Chapter 6

The Chain Rule Inequality and its Perturbations

In the previous chapter we showed that the chain rule operator equation shows a remarkable stability and rigidity, under modifications of operators or additive perturbations. In this chapter we study a different modification of the chain rule, replacing equalities by inequalities. Suppose $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ is a map satisfying the *chain rule inequality*

$$T(f \circ g) \leq Tf \circ g \cdot Tg, \quad f, g \in C^1(\mathbb{R}). \tag{6.1}$$

Under mild assumptions on T, we determine the form of all operators T satisfying this inequality, provided that the image of T contains functions attaining negative values. There will be an assumption of non-degeneration of T which is a weak surjectivity type requirement. Moreover, we impose a weak continuity condition on T. In the case of the chain rule equation, the continuity of the operators was not assumed, but it was a consequence of the solution formulas. Here we have less information on T, and we require T to be pointwise continuous, as defined below. Remarkably, for functions f with positive derivative, the solutions Tf of the chain rule inequality (6.1) turn out to be the same as for the chain rule equation. For general functions the solutions of the chain rule inequality are bounded from above by corresponding solutions of the chain rule equality. This is a similar phenomenon as in Gronwall's inequality in its differential form, cf. Gronwall [G] or Hartman [H], where the solution of a differential inequality is bounded by the solution of the corresponding differential equation. We also state results for the opposite inequality $T(f \circ g) \geq Tf \circ g \cdot Tg$. The proofs are based in part on a result about submultiplicative functions on $\mathbb{R}$, which is of independent interest.

H. König, V. Milman, *Operator Relations Characterizing Derivatives*,
https://doi.org/10.1007/978-3-030-00241-1_6

6.1 The chain rule inequality

Studying the chain rule inequality, we will impose the following two conditions.

Definition. An operator $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ is *non-degenerate* provided that, for any open interval $I \subset \mathbb{R}$ and any $x \in I$, there exists a function $g \in C^1(\mathbb{R})$ with $g(x) = x$, $\mathrm{Im}(g) \subset I$ and $Tg(x) > 1$. Let us call T *negatively non-degenerate* if there is $g \in C^1(\mathbb{R})$ with $g(x) = x$, $\mathrm{Im}(g) \subset I$ and $Tg(x) < -1$.

Definition. An operator $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ is *pointwise continuous* if for any sequence of functions $f_n \in C^1(\mathbb{R})$ and $f \in C^1(\mathbb{R})$ with $f_n \to f$ and $f_n' \to f'$ converging uniformly on all compact subsets of $\mathbb{R}$, we have the pointwise convergence of $\lim_{n\to\infty} Tf_n(x) = Tf(x)$ for all $x \in \mathbb{R}$.

Theorem 6.1 (Chain rule inequality)**.** *Let $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ be an operator such that the chain rule inequality holds:*

$$T(f \circ g) \le Tf \circ g \cdot Tg, \quad f, g \in C^1(\mathbb{R}). \tag{6.1}$$

Assume in addition that T is non-degenerate and pointwise continuous. Suppose further that there exists $x \in \mathbb{R}$ with $T(-\mathrm{Id})(x) < 0$. Then there is a continuous function $H \in C(\mathbb{R})$, $H > 0$, and there are real numbers $p > 0$ and $A \ge 1$, such that T has the form

$$Tf = \begin{cases} \frac{H\circ f}{H} f'^p, & f' \ge 0, \\ -A\frac{H\circ f}{H}|f'|^p, & f' < 0, \end{cases} \qquad f \in C^1(\mathbb{R}). \tag{6.2}$$

Remarks. (a) Let $Sf := \frac{H\circ f}{H}|f'|^p \operatorname{sgn} f'$. Then S satisfies the chain rule equation $S(f\circ g) = Sf \circ g \cdot Sg$. Equation (6.2) means that $Tf \le Sf$. Thus, the solutions of the chain rule inequality are bounded from above by solutions of the chain rule equation for which $A = 1$. Note that $-A = T(-\mathrm{Id})(0) \le -1$. Thus under the additional assumption $T(-\mathrm{Id})(0) = -1$ in Theorem 6.1, T satisfies the chain rule equation.

(b) Let $c > 0$. The modified operator inequality $T(f\circ g) \le c\cdot Tf \circ g \cdot Tg$ may be treated by considering $T_1 := c \cdot T$ which would satisfy $T_1(f\circ g) \le T_1 f \circ g \cdot T_1 g$.

(c) The condition $T(-\mathrm{Id})(x) < 0$ guarantees that there are sufficiently many negative functions in the range of T. If this is violated, there are many *positive* solution operators T of (6.1): Examples for non-negative solutions can be given by

$$Tf(x) = F\big(x, f(x), |f'(x)|\big),$$

where $F : \mathbb{R}^2 \times \mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}$ is a continuous function satisfying

$$F(x, z, \alpha\beta) \le F(y, z, \alpha) F(x, y, \beta) \tag{6.3}$$

for all $x, y, z, \alpha, \beta \in \mathbb{R}$. We might take, e.g.,

$$F(x, z, \alpha) = \exp(d(x, z)) \cdot K(\alpha),$$

where d is either a metric on $\mathbb{R}$ or $d(x, z) = z - x$, and $K : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is continuous and *submultiplicative*, $K(\alpha\beta) \leq K(\alpha) \cdot K(\beta)$ for $\alpha, \beta \geq 0$. Non-trivial examples of such maps K besides power-type functions α^p, $p > 0$ and maxima of such functions are given, e.g., by $K(\alpha) = \ln(\alpha + c)$ with $c \geq e$, cf. Gustavsson, Maligranda, Peetre [GMP], and products of submultiplicative functions. Moreover, for any continuous submultiplicative function $K : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, $\widetilde{F} := K \circ F$ will also satisfy (6.3) if F does. There does not seem to be much hope of classifying the solutions of (6.1) without any negativity assumption like $T(-\operatorname{Id})(x) < 0$ for some $x \in \mathbb{R}$.

6.2 Submultiplicative functions

Let $K : \mathbb{R} \to \mathbb{R}$ be continuous and define $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ by $Tf(x) := K(f'(x))$. This operator T will satisfy (6.1) if and only if K is submultiplicative, i.e., $K(\alpha\beta) \leq K(\alpha)K(\beta)$ for all $\alpha, \beta \in \mathbb{R}$. Hence, as a special case in the proof of Theorem 6.1, we have to classify submultiplicative functions on $\mathbb{R}$ attaining also negative values. This result is of independent interest and we formulate it as Theorem 6.2.

Theorem 6.2 (Submultiplicative functions)**.** *Let $K : \mathbb{R} \to \mathbb{R}$ be a measurable function which is continuous in* 0 *and in* 1 *and* submultiplicative*, i.e.,*

$$K(\alpha\beta) \leq K(\alpha)K(\beta), \quad \alpha, \beta \in \mathbb{R}.$$

Assume further that $K(-1) < 0 < K(1)$. Then there exist real numbers $p > 0$ and $A \geq 1$ such that

$$K(\alpha) = \begin{cases} \alpha^p, & \alpha \geq 0, \\ -A|\alpha|^p, & \alpha < 0. \end{cases}$$

Hence, $K(-1) = -A \leq 1$. Note that $K|_{\mathbb{R}_{\geq 0}}$ is multiplicative, and if $K(-1) = -1$, K is multiplicative on $\mathbb{R}$, i.e., $K(\alpha) = |\alpha|^p \operatorname{sgn} \alpha$. As mentioned in Remark (b) above, there are many continuous submultiplicative functions $K : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ besides powers $K(\alpha) = \alpha^p$. However, these cannot be extended to continuous submultiplicative functions $K : \mathbb{R} \to \mathbb{R}$ with $K(-1) < 0$. There is a corresponding result for *supermultiplicative* functions on $\mathbb{R}$, $K(\alpha\beta) \geq K(\alpha)K(\beta)$, which gives the same form of K, except that then $0 < A \leq 1$.

Examples. (a) The measurability assumption in Theorem 6.2 is necessary. Otherwise, we may take a non-measurable additive function $f : \mathbb{R} \to \mathbb{R}$ as given in the comments following Proposition 2.1 and $A > 1$, and define $K(0) := 0$ and

$$K(\alpha) := \begin{cases} \exp(f(\ln \alpha)), & \alpha > 0, \\ -A \exp(f(\ln |\alpha|)), & \alpha < 0. \end{cases}$$

Then $K : \mathbb{R} \to \mathbb{R}$ is non-measurable and submultiplicative with $K(-1) < 0 < K(1)$.

(b) Let $d \geq 1$, $c \geq 0$, $c \neq d$, and put

$$K(\alpha) := \begin{cases} 1, & \alpha = 1, \\ -c, & \alpha = 0, \\ -d, & \alpha \notin \{0,1\}. \end{cases}$$

Then K is measurable and submultiplicative with $K(-1) < 0 < K(1)$, but discontinuous at 0 and at 1.

The result corresponding to Theorem 6.1 in the supermultiplicative situation is

Theorem 6.3. *Let $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ be an operator such that*

$$T(f \circ g) \geq Tf \circ g \cdot Tg, \quad f, g \in C^1(\mathbb{R}).$$

Assume also that T is negatively non-degenerate and pointwise continuous with $T(-\operatorname{Id})(x) < 0$ for some $x \in \mathbb{R}$. Then there exist numbers $p > 0$, $0 < B \leq 1$ and a function $H \in C(\mathbb{R})$, $H > 0$ such that

$$Tf = \begin{cases} \frac{H \circ f}{H} f'^p, & f' \geq 0, \\ -B \frac{H \circ f}{H} |f'|^p, & f' < 0, \end{cases} \qquad f \in C^1(\mathbb{R}).$$

We first prove Theorem 6.2 which is used in the proof of Theorem 6.1. For this, we need two lemmas.

Lemma 6.4. *Let $K : \mathbb{R} \to \mathbb{R}$ be submultiplicative with $K(-1) < 0 < K(1)$. Assume that K is continuous in 0 and in 1. Then:*

(i) $K(0) = 0$, $K(1) = 1$ *and* $K|_{\mathbb{R}_{<0}} < 0 < K|_{\mathbb{R}_{>0}}$.

(ii) *There is* $0 < \epsilon < 1$ *such that* $0 < K(\alpha) < 1$ *for all* $\alpha \in (0, \epsilon)$ *and* $1 < K(\alpha) < \infty$ *for all* $\alpha \in (1/\epsilon, \infty)$.

Proof. Since $0 < K(1) = K(1^2) \leq K(1)^2$, $K(1) \geq 1$. Then $1 \leq K(1) = K((-1)^2) \leq K(-1)^2$, implying $K(-1) \leq -1$. By submultiplicativity $K(-1) \leq K(1)K(-1)$, $|K(-1)| \geq K(1)|K(-1)|$. Hence, $K(1) \leq 1$, $K(1) = 1$. Since K is continuous at 1, there is $\epsilon > 0$ such that $K|_{[1/(1+\epsilon), 1+\epsilon]} > 0$. For any $\alpha \in [1/(1+\epsilon), 1+\epsilon]$, $K(\alpha) > 0$ and $K(1/\alpha) > 0$. Hence, $0 < K(\alpha) \leq K(1/\alpha)K(\alpha^2)$, implying that $K(\alpha^2) > 0$, i.e., $K|_{[1/(1+\epsilon)^2, (1+\epsilon)^2]} > 0$. Inductively, we get that $K|_{\mathbb{R}_{>0}} > 0$, since $\mathbb{R}_{>0} = \bigcup_{n \in \mathbb{N}} [1/(1+\epsilon)^n, (1+\epsilon)^n]$. The inequality $K(0) = K((-1) \cdot 0) \leq K(-1) \cdot K(0)$ with $K(-1) < 0$ shows that $K(0) \leq 0$. Since $K|_{\mathbb{R}_{>0}} > 0$ and K is continuous in 0, we get $K(0) = 0$. Then there is $\epsilon > 0$ with $0 < K|_{(0,\epsilon)} < 1$. Since $1 \leq K(1) \leq K(\alpha) \cdot K(1/\alpha)$, it follows that $K|_{(1/\epsilon, \infty)} > 1$. Moreover, for any $\alpha > 0$, $K(-\alpha) \leq K(-1)K(\alpha) < 0$, i.e., $K|_{\mathbb{R}_{<0}} < 0$. □

The second lemma is a well-known fact on subadditive functions on $\mathbb{R}$.

Lemma 6.5. *Assume that $f : \mathbb{R} \to \mathbb{R}$ is measurable and* subadditive*, i.e.,*

$$f(s+t) \leq f(s) + f(t), \quad s,t \in \mathbb{R}.$$

Define $p := \sup_{t<0} \frac{f(t)}{t}$ and $q := \inf_{t>0} \frac{f(t)}{t}$. Then f is bounded on compact intervals, $-\infty < p \leq q < \infty$ and $f(0) \geq 0$. Moreover, the limits $\lim_{t\to-\infty} \frac{f(t)}{t}$, $\lim_{t\to\infty} \frac{f(t)}{t}$ exist and $p = \lim_{t\to-\infty} \frac{f(t)}{t}$, $q = \lim_{t\to\infty} \frac{f(t)}{t}$.

Proof. (a) We first show that f is bounded from above on each compact subset of $(0,\infty)$. Fix $a > 0$ and put $A := f(a)$. Let $E := \{t \in (0,a) \mid f(t) \geq A/2\}$. Then E is measurable since f is measurable. Moreover, $(0,a) = E \cup (\{a\} - E)$, since $t_1, t_2 > 0$ with $a = t_1 + t_2$ implies that $t_1 \in E$ or $t_2 \in E$. Suppose there are $0 < \alpha < \beta < \infty$ such that $f|_{[\alpha,\beta]}$ is not bounded from above. There there is a sequence (t_n), $\alpha \leq t_n \leq \beta$ with $t_n \to t_0$, $\alpha \leq t_0 \leq \beta$ and $f(t_n) \geq 2n$. Let $E'_n := \{t \in (0,\beta) \mid f(t) \geq n\}$. For a fixed $n \in \mathbb{N}$, choose above $a = t_n$, $A = f(t_n) \geq 2n$. Then

$$|E'_n| \geq \big|\{t \in (0,t_n) \mid f(t) \geq n\}\big| =: |E_n| \geq \frac{t_n}{2} \geq \frac{\alpha}{2} > 0.$$

Since $E'_n \geq E'_{n+1}$, we get that $\left|\bigcap_{n\in\mathbb{N}} E'_n\right| \geq \frac{\alpha}{2}$. Therefore, $f|_{(0,\beta)}$ is infinite on a set of strictly positive measure, which is a contradiction. Therefore, f is bounded from above on any compact subset of $(0,\infty)$.

(b) Since $f(0) \leq f(0) + f(0)$, we have $f(0) \geq 0$. Also, for any $t \in \mathbb{R}$, $0 \leq f(0) \leq f(t) + f(-t)$. Hence, $\frac{f(-t)}{-t} \leq \frac{f(t)}{t}$ for any $t > 0$, and therefore $\overline{\lim}_{t\to-\infty} \frac{f(t)}{t} \leq \underline{\lim}_{t\to\infty} \frac{f(t)}{t}$.

Let $q := \inf_{t>0} \frac{f(t)}{t}$. We claim that the limit $\lim_{t\to\infty} \frac{f(t)}{t}$ exists and that $q = \lim_{t\to\infty} \frac{f(t)}{t}$. Assume that $q \in \mathbb{R}$; the case $q = -\infty$ is treated similarly. Let $\epsilon > 0$ and choose $b > 0$ with $\frac{f(b)}{b} \leq q + \epsilon$. For any $t \geq 3b$, there is $n \in \mathbb{N}$ with $t \in [(n+2)b, (n+3)b]$. Using the subadditivity of f, and the definition of q and b, we find

$$\begin{aligned} q \leq \frac{f(t)}{t} &= \frac{f(nb + (t-nb))}{t} \leq \frac{n\,f(b) + f(t-nb)}{t} \\ &= \frac{nb}{t}\frac{f(b)}{b} + \frac{f(t-nb)}{t} \leq \frac{nb}{t}(q+\epsilon) + \frac{f(t-nb)}{t}. \end{aligned}$$

By part (a), f is bounded from above on $[2b, 3b]$. Let $M > 0$ be such that $f_{[2b,3b]} \leq M$. Since $t - nb \in [2b, 3b]$, we get

$$q \leq \frac{f(t)}{t} \leq \frac{nb}{t}(q+\epsilon) + \frac{M}{t}.$$

For $t \to \infty$, $\frac{nb}{t} \to 1$ and $\frac{M}{t} \to 0$. Therefore, for any $\epsilon > 0$,

$$\overline{\lim_{t\to\infty}} \frac{f(t)}{t} \le q + \epsilon = \inf_{t>0} \frac{f(t)}{t} + \epsilon,$$

which shows that the limit $\lim_{t\to\infty} \frac{f(t)}{t}$ exists and is equal to q.

(c) Consider similarly $g(t) := f(-t)$, $t > 0$. Then by (b)

$$p := \sup_{t>0} \frac{f(-t)}{-t} = -\inf_{t>0} \frac{g(t)}{t} = -\lim_{t\to\infty} \frac{g(t)}{t} = \lim_{t\to-\infty} \frac{f(-t)}{-t}.$$

Now $\frac{f(-t)}{-t} \le \frac{f(t)}{t}$ implies for $t \to \infty$ that $p \le q$. Moreover, since $p > -\infty$, $q > -\infty$, and since $q < \infty$, $p < \infty$. Therefore, $-\infty < p \le q < \infty$. □

As a consequence of Lemma 6.5, we have that

$$f(t) = pt + a(t), \quad t < 0,$$
$$f(t) = qt + a(t), \quad t > 0,$$

where $a(t) \ge 0$ for all $t \ne 0$.

Proof of Theorem 6.2. (a) Let $K : \mathbb{R} \to \mathbb{R}$ be measurable and submultiplicative, continuous in 0 and in 1 with $K(-1) < 0 < K(1)$. By Lemma 6.4, $K(0) = 0$, $K(1) = 1$, $K|_{\mathbb{R}_{<0}} < 0 < K|_{\mathbb{R}_{>0}}$ and for a suitable $0 < \epsilon < 1$, $0 < K(\alpha) < 1$ for all $\alpha \in (0, \epsilon)$ and $1 < K(\alpha) < \infty$ for all $\alpha \in (1/\epsilon, \infty)$. Define $f(t) := \ln K(\exp(t))$, $t \in \mathbb{R}$. Then f is measurable and subadditive, and we have by Lemma 6.5

$$-\infty < p := \sup_{t<0} \frac{f(t)}{t} = \lim_{t\to-\infty} \frac{f(t)}{t} \le q := \inf_{t>0} \frac{f(t)}{t} = \lim_{t\to\infty} \frac{f(t)}{t} < \infty.$$

Since f is negative for $t < 0$ and positive for $t > 0$, we have that $0 \le p \le q < \infty$, with

$$f(t) =: pt + a(t), \quad t < 0, \quad f(t) =: qt + a(t), \quad t > 0,$$

where $a(t) \ge 0$ for all t and $\lim_{t\to-\infty} \frac{a(t)}{t} = \lim_{t\to\infty} \frac{a(t)}{t} = 0$. This means, for all $0 < \alpha < 1$, that

$$K(\alpha) = \exp\big(f(\ln \alpha)\big) = \alpha^p \exp\big(a(\ln \alpha)\big) \ge \alpha^p,$$

and, for all $1 < \alpha < \infty$, that

$$K(\alpha) = \exp\big(f(\ln \alpha)\big) = \alpha^q \exp\big(a(\ln \alpha)\big) \ge \alpha^q.$$

(b) We claim that $p = q > 0$ holds. Using $K|_{\mathbb{R}_{<0}} < 0 < K|_{\mathbb{R}_{>0}}$, the submultiplicativity of K implies that for all $\beta < 0 < \alpha$

$$K(\alpha\beta) \le K(\alpha)K(\beta), \quad |K(\alpha\beta)| \ge K(\alpha)|K(\beta)|.$$

Since $K(1) = 1$, $f(0) = 0$. Fix $t < 0$ and choose $\alpha < -1$, $0 < \beta < 1$, with $t = \alpha\beta$. Then, by submultiplicativity,

$$K(t) = K\big((-1)|\alpha|\beta\big) \le K(-1)K(|\alpha|)K(\beta) \le 0,$$
$$|K(t)| \ge |K(-1)|K(|\alpha|)K(\beta) \ge |\alpha|^q\beta^p = |t|^q\beta^{p-q},$$

using that $K(-1) \le -1$ since $1 = K(1) \le K(-1)^2$. Assuming $p \ne q$, i.e., $p < q$, and letting β tend to 0 (and α to $-\infty$), would yield the contradiction $|K(t)| = \infty$. Hence, $0 \le p = q < \infty$. In fact, $0 < p = q$ since K is continuous at 0 with $K(0) = 0$ and $K(\beta) \ge \beta^p$ for $0 < \beta < 1$.

(c) Let $g(t) := \ln|K(-\exp(t))|$ for all $t \in \mathbb{R}$. Since, for any $s, t \in \mathbb{R}$,

$$K\big(-\exp(s)\exp(t)\big) \le K\big(-\exp(s)\big)K\big(\exp(t)\big) \le 0,$$

we get that

$$\begin{aligned} g(s+t) &= \ln\big|K\big(-\exp(s)\exp(t)\big)\big| \ge \ln\big|K(-\exp(s))\big| + \ln K(\exp(t)) \\ &= g(s) + f(t) = g(s) + pt + a(t), \end{aligned}$$

with $a(t) \ge 0$, for all t, and $\lim_{t\to\pm\infty} \frac{a(t)}{t} = 0$. Since $f(0) = 0$, $a(0) = 0$. Putting $s = 0$ yields

$$g(t) \ge g(0) + pt + a(t).$$

Putting $t = -s$ and renaming s by t gives

$$g(0) \ge g(t) - pt + a(-t).$$

Hence,

$$g(0) + pt + a(t) \le g(t) \le g(0) + pt - a(-t).$$

Since $a \ge 0$, this implies that $a = 0$ on $\mathbb{R}$. Therefore, for all $t \in \mathbb{R}$, $f(t) = pt$ and $g(t) = g(0) + pt$. We then find, for all $\beta < 0 < \alpha$,

$$K(\alpha) = \alpha^p, \quad |K(\beta)| = \exp\big(g(\ln|\beta|)\big) = \exp(g(0))|\beta|^p.$$

Since $\exp(g(0)) = |K(-1)| \ge 1$, $g(0) \ge 0$. Thus, $K(\beta) = K(-1)|\beta|^p$, proving Theorem 6.2 with $-A = K(-1) \le -1$. □

6.3 Localization and Proof of Theorem 6.1

As the first step in the proof of Theorem 6.1 on the chain rule inequality , we show that T is locally defined. More precisely, $Tf(x)$ only depends on x, $f(x)$ and $f'(x)$.

Proposition 6.6. *Let $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ be non-degenerate, pointwise continuous and satisfy the chain rule inequality* (6.1)*. Assume also that there exists $x_0 \in \mathbb{R}$ such that $T(-\operatorname{Id})(x_0) < 0$. Then there is a function $F : \mathbb{R}^3 \to \mathbb{R}$ such that, for all $f \in C^1(\mathbb{R})$ and all $x \in \mathbb{R}$,*

$$Tf(x) = F\big(x, f(x), f'(x)\big). \tag{6.4}$$

To show this, we need a lemma.

Lemma 6.7. *Under the assumptions of Proposition* 6.6 *we have for any open interval $I \subset \mathbb{R}$:*

(a) *For all $x \in I$, there is $g \in C^1(I)$ with $g(x) = x$, $\operatorname{Im}(g) \subset I$ and $Tg(x) < -1$.*

(b) *For $c \in \mathbb{R}$, $f \in C^1(\mathbb{R})$ with $f|_I = c$, we have $Tf|_I = 0$.*

(c) *For $f \in C^1(\mathbb{R})$ with $f|_I = \operatorname{Id}|_I$, we have $Tf|_I = 1$.*

(d) *Take $f_1, f_2 \in C^1(\mathbb{R})$ with $f_1|_I = f_2|_I$ and assume that f_2 is invertible. Then $Tf_1|_I \leq Tf_2|_I$. Hence, if f_1 is invertible, too, $Tf_1|_I = Tf_2|_I$.*

Proof. (a) By (6.1), $T(\operatorname{Id})(x) \leq T(\operatorname{Id})(x)^2$ for all $x \in \mathbb{R}$. Hence, $T(\operatorname{Id})(x) \geq 1$ or $T(\operatorname{Id})(x) \leq 0$. If there would be $x_1 \in \mathbb{R}$ with $T(\operatorname{Id})(x_1) \leq 0$, use that by non-degeneration of T there is $g \in C^1(\mathbb{R})$, $g(x_1) = x_1$ and $Tg(x_1) > 1$. Then,

$$1 \leq Tg(x_1) = T(g \circ \operatorname{Id})(x_1) \leq Tg(x_1)T(\operatorname{Id})(x_1) \leq 0,$$

a contradiction. Hence $T(\operatorname{Id})(x) \geq 1$ for all $x \in I$. Also, $T(-\operatorname{Id})(x) < 0$: $1 \leq T(\operatorname{Id})(x) = T((-\operatorname{Id})^2)(x) \leq T(-\operatorname{Id})(-x)T(-\operatorname{Id})(x)$. By assumption, there is $x_0 \in \mathbb{R}$, with $T(-\operatorname{Id})(x_0) < 0$. If there would be $x_1 \in \mathbb{R}$ with $T(-\operatorname{Id})(x_1) > 0$, by continuity of the function $T(-\operatorname{Id})$ there would be $x_2 \in \mathbb{R}$ with $T(-\operatorname{Id})(x_2) = 0$, contradicting $1 \leq T(-\operatorname{Id})(-x_2)T(-\operatorname{Id})(x_2)$. Hence, $T(-\operatorname{Id})(x) < 0$ for all $x \in \mathbb{R}$. Also $1 \leq T(-\operatorname{Id})(0)^2$ yields $T(-\operatorname{Id})(0) \leq -1$.

Now let $I \subset \mathbb{R}$ be an open interval and $x_1 \in I$. Let $\epsilon > 0$ with $J = (x_1 - \epsilon, x_1 + \epsilon) \subset I$, $\widetilde{J} := J - \{x_1\} = (-\epsilon, \epsilon)$. Since T is non-degenerate, there is $f \in C^1(\mathbb{R})$ with $f(0) = 0$, $\operatorname{Im}(f) \subset \widetilde{J}$ and $Tf(0) > 1$. Then

$$T(-f)(0) \leq T(-\operatorname{Id})(0)Tf(0) < -1,$$

and $\operatorname{Im}(-f) \subset \widetilde{J}$. We transport $-f$ back to J by conjugation with a shift. For $y \in \mathbb{R}$, let $S_y := \operatorname{Id} + y \in C^1(\mathbb{R})$ denote the shift by y. Since for $y_n \to y$, $S_{y_n} \to S_y$ and $S'_{y_n} \to S'_y$ converge uniformly on compacta, by the pointwise continuity of T, we have that $T(S_{y_n})(x) \to T(S_y)(x)$ for all $x \in \mathbb{R}$, i.e., $T(S_y)(x)$ is continuous in y for every fixed $x \in \mathbb{R}$. Since

$$1 \leq T(\operatorname{Id})(x_1) \leq T(S_{x_1})(0)T(S_{-x_1})(x_1),$$

we have $T(S_{x_1})(0) \neq 0$. Using $T(S_0)(0) = T(\mathrm{Id})(0) \geq 1$, the continuity of $T(S_y)(0)$ in y implies that $T(S_{x_1})(0) > 0$. Let $g := S_{x_1} \circ (-f) \circ S_{-x_1}$. Then $g(x_1) = x_1$, $\mathrm{Im}(g) \subset J \subset I$, and

$$Tg(x_1) \leq T(S_{x_1})(0)T(-f)(0)T(S_{-x_1})(x_1) < -1,$$

using $T(-f)(0) < -1$ and $1 \leq T(S_{x_1})(0)T(S_{-x_1})(x_1)$.

(b) For the constant function c, $c \circ g = c$, hence $Tc(x) \leq Tc(g(x))Tg(x)$ for all $g \in C^1(\mathbb{R})$. By non-degeneration of T and (a), there are $g_1, g_2 \in C^1(\mathbb{R})$ with $g_j(x) = x$, $\mathrm{Im}(g_j) \subset I$ $(j \in \{1, 2\})$, and $Tg_2(x) < -1$, $Tg_1(x) > 1$. Applying the previous inequality to $g = g_1, g_2$, we find $Tc(x) = 0$.

Now suppose $f \in C^1(\mathbb{R})$ satisfies $f|_I = c$. Let $x \in I$ and g_1, g_2 be as before. Since $f \circ g_j = c$, for any $x \in I$, we have $0 = Tc(x) \leq Tf(x)Tg_j(x)$, yielding $Tf(x) = 0$. Hence $Tf|_I = 0$.

(c) Assume that $f \in C^1(\mathbb{R})$ satisfies $f|_I = \mathrm{Id}|_I$. Let $x \in I$ and choose g_1, g_2 as in part (b). Then $f \circ g_j = g_j$ $(j = 1, 2)$ and

$$Tg_j(x) = T(f \circ g_j)(x) \leq Tf(x)Tg_j(x).$$

This inequality for g_1 yields $Tf(x) \geq 1$, the one for g_2 that

$$|Tg_2(x)| \geq Tf(x)|Tg_2(x)|, \quad Tf(x) \leq 1.$$

Hence, $Tf(x) = 1$, $Tf|_I = 1$.

(d) Assume that $f_1|_I = f_2|_I$ and that f_2 is invertible. Let $g := f_2^{-1} \circ f_1$. Then $g \in C^1(\mathbb{R})$ with $f_1 = f_2 \circ g$ and $g|_I = \mathrm{Id}|_I$. By (c), $Tg|_I = 1$. Hence, for any $x \in I$, we have $g(x) = x$ and

$$Tf_1(x) = T(f_2 \circ g)(x) \leq Tf_2(x)Tg(x) = Tf_2(x).$$

Therefore, $Tf_1|_I \leq Tf_2|_I$. □

Proof of Proposition 6.6. (i) Let $\mathcal{C} := \{f \in C^1(\mathbb{R}) \mid f \text{ is invertible and } f'(x) \neq 0 \text{ for all } x \in \mathbb{R}\}$. For any open interval $I \subset \mathbb{R}$ and $f_1, f_2 \in \mathcal{C}$ with $f_1|_I = f_2|_I$ we have by Lemma 6.7(d) that $Tf_1|_I = Tf_2|_I$, i.e., localization on intervals. Replacing a function $f \in C^1(\mathbb{R})$ by its tangent line approximation on the right side of a point x, and f on the left side of x is an operation inside $\mathcal{C}$. Therefore, the method of the proof of Proposition 3.3 yields that there is a function $F : \mathbb{R}^2 \times (\mathbb{R} \setminus \{0\}) \to \mathbb{R}$ such that for all $f \in \mathcal{C}$ and all $x \in \mathbb{R}$,

$$Tf(x) = F\big(x, f(x), f'(x)\big).$$

(ii) We now consider functions $f \in C^1(\mathbb{R})$ which are not invertible. Suppose $I := (y_0, y_1)$ is an interval where f is strictly increasing with $f'(x) > 0$, $x \in I$ and $f'(y_0) = f'(y_1) = 0$ (or $y_0 = -\infty$, $f'(y_1) = 0$ or $f'(y_0) = 0$, $y_1 = \infty$, with obvious modifications in the following). For $\epsilon > 0$ sufficiently small, $f'(y_0+\epsilon) > 0$, $f'(y_1 - \epsilon) > 0$ for all $0 < \epsilon \le \epsilon_0$. Define $\widetilde{f} \in C^1(\mathbb{R})$ by

$$\widetilde{f}(x) := \begin{cases} f(y_0), & x \le y_0, \\ f(x), & x \in I, \\ f(y_1), & x \ge y_1. \end{cases}$$

Then $\widetilde{f}'(y_0) = \widetilde{f}'(y_1) = 0$ and $\widetilde{f}$ is the limit of functions $\widetilde{f}_\epsilon \in \mathcal{C}$ in the sense that $\widetilde{f}_\epsilon \to \widetilde{f}$ and $\widetilde{f}'_\epsilon \to \widetilde{f}'$ converge uniformly on compacta. One may choose

$$\widetilde{f}_\epsilon(x) := \begin{cases} f(y_0+\epsilon) + f'(y_0+\epsilon)\big(x - (y_0+\epsilon)\big), & x \le y_0 + \epsilon, \\ f(x), & x \in (y_0+\epsilon, y_1 - \epsilon), \\ f(y_1-\epsilon) + f'(y_1-\epsilon)\big(x - (y_1-\epsilon)\big), & x \ge y_1 - \epsilon. \end{cases}$$

Note that $\widetilde{f}_\epsilon \in \mathcal{C}$ for any $0 < \epsilon \le \epsilon_0$ since $\widetilde{f}_\epsilon$ is invertible with $\widetilde{f}'_\epsilon(x) > 0$ for all $x \in \mathbb{R}$. By (i) for any $x \in I_\epsilon := (y_0+\epsilon, y_1-\epsilon)$

$$T\widetilde{f}_\epsilon(x) = F\big(x, \widetilde{f}_\epsilon(x), \widetilde{f}'_\epsilon(x)\big) = F\big(x, f(x), f'(x)\big).$$

Since T is pointwise continuous, for any $x \in (y_0, y_1)$

$$T\widetilde{f}(x) = \lim_{\epsilon \to 0} T\widetilde{f}_\epsilon(x) = F\big(x, f(x), f'(x)\big).$$

By definition of $\widetilde{f}_\epsilon$, $f|_{I_\epsilon} = \widetilde{f}_\epsilon|_{I_\epsilon}$. Since $\widetilde{f}_\epsilon \in \mathcal{C}$, we have by Lemma 6.7(d) that $Tf(x) \le T\widetilde{f}_\epsilon(x) = F(x, f(x), f'(x))$ for any $x \in I_\epsilon$. For $\epsilon \to 0$ this shows that

$$Tf(x) \le T\widetilde{f}(x) = F\big(x, f(x), f'(x)\big), \quad x \in (y_0, y_1).$$

(iii) We now show the converse inequality $T\widetilde{f}(x) \le Tf(x)$ for $x \in (y_0, y_1)$. We may write $\widetilde{f} = f \circ g$ where

$$g(x) = \begin{cases} y_0, & x \le y_0, \\ x, & x \in (y_0, y_1), \\ y_1, & x \ge y_1. \end{cases}$$

If g were in $C^1(\mathbb{R})$, $g|_{(y_0,y_1)} = \mathrm{Id}$, $Tg|_{(y_0,y_1)} = 1$ so that

$$T\widetilde{f}(x) \le Tf(x)Tg(x) = Tf(x),$$

which would prove the claim. However, $g \notin C^1(\mathbb{R})$. Therefore, we approximate g

by smooth functions $g_\epsilon \in C^1(\mathbb{R})$. Let

$$g_\epsilon(x) := \begin{cases} y_0 + \frac{\epsilon}{2}, & x < y_0, \\ y_0 + \frac{\epsilon^2+(x-y_0)^2}{2\epsilon}, & y_0 \le x \le y_0 + \epsilon, \\ x, & y_0 + \epsilon \le x \le y_1 - \epsilon, \\ y_1 - \frac{\epsilon^2+(x-y_1)^2}{2\epsilon}, & y_1 - \epsilon \le x \le y_1, \\ y_1 - \frac{\epsilon}{2}, & x \ge y_1. \end{cases}$$

Then $g_\epsilon(y_1) = y_1 - \frac{\epsilon}{2}$, $g_\epsilon'(y_1) = 0$, $g_\epsilon(y_1 - \epsilon) = y_1 - \epsilon$, $g_\epsilon'(y_1 - \epsilon) = 1$, and similar equations hold for y_0 and $y_0 + \epsilon$ so that $g_\epsilon \in C^1(\mathbb{R})$ for any $\epsilon > 0$. Note that $f \circ g_\epsilon \to \widetilde{f}$, $(f \circ g_\epsilon)' \to \widetilde{f}'$ uniformly on compacta, with $f \circ g_\epsilon$, $\widetilde{f} \in C^1(\mathbb{R})$: Namely, we have $g_\epsilon' = 1$ in $(y_0 + \epsilon, y_1 - \epsilon)$ and $0 \le g_\epsilon' \le 1$ in $(y_1 - \epsilon, y_1)$, $g_\epsilon' = 0$ in (y_1, ∞). Since $g_\epsilon|_{I_\epsilon} = \mathrm{Id}|_{I_\epsilon}$, we have $Tg_\epsilon|_{I_\epsilon} = 1$ by Lemma 6.7(c). Thus by (6.1) for all $x \in I_\epsilon$

$$T(f \circ g_\epsilon)(x) \le Tf(g_\epsilon(x))Tg_\epsilon(x) = Tf(x).$$

Now the pointwise continuity of T implies for all $x \in (y_0, y_1)$

$$T\widetilde{f}(x) = \lim_{\epsilon \to 0} T(f \circ g_\epsilon)(x) \le Tf(x).$$

Together with part (ii), we get

$$T\widetilde{f}(x) = Tf(x) = F\big(x, f(x), f'(x)\big), \quad x \in (y_0, y_1).$$

(iv) We now know that (6.4) holds for all $f \in C^1(\mathbb{R})$ and all open intervals (y_0, y_1) of strict monotonicity of f. On intervals J where f is constant, $Tf|_J = 0$ by Lemma 6.7(b), and $F(x, y, 0) = 0$ is a result of continuity arguments like $\lim_{\epsilon \to 0} T\widetilde{f}_\epsilon(x) = T\widetilde{f}(x)$ for boundary points of J together with $Tf|_J = 0$. Equation (6.4) then means $0 = 0$. Equation (6.4) similarly extends to limit points of intervals of monotonicity of f or to limit points of intervals of constancy of f. Hence (6.4) holds for all $f \in C^1(\mathbb{R})$ and all $x \in I$. □

Proof of Theorem 6.1. (a) By Proposition 6.6 there is $F : \mathbb{R}^3 \to \mathbb{R}$ such that for all $f \in C^1(\mathbb{R})$, $x \in \mathbb{R}$,

$$Tf(x) = F\big(x, f(x), f'(x)\big).$$

The chain rule inequality is equivalent to the functional inequality for F,

$$F(x, z, \alpha\beta) \le F(y, z, \alpha)F(x, y, \beta) \tag{6.5}$$

for all $x, y, z, \alpha, \beta \in \mathbb{R}$. Just choose $f, g \in C^1(\mathbb{R})$ with $g(x) = y$, $f(y) = z$, $g'(x) = \beta$, $f'(y) = \alpha$. The equations $Tc = 0$, $T(\mathrm{Id}) = 1$ imply that

$$F(x, y, 0) = 0, \quad F(x, x, 1) = 1. \tag{6.6}$$

Note that $F(x,y,1) = T(S_{y-x})(x)$ where $S_{y-x} = \mathrm{Id} + (y-x)$ is the shift by $y-x$. Since $T(S_{y-x})(x)$ depends continuously on $y-x$, cf. the proof of (a) of Lemma 6.7, and since by (6.5) and (6.6)

$$1 = F(x,x,1) \le F(y,x,1)F(x,y,1),$$

we first get that $F(x,y,1) \neq 0$ and then $F(x,y,1) > 0$ for all $x, y \in \mathbb{R}$. We showed in the proof of (a) of Lemma 6.7 that $T(-\mathrm{Id})(0) \le -1$. Hence, $F(0,0,-1) = T(-\mathrm{Id})(0) \le -1$ and for any $x \in \mathbb{R}$

$$F(x,x,-1) \le F(0,x,1)F(0,0,-1)F(x,0,1) \le -1$$

using $1 = F(0,0,1) \le F(0,x,1)F(x,0,1)$.

(b) Fix $x_0 \in \mathbb{R}$ and put $K(\alpha) := F(x_0,x_0,\alpha)$ for $\alpha \in \mathbb{R}$. By (6.5) for $x = y = z = x_0$, K is submultiplicative on $\mathbb{R}$ with $K(-1) < 0 < K(1)$. Further K is continuous as implied by the pointwise continuity of T: Assume $\alpha_n \to \alpha$ in $\mathbb{R}$. Consider $f_n(x) := \alpha_n(x-x_0)+x_0$, $f(x) := \alpha(x-x_0)+x_0$. Then $f_n(x_0) = f(x_0) = x_0$ and $f_n'(x) = \alpha_n \to \alpha = f'(x)$. Hence, $f_n \to f$, $f_n' \to f'$ converge uniformly on compacta and therefore $Tf_n(x_0) \to Tf(x_0)$, which means

$$K(\alpha_n) = F(x_0,x_0,\alpha_n) = Tf_n(x_0) \to Tf(x_0) = F(x_0,x_0,\alpha) = K(\alpha).$$

Theorem 6.2 yields that there are $p(x_0) > 0$ and $A(x_0) = |F(x_0,x_0,-1)| \ge 1$ such that

$$K(\alpha) = \begin{cases} \alpha^{p(x_0)}, & \alpha \ge 0, \\ -A(x_0)|\alpha|^{p(x_0)}, & \alpha < 0. \end{cases} \tag{6.7}$$

For any $x, y, z \in \mathbb{R}$ by (6.5)

$$F(x,x,\alpha) \le F(z,x,1)F(z,z,\alpha)F(x,z,1) = d(x,z)F(z,z,\alpha),$$

where $d(x,z) := F(z,x,1)F(x,z,1) \ge 1$ is a number independent of α. Fixing x, z with $x \neq z$, we have for all $\alpha > 0$ that $\alpha^{p(x)-p(z)} \le d(x,z)$. If $p(x) \neq p(z)$, we would get a contradiction for either $\alpha \to 0$ or for $\alpha \to \infty$. Hence, the exponent $p := p(x)$ is independent of $x \in \mathbb{R}$.

(c) We next analyze the form of $F(x,z,\alpha)$ for $x \neq z$. Let $x, z \in \mathbb{R}$, $x \neq z$. By (6.5) and (6.7) for all $\alpha > 0$, $\beta \in \mathbb{R}$,

$$F(x,z,\alpha\beta) \le F(x,z,\beta)F(x,x,\alpha) = \alpha^p F(x,z,\beta)$$

and

$$F(x,z,\beta) \le F(x,z,\alpha\beta)F\left(x,x,\frac{1}{\alpha}\right) = \frac{1}{\alpha^p}F(x,z,\alpha\beta).$$

Therefore,

$$F(x,z,\alpha\beta) \le \alpha^p F(x,z,\beta) \le F(x,z,\alpha\beta),$$

and we have equality $F(x,z,\alpha\beta) = \alpha^p F(x,z,\beta)$. Putting here $\beta = 1$ and $\beta = -1$, we find that

$$F(x,z,\alpha) = \begin{cases} F(x,z,1)\alpha^p, & \alpha \geq 0, \\ F(x,z,-1)|\alpha|^p, & \alpha < 0. \end{cases} \tag{6.8}$$

We know that $F(x,z,1) > 0$. On the other hand,

$$F(x,z,-1) \leq F(0,z,1)F(0,0,-1)F(x,0,1) < 0.$$

Let $c_\pm(x,z) := F(x,z,\pm 1)$ and $a(x,z) := |c_-(x,z)|/c_+(x,z)$. Since

$$c_-(x,z) = F(x,z,-1) \leq F(x,z,1)F(x,x,-1) \leq -F(x,z,1) = -c_+(x,z),$$

we have $a(x,z) \geq 1$ for all $x, z \in \mathbb{R}$. Choose $\alpha, \beta \in \{+1,-1\}$ in (6.5) to find that

$$\begin{aligned} c_+(x,z) &\leq c_+(y,z)c_+(x,y), \\ c_-(x,z) &\leq c_-(y,z)c_+(x,y) \quad \text{and} \\ c_-(x,z) &\leq c_+(y,z)c_-(x,y). \end{aligned}$$

Using these inequalities and $c_-(x,z) < 0$, we get

$$\begin{aligned} c_+(x,z)\max\bigl(a(y,z),a(x,y)\bigr) &\leq c_+(y,z)c_+(x,y)\max\bigl(a(y,z),a(x,y)\bigr) \\ &= \max\bigl(|c_-(y,z)|c_+(x,y), c_+(y,z)|c_-(x,y)|\bigr) \\ &\leq |c_-(x,z)| = c_+(x,z)a(x,z). \end{aligned} \tag{6.9}$$

Since $c_+(x,z) > 0$, this implies for all $x,y,z \in \mathbb{R}$ that $\max(a(y,z),a(x,y)) \leq a(x,z)$, which yields $a(x,y) \leq a(x,0) \leq a(0,0)$ and $a(0,0) \leq a(x,0) \leq a(x,y)$. Therefore, a is constant, $a(x,y) = a(0,0)$ for all $x,y \in \mathbb{R}$. Let $A := a(0,0)$. Then $A \geq 1$ and $c_-(x,z) = -A\,c_+(x,z)$. Since we now have equalities everywhere in (6.9), we conclude $c_+(x,z) = c_+(y,z)c_+(x,y)$. For $y = 0$, $c_+(x,z) = c_+(0,z)c_+(x,0)$, $1 = c_+(x,x) = c_+(0,x)c_+(x,0)$. Put $H(x) := c_+(0,x)$. Then $H > 0$ and $c_+(x,z) = \frac{H(z)}{H(x)}$. Hence, by (6.8),

$$F(x,z,\alpha) = \begin{cases} \frac{H(z)}{H(x)}\alpha^p, & \alpha \geq 0, \\ -A\frac{H(z)}{H(x)}|\alpha|^p, & \alpha < 0. \end{cases}$$

Note that $H(z) = F(0,z,1) = T(S_z)(0)$ depends continuously on z. Finally, using (6.4), we have

$$Tf(x) = \left\{ \begin{matrix} \frac{H\circ f(x)}{H(x)} f'(x)^p & f'(x) \geq 0 \\ -A\frac{H\circ f(x)}{H(x)}|f'(x)|^p & f'(x) < 0 \end{matrix} \right\}; \qquad f \in C^1(\mathbb{R}),\ x \in \mathbb{R}.$$

This ends the proof of Theorem 6.1. $\square$

The proof of Theorem 6.3 is similar to the one of Theorem 6.1.

6.4 Rigidity of the chain rule

In Theorem 5.8 we showed that the chain rule is rigid: the perturbed chain rule equation

$$T(f \circ g)(x) - Tf \circ g(x) \cdot Tg(x) = B\big(x, f \circ g(x), g(x)\big) \tag{6.10}$$

under weak conditions implies that $B \equiv 0$ and that (6.10) has the same solutions as the unperturbed chain rule. We now consider an extension of (6.10) and study the more general inequality

$$\big|T(f \circ g)(x) - Tf \circ g(x) \cdot Tg(x)\big| \le B\big(x, f \circ g(x), g(x)\big). \tag{6.11}$$

Theorem 5.8 required no continuity assumption on T. Since (6.11) allows more freedom than (6.10), we need a stronger condition of non-degeneration of T to solve (6.11). We also assume that T is pointwise continuous.

Definition. An operator $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ is *strongly non-degenerate* provided that, for all open intervals $I \subset \mathbb{R}$, all $x \in I$ and all $t > 0$, there are functions $f_1, f_2 \in C^1(\mathbb{R})$ with $f_1(x) = f_2(x) = x$, $\operatorname{Im}(f_1) \subset I$, $\operatorname{Im}(f_2) \subset I$, and $Tf_1(x) > t$, $Tf_2(x) < -t$.

Note that the model chain rule equality has derivative-type solutions, and then these assumptions are clearly satisfied.

We then have the following rigidity result for the chain rule.

Theorem 6.8 (Strong rigidity of the chain rule)**.** *Assume that $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ is strongly non-degenerate and pointwise continuous. Suppose there is a function $B : \mathbb{R}^3 \to \mathbb{R}$ such that T satisfies*

$$\big|T(f \circ g)(x) - Tf \circ g(x) \cdot Tg(x)\big| \le B\big(x, f \circ g(x), g(x)\big). \tag{6.11}$$

for all $f, g \in C^1(\mathbb{R})$, $x \in \mathbb{R}$. Assume also that there is $x_0 \in \mathbb{R}$ such that $T(-\mathrm{Id})(x_0) < 0$. Then (6.11) *has the same solutions as the unperturbed chain rule, i.e., B can be chosen to be zero: There is $p > 0$ and a function $H \in C(\mathbb{R})$, $H > 0$, such that*

$$Tf(x) = \frac{H \circ f(x)}{H(x)} |f'(x)|^p \operatorname{sgn} f'(x), \quad f \in C^1(\mathbb{R}),\ x \in \mathbb{R}.$$

The proof of this theorem relies on the follow localization result.

Proposition 6.9. *Under the assumptions of Theorem* 6.8*, there is a function $F : \mathbb{R}^3 \to \mathbb{R}$ such that, for all $f \in C^1(\mathbb{R})$ and all $x \in \mathbb{R}$,*

$$Tf(x) = F\big(x, f(x), f'(x)\big).$$

Proof. Using Proposition 3.3, it suffices to show that for any open interval $I \subset \mathbb{R}$ and $f_1, f_2 \in C^1(\mathbb{R})$ with $f_1|_I = f_2|_I$ we have $Tf_1|_I = Tf_2|_I$. Let $x \in I$. Since T

is strongly non-degenerate, we may choose functions $g_n \in C^1(\mathbb{R})$ with $g_n(x) = x$, $\mathrm{Im}(g_n) \subset I$ and $\lim_{n\to\infty} Tg_n(x) = \infty$. Then by (6.11)

$$-B\big(x, f_1(x), x\big) \le T(f_1 \circ g_n)(x) - Tf_1(x) \cdot Tg_n(x) \le B\big(x, f_1(x), x\big).$$

Since $\lim_{n\to\infty} \frac{B(x,f_1(x),x)}{Tg_n(x)} = 0$, we get by dividing the previous inequalities by $Tg_n(x)$ that

$$Tf_1(x) = \lim_{n\to\infty} \frac{T(f_1 \circ g_n)(x)}{Tg_n(x)},$$

where the limit exists. Note that $f_1 \circ g_n = f_2 \circ g_n$ since $f_1|_I = f_2|_I$. Therefore, $Tf_1(x) = Tf_2(x)$ and consequently $Tf_1|_I = Tf_2|_I$. □

Using Proposition 6.9, the operator chain rule inequality (6.11) for T is equivalent to the functional inequality for F:

$$\big|F(x, z, \alpha\beta) - F(y, z, \alpha)F(x, y, \beta)\big| \le B(x, z, y), \tag{6.12}$$

for all $x, y, z, \alpha, \beta \in \mathbb{R}$. For $x = y = z$ and $\phi_x := F(x, x, \cdot)$, $d_x := B(x, x, x)$, this means

$$\big|\phi_x(\alpha\beta) - \phi_x(\alpha)\phi_x(\beta)\big| \le d_x. \tag{6.13}$$

Since T is strongly non-degenerate, $\overline{\lim}_{\alpha\in\mathbb{R}} \phi_x(\alpha) = \infty$, $\underline{\lim}_{\alpha\in\mathbb{R}} \phi_x(\alpha) = -\infty$. Actually, we can show that $\overline{\lim}_{\alpha\to\infty} \phi_x(\alpha) = \infty$, $\underline{\lim}_{\alpha\to-\infty} \phi_x(\alpha) = -\infty$, cf. [KM10]. The pointwise continuity of T implies that $\phi_x : \mathbb{R} \to \mathbb{R}$ is continuous. These facts suffice to show that the *nearly multiplicative* function ϕ_x is actually multiplicative:

Proposition 6.10. *Suppose that* $\phi : \mathbb{R} \to \mathbb{R}$ *is continuous with* $\overline{\lim}_{\alpha\to\infty} \phi_x(\alpha) = \infty$ *and* $\underline{\lim}_{\alpha\to-\infty} \phi_x(\alpha) = -\infty$. *Assume also that there is* $d \in \mathbb{R}$ *such that for all* $\alpha, \beta \in \mathbb{R}$

$$\big|\phi(\alpha\beta) - \phi(\alpha)\phi(\beta)\big| \le d. \tag{6.14}$$

Then ϕ *is multiplicative, i.e.,* d *may be chosen zero, and there is* $p > 0$ *such that*

$$\phi(\alpha) = |\alpha|^p \operatorname{sgn} \alpha.$$

Proof. Choose $\beta_n \in \mathbb{R}$ such that $0 < \phi(\beta_n) \to \infty$. Then by (6.14)

$$\left|\frac{\phi(\alpha\beta_n)}{\phi(\beta_n)} - \phi(\alpha)\right| \le \frac{d}{\phi(\beta_n)} \to 0,$$

and hence $\phi(\alpha) = \lim_{n\to\infty} \frac{\phi(\alpha\beta_n)}{\phi(\beta_n)}$, where the limit exists for all $\alpha \in \mathbb{R}$. In particular, $\phi(0) = 0$, $\phi(1) = 1$. We conclude that for any $\alpha, \gamma \in \mathbb{R}$

$$\phi(\alpha)\phi(\gamma) = \lim_{n\to\infty} \frac{\phi(\alpha\beta_n)}{\phi(\beta_n)} \frac{\phi(\gamma\beta_n)}{\phi(\beta_n)}.$$

Now $\phi(\alpha\,\beta_n)\phi(\gamma\,\beta_n) \le \phi(\alpha\,\gamma\,\beta_n^2) + d$ and $\phi(\beta_n)\phi(\beta_n) \ge \phi(\beta_n^2) - d$. Hence

$$\phi(\alpha)\phi(\gamma) \le \lim_{n\to\infty} \frac{\phi(\alpha\gamma\beta_n^2) + d}{\phi(\beta_n^2) - d} = \lim_{n\to\infty} \frac{\phi(\alpha\gamma\beta_n^2)}{\phi(\beta_n^2)} = \phi(\alpha\gamma),$$

since $\phi(\beta_n^2) \to \infty$, too, in view of $|\phi(\beta_n)^2 - \phi(\beta_n^2)| \le d$. Similarly $\phi(\alpha\gamma) \ge \phi(\alpha)\phi(\gamma)$. Therefore ϕ is multiplicative and continuous, with negative values for $\alpha \to -\infty$. Proposition 2.3 implies that there is $p > 0$ such that $\phi(\alpha) = |\alpha|^p \operatorname{sgn}\alpha$. □

Proof of Theorem 6.8. By Proposition 6.9, $Tf(x) = F(x, f(x), f'(x))$, where F satisfies (6.12). We analyze the form of F. By Proposition 6.10 and (6.13), there is $p(x) > 0$ such that $F(x, x, \alpha) = \phi_x(\alpha) = \alpha^{p(x)}$ for any $\alpha > 0$. For $x \ne z$, by choosing successively $y = x$ and $y = z$ in (6.12), we find

$$\left|F(x, z, \alpha\beta) - F(x, z, \alpha)\beta^{p(x)}\right| \le B(x, z, x), \quad \beta > 0,\ \alpha \in \mathbb{R},$$

and

$$\left|F(x, z, \alpha\beta) - \alpha^{p(z)} F(x, z, \beta)\right| \le B(x, z, z), \quad \alpha > 0,\ \beta \in \mathbb{R}.$$

Exchange α and β in the first inequality. Then the triangle inequality yields $|\alpha^{p(x)} - \alpha^{p(z)}||F(x, z, \beta)| \le B(x, z, x) + B(x, z, z)$ for any $\alpha > 0$, $\beta \in \mathbb{R}$. This obviously implies $p(x) = p(z)$ for $\beta = 1$, $\alpha \to \infty$, since $F(x, z, 1) \ne 0$, which is an easy consequence of (6.12). Thus for any $x, \alpha \in \mathbb{R}$

$$F(x, x, \alpha) = |\alpha|^p \operatorname{sgn}\alpha$$

with $p := p(x) = p(z) > 0$. Since $\frac{B(x,z,x)}{\beta^p} \to 0$ for $\beta \to \infty$, we also conclude for all $\alpha \in \mathbb{R}$

$$F(x, z, \alpha) = \lim_{\beta\to\infty} \frac{F(x, z, \alpha\beta)}{\beta^p}. \tag{6.15}$$

For any $\alpha > 0$, $\alpha\beta \to \infty$ if $\beta \to \infty$, and therefore

$$F(x, z, \alpha) = \lim_{\beta\to\infty} \frac{F(x, z, \alpha\beta)}{\beta^p} = \alpha^p \lim_{\alpha\beta\to\infty} \frac{F(x, z, \alpha\beta)}{(\alpha\beta)^p} = \alpha^p F(x, z, 1)$$

for any $x, z \in \mathbb{R}$. For $\alpha < 0$ we have

$$F(x, z, \alpha) = \lim_{\beta\to\infty} \frac{F(x, z, \alpha\beta)}{\beta^p} = |\alpha|^p \lim_{|\alpha|\beta\to\infty} \frac{F(x, z, -|\alpha|\beta)}{|\alpha|^p\beta^p} = |\alpha|^p F(x, z, -1).$$

Dividing (6.12) by $(\alpha\beta)^p$ for $\alpha, \beta > 0$, we get

$$\left|\frac{F(x, z, \alpha\beta)}{\alpha^p\beta^p} - \frac{F(y, z, \alpha)}{\alpha^p}\frac{F(x, y, \beta)}{\beta^p}\right| \le \frac{B(x, z, y)}{\alpha^p\beta^p}.$$

By (6.15), this implies $F(x, z, 1) = F(y, z, 1)F(x, y, 1)$ for all $x, y, z \in \mathbb{R}$. For $\alpha \to \pm\infty, \beta \to \mp\infty$, a similar argument yields that for all $x, y, z \in \mathbb{R}$, $F(x, z, -1) =$

$F(y,z,1)F(x,y,-1) = F(y,z,-1)F(x,y,1)$. Since $F(x,x,1) = \phi_x(1) = 1^p = 1$ for all $x \in \mathbb{R}$, $1 = F(y,x,1)F(x,y,1)$ for all $x, y \in \mathbb{R}$. Let $H(y) := F(0,y,1)$. Then $F(y,0,1) = \frac{1}{H(y)}$ and

$$F(x,z,1) = F(0,z,1)F(x,0,1) = \frac{H(z)}{H(y)},$$
$$F(x,z,-1) = F(z,z,-1)F(x,z,1) = -F(x,z,1) = -\frac{H(z)}{H(y)},$$

using that $F(z,z,-1) = \phi_z(-1) = -1$. We conclude that

$$F(x,z,\alpha) = \frac{H(z)}{H(x)}|\alpha|^p \operatorname{sgn}\alpha, \quad x,z,\alpha \in \mathbb{R}.$$

Note that $H(y) = F(0,y,1) = T(S_y)(0)$ depends continuously on y, where S_y denotes as before the shift by y. Using Proposition 6.9, we get

$$Tf(x) = F\big(x, f(x), f'(x)\big) = \frac{H \circ f(x)}{H(x)}|f'(x)|^p \operatorname{sgn} f'(x),$$

for any $f \in C^1(\mathbb{R})$, $x \in \mathbb{R}$. This solves the chain rule operator equation, so that B in (6.11) can be chosen to be zero, and proves Theorem 6.8. □

We now turn to a further extension of the chain rule, the one-sided perturbed chain rule inequality. Let $B : \mathbb{R}^3 \to \mathbb{R}$ be a function and $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ be an operator satisfying

$$T(f \circ g)(x) - Tf \circ g(x) \cdot Tg(x) \le B\big(x, f \circ g(x), g(x)\big), \tag{6.16}$$

for all $f \in C^1(\mathbb{R})$, $x \in \mathbb{R}$. This is more general than the two-sided inequality considered in Theorem 6.8, and also more general than the one-sided chain rule inequality considered in Theorem 6.1.

In the results proved so far, the operator T was localized. The operator inequality (6.16), however, is too general that localization could always be shown, even under strong non-degeneration and continuity assumptions on T. We provide an example.

Example. Let $H \in C(\mathbb{R})$ be a non-constant function with $4 \le H \le 5$. For $f \in C^1(\mathbb{R})$, $x \in \mathbb{R}$, with $f'(x) \in (-1,0)$, let $I_{f,x}$ denote the interval $I_{f,x} := [x + f'(x)(1 + f'(x)), x]$. Then $0 < |I_{f,x}| \le 1/4$. Let $Jf(x) := \frac{1}{|I_{f,x}|}\int_{I_{f,x}} f(y)dy$ denote the average of f in $I_{f,x}$. Define an operator $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ by putting, for any $f \in C^1(\mathbb{R})$, $x \in \mathbb{R}$,

$$Tf(x) := \begin{cases} \frac{H \circ f(x)}{H(x)} f'(x), & f'(x) \ge 0, \\ \frac{H \circ f(x)}{H(x)} 4f'(x), & f'(x) \le -2, \\ \frac{H \circ f(x)}{H(x)} \left(7 + \frac{15}{2} f'(x)\right), & -2 < f'(x) \le -1, \\ \frac{H \circ Jf(x)}{H(x)} \frac{1}{2} f'(x), & -1 < f'(x) < 0. \end{cases}$$

Then T satisfies, for all $f, g \in C^1(\mathbb{R})$, $x \in \mathbb{R}$,

$$T(f \circ g)(x) - Tf \circ g(x) \cdot Tg(x) \leq 5. \tag{6.17}$$

Obviously T is not localized since it depends on the integral average Jf if $f'(x) \in (-1, 0)$. Note here that for $f'(x) \geq 0$ or $f'(x) \leq -2$, $Tf(x)$ has the form given in Theorem 6.1 for $B = 0$, with $p = 1$, $A = 4$. For $-2 < \alpha = f'(x) < 0$ there is a continuous perturbation of the line 4α by $\frac{1}{2}\alpha$ if $\alpha \in (-1, 0)$ and by $7 + \frac{15}{2}\alpha$ if $\alpha \in (-2, -1]$. Note that Tf is continuous for all $f \in C^1(\mathbb{R})$: if $x_n \in \mathbb{R}$ are such that $f'(x_n) \in (-1, 0)$ and $x_n \to x$ with $f'(x_n) \to -1$ or $f'(x_n) \to 0$, then $Jf(x_n) \to f(x)$ since $|I_{f,x_n}| \to 0$.

To prove (6.17), use $\frac{4}{5} \leq \frac{H(z)}{H(y)} \leq \frac{5}{4}$ for all $y, z \in \mathbb{R}$, and distinguish the following cases: (1) $\alpha, \beta \geq 0$; (2) $\alpha, \beta \leq -2$; (3) $\alpha, \beta \in (-2, 0)$; (4) $\alpha \leq -2$, $\beta \in (-2, -1]$; (5) $\alpha \leq -2$, $\beta \in (-1, 0)$; (6) $\alpha > 0$, $\alpha\beta \leq -2$; (7) $\alpha > 0$, $\alpha\beta \in (-2, -1]$; (8) $\alpha > 0$, $\alpha\beta \in (-1, 0)$. The estimates to show (6.17) are easy in each case but a bit tedious. They can be found in detail in [KM10]. □

Assuming localization in addition to (6.16), i.e., that there is a function $F : \mathbb{R}^3 \to \mathbb{R}$ such that for all $f \in C^1(\mathbb{R})$, $x \in \mathbb{R}$,

$$Tf(x) = F\big(x, f(x), f'(x)\big)$$

holds, the operator inequality (6.16) is equivalent to the functional inequality

$$F(x, z, \alpha\beta) \leq F(y, z, \alpha)F(x, z, \beta) + B(x, z, y) \tag{6.18}$$

for all $x, y, z, \alpha, \beta \in \mathbb{R}$. Similar to the two-sided case in (6.12), the most important special case to solve is the one of $x = y = z$ which means

$$\phi_x(\alpha\beta) \leq \phi_x(\alpha)\phi_x(\beta) + d_x,$$

where $\phi_x := F(x, x, \cdot\,)$ and $d_x := B(x, x, x)$. We have the following result on these *nearly submultiplicative functions.*

Theorem 6.11. *Let $\phi : \mathbb{R} \to \mathbb{R}$ be continuous with $\overline{\lim}_{\alpha\to\infty} \phi(\alpha) = \infty$. Suppose that there is $\alpha_0 < 0$ with $\phi(\alpha_0) < 0$ and that there is $d \in \mathbb{R}$ such that we have for all $\alpha, \beta \in \mathbb{R}$*

$$\phi(\alpha\beta) \leq \phi(\alpha)\phi(\beta) + d.$$

Then $d \geq 0$ and there are $p > 0$ and $A \geq 1$ such that for all $\alpha > 0$

$$\phi(\alpha) = \alpha^p\,, \quad -A\alpha^p \leq \phi(-\alpha) \leq \min\left(-\tfrac{1}{A}\alpha^p, -A\alpha^p + d\right).$$

Moreover, the limit $\lim_{\alpha\to\infty} \frac{\phi(-\alpha)}{-\alpha^p}$ exists and $A = \lim_{\alpha\to\infty} \frac{\phi(-\alpha)}{-\alpha^p}$.

Remarks. For $d \neq 0$, $\phi|_{\mathbb{R}_{<0}}$ is not of power type, but is close to the power-type function $-A|\alpha|^p$ for large $|\alpha|$, $\alpha < 0$. Interestingly enough, $\phi|_{\mathbb{R}_{\geq 0}}$ is of power type α^p. For $p = 1$, $A = 2$, the function

$$\phi(\alpha) := \begin{cases} \alpha, & \alpha \geq 0, \\ \frac{1}{2}\alpha, & \alpha \in [-1, 0), \\ 3 + \frac{7}{2}\alpha, & \alpha \in [-2, -1), \\ 2\alpha, & \alpha < -2 \end{cases}$$

provides an explicit example satisfying the assumptions of Theorem 6.11 with $d = \frac{3}{2}$, $\phi(\alpha\beta) \leq \phi(\alpha)\phi(\beta) + \frac{3}{2}$, which is not of power type on $\mathbb{R}_{<0}$.

The proof of Theorem 6.11 is an asymptotic modification of the one of Theorem 6.2 for submultiplicative functions when $d = 0$. We only provide the essential steps of the proof, which are:

(a) Show $\lim_{\alpha \to \infty} \phi(\alpha) = \infty$, $\lim_{\alpha \to -\infty} \phi(\alpha) = -\infty$.

(b) Choose $b > 1$ close to 1. Let $\phi_1 := b\phi$, $\phi_2 := \frac{1}{b}\phi$. Then for large $\gamma_0 = \gamma_0(b)$, $\phi_1(\alpha\beta) \leq \phi_1(\alpha)\phi_1(\beta)$ if $\alpha\beta \geq \gamma_0$ and $\phi_2(\alpha\beta) \leq \phi_2(\alpha)\phi_2(\beta)$ if $\alpha\beta \leq -\gamma_0$: ϕ_1 and ϕ_2 are submultiplicative for large $\alpha\beta$ in the positive, respectively negative range.

(c) Define $f(t) := \ln \phi_1(\exp(t))$, $g(t) := \ln |\phi_2(-\exp(t))|$ for $t \in \mathbb{R}$. Then, for $t_0 := \ln \gamma_0$,

$$0 \leq p := \inf_{t \geq t_0} \frac{f(t)}{t} = \lim_{t \to \infty} \frac{f(t)}{t} < \infty, \quad f(t) = pt + a(t).$$

(d) For $t, s \in \mathbb{R}$ with $t + s \geq t_0$,

$$g(t+s) \geq g(t) + f(s) - 2\ln b.$$

Using $0 \leq f(t) + f(-t)$ for all $t \in \mathbb{R}$, show that

$$\big|g(t) - [c + pt + a(t - t_0)]\big| \leq 2\ln b,$$

for all $t \geq t_0$, where $c := g(t_0) - pt_0$, and a satisfies $\lim_{t \to \infty} \frac{a(t)}{t} = 0$ with $a(t) \geq 0$ for $t \geq t_0$.

(e) Improve the bound for a to $a(t) \leq 6\ln b$ for all $t \geq t_1$, for a suitable $t_1 \geq t_0$. Then $\phi_1(\alpha) = \alpha^p \exp(a(\ln \alpha))$, $\alpha > 0$ is asymptotically α^p for large α, if b is close to 1 and thus $\ln b$ is close to 0. Further $\phi_2(\alpha) \simeq -A|\alpha|^p \exp\big(a\big(\ln\big|\frac{\alpha}{\alpha_1}\big|\big)\big)$, $\alpha_1 := \exp(t_1)$, for large negative α.

(f) Use $\phi = \frac{1}{b}\phi_1 = b\phi_2$, take the limit as $b \to 1$ and prove that $\phi(\alpha) = \alpha^p$ for $\alpha > 0$ and $\lim_{\alpha \to -\infty} \frac{\phi(\alpha)}{-|\alpha|^p} = A$.

We do not give the details here, but refer to [KM10].

Using Theorem 6.11, we may prove the following result on the one-sided perturbed chain rule inequality, assuming localization which cannot be guaranteed otherwise, as shown by the previous example.

Theorem 6.12. *Assume that $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ is strongly non-degenerate, pointwise continuous, and that there is $x_0 \in \mathbb{R}$ with $T(-\operatorname{Id})(x_0) < 0$. Suppose that there is a function $B : \mathbb{R}^3 \to \mathbb{R}$ such that the perturbed chain rule inequality*

$$T(f \circ g)(x) \le Tf \circ g(x) \cdot Tg(x) + B\big(x, f \circ g(x), g(x)\big)$$

holds for all $f, g \in C^1(\mathbb{R})$, $x \in \mathbb{R}$. Assume also that there is $F : \mathbb{R}^3 \to \mathbb{R}$, so that

$$Tf(x) = F\big(x, f(x), f'(x)\big), \quad f \in C^1(\mathbb{R}),\ x \in \mathbb{R}.$$

Then there are $p > 0$, $A \ge 1$, $H \in C(\mathbb{R})$, $H > 0$ and a function $K : \mathbb{R}^2 \times \mathbb{R}_{<0} \to \mathbb{R}_{<0}$ which is continuous in the second and the third variable satisfying

$$\begin{aligned} -A\alpha^p &\le K(x, z, -\alpha) \\ &\le \min\left(-\frac{1}{A}\alpha^p, -A\alpha^p + \frac{H(x)}{H(z)} \min\big[B(x,z,x), B(x,z,z)\big]\right) \end{aligned}$$

for all $x, z \in \mathbb{R}$, $\alpha > 0$, and for which $A = \lim_{\beta\to\infty} \frac{K(x,z,-\beta)}{-\beta^p}$ exists for all $x, z \in \mathbb{R}$, the limit A being independent of x and z, such that for all $f \in C^1(\mathbb{R})$ and $x \in \mathbb{R}$

$$Tf(x) = \begin{cases} \frac{H\circ f(x)}{H(x)} f'(x)^p, & f'(x) \ge 0, \\ \frac{H\circ f(x)}{H(x)} K\big(x, f(x), f'(x)\big), & f'(x) < 0. \end{cases}$$

The property of K means that for negative values of $f'(x)$, $Tf(x)$ is reasonably close to $-A\frac{H\circ f(x)}{H(x)}|f'(x)|^p$, deviating from this value by at most

$$\min[B(x, f(x), x), B(x, f(x), f(x))],$$

i.e., deviating by at most this amount from the solution in Theorem 6.1 for $B = 0$.

For the proof of Theorem 6.12 we refer to [KM10]. Theorem 6.11 has an analogue for nearly supermultiplicative functions $\phi(\alpha\beta) \ge \phi(\alpha)\phi(\beta) - d$ and Theorem 6.12 has an analogue for the perturbed supermultiplicative operator inequality

$$T(f \circ g)(x) \ge Tf \circ g(x) \cdot Tg(x) - B\big(x, f \circ g(x), g(x)\big).$$

6.5 Notes and References

The result on the chain rule inequality, Theorem 6.1 and Proposition 6.6 on the localization of the operator T were shown by König and Milman in [KM9]. Theorem 6.2 on submultiplicative functions on $\mathbb{R}$ is also found in [KM9].

The proof of Lemma 6.5 on subadditive functions on $\mathbb{R}$ follows Hille, Phillips, [HP, Chap. VII]. If additionally in this lemma f is continuous at 0 and $f(0) = 0$ holds, f is continuous at any $t \in \mathbb{R}$, cf. also Hille, Phillips, [HP]. A simpler variant of Lemma 6.5 for sequences goes back to Fekete [Fe], p. 233, cf. also Pólya, Szegö [PS], Problem 98.

The rigidity result Theorem 6.8 for the is taken from [KM10]. Theorems 6.11 and 6.12 as well as the example before Theorem 6.11 are shown in [KM10], too.

Submultiplicative maps may not only be considered on the real line as in Theorem 6.2, but also on function spaces like $C^k(I)$. But, as in the case of multiplicative operators, cf. [M], [MS], [AAFM] or [AFM], one assumes that the mapping is bijective. Let us call $T : C^k(I) \to C^k(I)$ *C^k-pointwise continuous* provided that for all sequences $f_n \in C^k(I)$ and all $f \in C^k(I)$ with $f_n^{(j)} \to f^{(j)}$ converging uniformly on compact subsets of I for all $j \in \{0, \dots, k\}$ we have that $Tf_n(x) \to Tf(x)$ converges for all $x \in I$. Then the following result holds for submultiplicative operators, cf. Faifman, König and Milman [FKM]:

Proposition 6.13. *Let $I \subset \mathbb{R}$ be open and $k \in \mathbb{N}_0$. Suppose that $T : C^k(I) \to C^k(I)$ is bijective, C^k-pointwise continuous and submultiplicative, i.e.,*

$$T(f \cdot g)(x) \le Tf(x) \cdot Tg(x), \quad f, g \in C^k(I),\ x \in I. \tag{6.19}$$

Assume also that $T(-\mathbb{1}) < 0$ and that $Tf \ge 0$ holds if and only if $f \ge 0$ for all $f \in C^k(I)$. Then there exist a homeomorphism $u : I \to I$ and two continuous functions $p, A \in C(I)$ with $A \ge 1$, $p > 0$ such that

$$(Tf)(u(x)) = \begin{cases} f(x)^{p(x)}, & f(x) \ge 0, \\ -A(x)\ |f(x)|^{p(x)}, & f(x) < 0. \end{cases}$$

Conversely, T defined this way satisfies (6.19).

For $k \in \mathbb{N}$, we have that $A = p = 1$ and that u is a C^k-diffeomorphism, so that

$$Tf(u(x)) = f(x).$$

Thus, for $k \in \mathbb{N}$, the operator is even *multiplicative and linear.*

We indicate some steps of the proof.

Step 1. For $x \in I$, an approximate indicator at x is a function $f \in C^k(I)$ with $f \ge 0$ such that there are open neighborhoods $x \in J_1 \subset J_2$ of x with $f|_{J_1} = \mathbb{1}$ and $f|_{I \setminus J_2} = 0$. Let AI_x denote the set of all approximate indicators at x. Define a set-valued map from I to the subsets of I by $u(x) := \bigcap_{f \in AI_x} \operatorname{supp}(Tf)$, where $\operatorname{supp}(Tf)$ denotes the support of Tf. One shows that $u(x)$ is either empty or consists of only one point and that for $f \in AI_x$, $Tf|_{u(x)} = 1$. Also for any $f \in C^k(I)$ and $x \in I$, $\operatorname{sgn} Tf|_{u(x)} = \operatorname{sgn} f(x)$. Here the fact that $f \ge 0$ implies $Tf \ge 0$ is used.

Step 2. Let G denote the set of all $x \in I$ for which there is an approximate indicator $f \in AI_x$ with compact support. Then obviously, $u(x)$ is not empty and hence consists of one point, and $u : G \to u(G) \subset I$ can be considered as a point map. Using among other things that $Tf \geq 0$ implies $f \geq 0$, one proves that $u(G)$ and G are dense in I and that $u : G \to u(G) \subset I$ is continuous and injective.

Step 3. One shows that for any open subset $J \subset I$ and any $f_1, f_2 \in C^k(I)$ with $f_1|_J = f_2|_J$ we have that $Tf_1|_{u(J)} = Tf_2|_{u(J)}$, after proving for any $h \in C^k(I)$ that $h|_J = \mathbb{1}$ implies $Th|_{u(J)} = \mathbb{1}$ and that $h|_J = -\mathbb{1}$ implies $Th|_{u(J)} = T(-\mathbb{1})|_{u(J)}$. This yields the localization of T on G: There is F such that

$$Tf(u(x)) = F(x, f(x), \ldots, f^{(k)}(x))$$

for any $f \in C^k(I)$ and $x \in G$. Moreover, $\operatorname{sgn} F(x, \alpha_0, \ldots, \alpha_k) = \operatorname{sgn} \alpha_0$.

Step 4. The operator inequality for T translates into a functional inequality for F. One proves that F does not depend on the variables $(\alpha_1, \ldots, \alpha_k)$. Theorem 6.2 then yields that for any $f \in C^k(I)$ and all $x \in G$

$$(Tf)(u(x)) = \begin{cases} f(x)^{p(x)}, & f(x) \geq 0, \\ -A(x)\ |f(x)|^{p(x)}, & f(x) < 0, \end{cases}$$

where $A \geq 1$ and $p \geq 0$ are continuous functions on G. The functions and operators are then extended by continuity to all of I, with $u : I \to I$ being a homeomorphism. For $k \in \mathbb{N}$, considering the inverse operator expressed with powers $\frac{1}{p(x)}$, shows that $A = p = 1$ and that u is a C^k-diffeomorphism. □