Proof. Let us start with the L^p case, $p < +\infty$. There holds

$$||u - T_k(u)||_{L^p(\Omega)}^p = \int_{\{u < -k\}} |u + k|^p dx + \int_{\{u > k\}} |u - k|^p dx$$

$$\leq \int_{\{u < -k\}} |u|^p dx + \int_{\{u > k\}} |u|^p dx$$

$$= \int_{\Omega} |u|^p \mathbf{1}_{|u| > k} dx \longrightarrow 0$$

when $k \to +\infty$ by monotone convergence. Concerning gradients, if u is in $W^{1,p}(\Omega)$, we similarly have

$$\|\nabla u - \nabla (T_k(u))\|_{L^p(\Omega)}^p = \int_{\Omega} |1 - T_k'(u)|^p |\nabla u|^p dx = \int_{\Omega} |\nabla u|^p \mathbf{1}_{|u|>k} dx \longrightarrow 0,$$

when $k \to +\infty$.

Finally, when
$$p = +\infty$$
, $T_k(u) = u$ as soon as $k \ge ||u||_{L^{\infty}(\Omega)}$.

Remark 3.10. i) In all cases, there holds $T_k(u) \in L^{\infty}(\Omega)$ with $||T_k(u)||_{L^{\infty}(\Omega)} \le k$. ii) If $u \in W_0^{1,p}(\Omega)$ then $T_k(u) \in W_0^{1,p}(\Omega)$ since $T_k(0) = 0$.

We conclude this general study with a few remarks.

Remark 3.11. i) The superposition operators do not in general operate on Sobolev spaces of order higher than 1. Thus, for example, if $u \in H^2(\Omega)$, it is not necessarily the case that $u_+ \in H^2(\Omega)$. This is obvious in one dimension of space, since $H^2(\Omega) \hookrightarrow C^1(\bar{\Omega})$ in this case. The problem however is not only connected to a lack of regularity of the function T. Thus, even if T is of class C^{∞} with T' and T'' bounded, and $u \in C^{\infty}(\Omega) \cap H^2(\Omega)$, we do not always have $T(u) \in H^2(\Omega)$.

In effect, by the classical chain rule, $\partial_i(T(u)) = T'(u)\partial_i u$ and $\partial_{ij}(T(u)) = T''(u)\partial_i u\partial_j u + T'(u)\partial_{ij}u$. The second term in the expression of second derivatives actually belongs to $L^2(\Omega)$. However, for the first term, in general $\partial_i u\partial_j u \notin L^2(\Omega)$ (except when $d \leq 4$ by the Sobolev embeddings). Let us mention a more general result in this direction: if T is C^{∞} , then for real s, if $u \in W^{s,p}(\Omega)$ then $T(u) \in W^{s,p}(\Omega)$ as soon as $s - \frac{d}{p} > 0$, see for example [51].

ii) The vector-valued case is comparable to the scalar case. If $T: \mathbb{R}^m \to \mathbb{R}$ is globally Lipschitz, then for all $u \in W^{1,p}(\Omega; \mathbb{R}^m)$, $T(u) \in W^{1,p}(\Omega)$. On the other hand, the chain rule formula is not valid as such, because $DT(u)\nabla u$ makes no sense in general. Consider for example, for m = 2, $T(u_1, u_2) = \max(u_1, u_2)$. If $u \in H^1(\Omega; \mathbb{R}^2)$ is of the form u = (v, v) with $v \in H^1(\Omega)$, then DT(u) is nowhere defined on Ω whereas ∇u is not zero almost everywhere on Ω . So we would be hard pressed to give a reasonable definition of such a product as $DT(u)\nabla u$. There are however more complicated formulas to describe this gradient.