 h 就可以逼近任意 y^p p odd #