Chapter 4

The Chain Rule

4.1 The chain rule on $C^k(\mathbb{R})$

The derivative $D : C^1(\mathbb{R}) \to C(\mathbb{R})$ satisfies the chain rule

$$D(f \circ g) = (Df) \circ g \cdot Dg$$

for all $f, g \in C^1(\mathbb{R})$. In this chapter, we study the question to which extent the chain rule formula characterizes the derivative. We consider general operators $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ satisfying the *chain rule operator equation*

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

Due to the multiplicative structure of this equation, if T_1 and T_2 are operators satisfying the chain rule, so does the pointwise product $T_1 \cdot T_2$, and also do the positive powers of the pointwise modulus $|T_1|$. Suppose $H \in C(\mathbb{R})$ is a strictly positive continuous function. Then $Tf := H \circ f / H$ defines a map satisfying the chain rule as well. It is even defined on $C(\mathbb{R})$, not only on $C^1(\mathbb{R})$. Another example of a map $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ verifying the chain rule is given by

$$Tf := \begin{cases} f', & f \in C^1(\mathbb{R}) \text{ is bijective,} \\ 0, & f \in C^1(\mathbb{R}) \text{ is not bijective.} \end{cases}$$

To avoid degenerate cases like this one, we impose the condition that T should not be identically zero on the half-bounded differentiable functions

$$C_b^1(\mathbb{R}) := \{ f \in C^1(\mathbb{R}) \mid f \text{ is bounded from above or from below} \},$$

i.e., that there exist $f \in C_b^1(\mathbb{R})$ and $x \in \mathbb{R}$ such that $Tf(x) \neq 0$. For integers $k \in \mathbb{N}$, we also let $C_b^k(\mathbb{R}) := C^k(\mathbb{R}) \cap C_b^1(\mathbb{R})$.

Our main result states that a multiplicative combination of the previous examples, together with a possible factor $\operatorname{sgn} f'$, creates all possible solutions of

H. König, V. Milman, *Operator Relations Characterizing Derivatives*,
https://doi.org/10.1007/978-3-030-00241-1_4

the chain rule equation not only on $C^1(\mathbb{R})$ but also on $C^k(\mathbb{R})$ for any $k \in \mathbb{N}$. Again, all solutions operators are local, i.e., pointwise defined.

Theorem 4.1 (Chain rule)**.** *Let $k \in \mathbb{N} \cup \{\infty\}$ and $T : C^k(\mathbb{R}) \to C(\mathbb{R})$ be an operator satisfying the* chain rule equation

$$T(f \circ g) = (Tf) \circ g \cdot Tg, \quad f, g \in C^k(\mathbb{R}). \tag{4.1}$$

Assume that $T|_{C_b^k(\mathbb{R})} \not\equiv 0$. Then there exist $p \geq 0$ and a positive continuous function $H \in C(\mathbb{R})$, $H > 0$, such that either

$$Tf = \frac{H \circ f}{H} |f'|^p \tag{4.2}$$

or

$$Tf = \frac{H \circ f}{H} |f'|^p \operatorname{sgn} f'. \tag{4.3}$$

In the second case we need $p > 0$ to guarantee that the image of T consists of continuous functions.

If $k = 0$ and $T : C(\mathbb{R}) \to C(\mathbb{R})$ satisfies (4.1)*, all solutions of T are given by $Tf = \frac{H \circ f}{H}$.*

Conversely, the operators given by (4.2) *or* (4.3) *satisfy the chain rule equation* (4.1)*.*

Under the additional initial condition $T(2\,\mathrm{Id}) = 2$ *(constant function), T has the form $Tf = f'$ or $Tf = |f'|$.*
If additionally to (4.1)*, $T(-2\,\mathrm{Id}) = -2$ holds, T is the derivative, $Tf = f'$.*

In the formulation of similar results later we will combine statements like (4.2) and (4.3) by writing

$$Tf = \frac{H \circ f}{H} |f'|^p \{\operatorname{sgn} f'\},$$

the brackets $\{\cdot\}$ indicating that two possible solutions are given, one with the expression $\operatorname{sgn} f'$ and one without. Formulas (4.1), (4.2) and (4.3) are meant pointwise, e.g.,

$$\begin{aligned} T(f \circ g)(x) &= (Tf)(g(x)) \cdot Tg(x), \quad x \in \mathbb{R}, \\ Tf(x) &= \frac{H(f(x))}{H(x)} |f'(x)|^p \{\operatorname{sgn} f'(x)\}, \quad x \in \mathbb{R}. \end{aligned}$$

Remarks. (a) Note that we do not impose any a priori continuity condition on the operator T. A suitable level of continuity of T, however, is an a-posteriori consequence of the result.

(b) The proof will show that p and H are completely determined by the function $T(2\,\mathrm{Id}) \in C(\mathbb{R})$: we have $T(2\,\mathrm{Id}) > 0$, $p = \log_2 T(2\,\mathrm{Id})(0)$ and $H(x) = \prod_{n\in\mathbb{N}} \varphi(x/2^n)$ where φ is defined by $\varphi(x) = T(2\,\mathrm{Id})(x)/T(2\,\mathrm{Id})(0)$, and where the product converges uniformly on compact subsets of $\mathbb{R}$, with normalization $H(0) = 1$.

(c) For $C^\infty(\mathbb{R})$, there are not more solutions of the chain rule equation than for $C^1(\mathbb{R})$. Therefore, in the setup of spaces $C^k(\mathbb{R})$, the space $C^1(\mathbb{R})$ constitutes the natural domain of the chain rule. Of course, for $k = 0$, in $C(\mathbb{R})$ there is the non-surjective solution $Tf = \frac{H\circ f}{H}$ which does not depend on the derivative.

(d) For $p > 0$, let G be the antiderivative of $H^{1/p} > 0$. Then G is a strictly monotone $C^1(\mathbb{R})$-function and

$$Tf = \left|\frac{(G\circ f)'}{G'}\right|^p \{\operatorname{sgn} f'\} = \left|\frac{d(G\circ f)}{dG}\right|^p \left\{\operatorname{sgn}\left(\frac{d(G\circ f)}{dG}\right)\right\}.$$

In this sense, all solutions of (4.1) are p-th powers of some derivatives, up to signs.

As a consequence, the derivative is characterized by the Leibniz rule and the chain rule:

Corollary 4.2. *Let $k \in \mathbb{N}$ and suppose that $T : C^k(\mathbb{R}) \to C(\mathbb{R})$ satisfies the Leibniz rule and the chain rule,*

$$T(f\cdot g) = Tf\cdot g + f\cdot Tg, \quad T(f\circ g) = (Tf)\circ g\cdot Tg; \qquad f, g \in C^k(\mathbb{R}).$$

Then $T = 0$ or T is the derivative, $Tf = f'$ for all $f \in C^k(\mathbb{R})$.

Again, no continuity assumption on T is required here.

Proof of Corollary 4.2. By Theorem 3.1, T has the form $Tf = c\,f \ln|f| + d\,f'$ for suitable functions $c, d \in C(\mathbb{R})$. If $T \not\equiv 0$, c or d do not vanish identically and therefore T satisfies $T|_{C_b^k(\mathbb{R})} \not\equiv 0$. Hence, by Theorem 4.1, $Tf = \frac{H\circ f}{H}|f'|^p\{\operatorname{sgn} f'\}$ for some $p \geq 0$ and $H \in C(\mathbb{R})$, $H > 0$. Both forms of T can coincide only if $p = 1$, H is constant and $c = 0$, $d = 1$ and the $\operatorname{sgn} f'$-term occurs. Then $Tf = f'$, $f \in C(\mathbb{R})$. $\square$

Example. On suitable subsets of $C^k(I)$ or even $C(I)$, we may define operations T which satisfy the Leibniz rule and chain rule but are neither zero nor the derivative: Let $I = (1, \infty)$ and $C_+(I) := \{f : I \to I \mid f \text{ is continuous}\}$. Define $H \in C(I)$ by $H(x) = x \ln x$. Then the operator $T : C_+(I) \to C(I)$ given by $Tf = \frac{H\circ f}{H}$ is well defined and satisfies the Leibniz rule and the chain rule.

We now state a stronger version of Corollary 4.2: The derivative is also the only operator satisfying both the chain rule and the *extended* Leibniz rule studied in Theorem 3.7:

Corollary 4.3. *Suppose $T, A : C^1(\mathbb{R}) \to C(\mathbb{R})$ satisfy the chain rule and the extended Leibniz rule for all $f, g \in C^1(\mathbb{R})$,*

$$T(f \circ g) = Tf \circ g \cdot Tg \ ,$$
$$T(f \cdot g) = Tf \cdot Ag + Af \cdot Tg \ ,$$

and that T does not vanish identically on the half-bounded functions and attains some negative values. Then T is the derivative, $Tf = f'$, and $Af = f$ for all $f \in C^1(\mathbb{R})$.

Proof of Corollary 4.3. Theorem 4.1 yields that Tf is given by

$$Tf = \frac{H \circ f}{H} \ |f'|^p \ \operatorname{sgn} f'$$

for a suitable function $H \in C(\mathbb{R})$, $H > 0$ and $p > 0$. This form of Tf has to coincide with one of the solutions of the extended Leibniz rule (3.7) for $k = 1$, which were given by (3.8), (3.9) or (3.10) in Theorem 3.7. This is only possible for the first solution (3.8), and then only in the special case when $a(x) = d_1(x) = p(x) = 1$, $d_0(x) = 0$, and if the above function H satisfies $H = \mathbb{1}$ and $p = 1$, yielding $Tf = f'$, $Af = f$ for all $f \in C^1(\mathbb{R})$. $\square$

To prove Theorem 4.1 we first show, as in Chapter 3, that the operator T is localized. For this, we need that there are sufficiently many non-zero functions in the range of T.

Lemma 4.4. *Suppose the assumptions of Theorem* 4.1 *hold. Then for any open half-bounded interval $I = (c, \infty)$ or $I = (-\infty, c)$ with $c \in \mathbb{R}$, any $y \in I$ and any $x \in \mathbb{R}$, there exists $g \in C^k(\mathbb{R})$ such that $g(x) = y$, $\operatorname{Im}(g) \subset I$ and $(Tg)(x) \neq 0$.*

Proof. (i) Let $x \in \mathbb{R}$. We show that $(Tg)(x) \neq 0$ for a suitable function $g \in C_b^k(\mathbb{R})$: Since $T|_{C_b^k(\mathbb{R})} \not\equiv 0$, there is $x_1 \in \mathbb{R}$ and a half-bounded function $h \in C_b^k(\mathbb{R})$ with $(Th)(x_1) \neq 0$. Define $\varphi, g \in C_b^k(\mathbb{R})$ by

$$\varphi(s) := s + x - x_1, \quad g(s) := h \circ \varphi^{-1}(s); \qquad s \in \mathbb{R}.$$

Then $h = g \circ \varphi$, $\varphi(x_1) = x$ and

$$0 \neq (Th)(x_1) = (Tg)(\varphi(x_1)) \cdot (T\varphi)(x_1) = (Tg)(x) \cdot (T\varphi)(x_1),$$

which implies $(Tg)(x) \neq 0$. Clearly $g \in C_b^k(\mathbb{R})$.

(ii) Suppose $I = (c, \infty)$ with $c \in \mathbb{R}$. Pick any $y \in I$ and $x \in \mathbb{R}$. By (i) there is $g \in C_b^k(\mathbb{R})$ with $(Tg)(x) \neq 0$. Let J be an open half-bounded interval with $\operatorname{Im}(g) \subset J$. Choose a bijective C^k-map $f : I \to J$ with $f(y) = g(x)$, noting that $g(x) \in J$. This may be done in such a way that f is extendable to a C^k-map $\widetilde{f} : \mathbb{R} \to \mathbb{R}$ on $\mathbb{R}$, $\widetilde{f}|_I = f$. Let

$$g_1 := f^{-1} \circ g : \mathbb{R} \longrightarrow I \subset \mathbb{R}.$$

Then $g_1 \in C^k(\mathbb{R})$, $g_1(x) = y$ and $\mathrm{Im}(g_1) \subset I$. Since $g = f \circ g_1 = \widetilde{f} \circ g_1$, we find, using the chain rule equation (4.1),

$$0 \neq (Tg)(x) = (T\widetilde{f})(y) \cdot (Tg_1)(x).$$

Hence $(Tg_1)(x) \neq 0$, $g_1(x) = y$ and $\mathrm{Im}(g_1) \subset I$. □

Lemma 4.5. *Under the assumptions of Theorem* 4.1, *we have for any open, half-bounded interval I and any $f, f_1, f_2 \in C^k(\mathbb{R})$:*

(i) *If $f|_I = \mathrm{Id}$, then $(Tf)|_I = 1$.*

(ii) *If $f_1|_I = f_2|_I$, then $(Tf_1)|_I = (Tf_2)|_I$.*

Proof. (i) Assume $f|_I = \mathrm{Id}$. Take any $y \in I$, $x \in \mathbb{R}$. By Lemma 4.4, there is $g \in C^k(\mathbb{R})$ with $g(x) = y$, $\mathrm{Im}(g) \subset I$ and $(Tg)(x) \neq 0$. Then $f \circ g = g$ so that by (4.1)

$$0 \neq (Tg)(x) = T(f \circ g)(x) = (Tf)(y) \cdot (Tg)(x),$$

which implies that $(Tf)(y) = 1$. Since $y \in I$ was arbitrary, we conclude $(Tf)|_I = 1$.

(ii) Let $f_1|_I = f_2|_I$ and $x \in I$ be arbitrary. Choose a smaller open half-bounded interval $J \subset I$ and a function $g \in C^k(\mathbb{R})$ such that $x \in J$, $\mathrm{Im}(g) \subset I$ and $g|_J = \mathrm{Id}$. Then $f_1 \circ g = f_2 \circ g$ and $g(x) = x$. By part (i), $(Tg)|_J = 1$. Hence, again using the chain rule (4.1),

$$\begin{aligned}(Tf_1)(x) &= (Tf_1)(g(x)) \cdot Tg(x) = T(f_1 \circ g)(x) \\ &= T(f_2 \circ g)(x) = (Tf_2)(g(x)) \cdot Tg(x) = (Tf_2)(x),\end{aligned}$$

which shows $(Tf_1)|_I = (Tf_2)|_I$. □

Proposition 4.6. *Let $k \in \mathbb{N}_0 \cup \{\infty\}$ and $T : C^k(\mathbb{R}) \to C(\mathbb{R})$ satisfy the chain rule equation* (4.1). *Assume that $T|_{C_b^k(\mathbb{R})} \not\equiv 0$. Then there is a function $F : \mathbb{R}^{k+2} \to \mathbb{R}$ such that for all $f \in C^k(\mathbb{R})$ and $x \in \mathbb{R}$*

$$Tf(x) = F\big(x, f(x), \ldots, f^{(k)}(x)\big). \tag{4.4}$$

In the case $k = \infty$, this is supposed to mean that $Tf(x)$ depends on x and on all derivative values $f^{(j)}(x)$.

Proof. The result follows immediately from Proposition 3.3 for $I = \mathbb{R}$ and Lemma 4.5(ii). Note that (3.3) is used in the proof of Proposition 3.3 only for half-bounded intervals J. □

Proof of Theorem 4.1. (i) Let $k \in \mathbb{N} \cup \{\infty\}$. We first show that $Tf(x)$ does not depend on any derivative values $f^{(j)}(x)$ of order $j \geq 2$. Let $x_0, y_0, z_0 \in \mathbb{R}$ and

$f, g \in C^k(\mathbb{R})$ satisfy $g(x_0) = y_0$, $f(y_0) = z_0$. Using the representation (4.4) of T, the chain rule equation (4.1) for T turns into a functional equation for F,

$$\begin{aligned} T(f \circ g)(x_0) &= F\big(x_0, z_0, f'(y_0)g'(x_0), (f \circ g)''(x_0), \dots\big) \\ &= (Tf)(y_0)Tg(x_0) \\ &= F\big(y_0, z_0, f'(y_0), f''(y_0), \dots\big)F\big(x_0, y_0, g'(x_0), g''(x_0), \dots\big). \end{aligned} \quad (4.5)$$

If $z_0 = x_0$, also $(g \circ f)(y_0)$ is defined and

$$T(f \circ g)(x_0) = Tf(y_0)Tg(x_0) = Tg(x_0)Tf(y_0) = T(g \circ f)(y_0),$$

i.e.,

$$\begin{aligned} &F\big(x_0, x_0, f'(y_0)g'(x_0), (f \circ g)''(x_0), \dots\big) \\ &\qquad = F\big(y_0, y_0, g'(x_0)f'(y_0), (g \circ f)''(y_0), \dots\big). \end{aligned} \quad (4.6)$$

By the Faà di Bruno formula, cf. Spindler [Sp], the derivatives of $(f \circ g)$ have the form

$$(f \circ g)^{(j)} = f^{(j)} \circ g \cdot (g')^j + \varphi_j(f' \circ g, \dots, f^{(j-1)} \circ g, g', \dots g^{(j-1)}) + f' \circ g \cdot g^{(j)},$$

for $2 \le j \le k$, where φ_j depends only on the lower-order derivatives of f and g, up to order $(j-1)$ (at y_0 and x_0). We have, e.g., $\varphi_2 = 0$, $\varphi_3(f' \circ g, f'' \circ g, g', g'') = 3f'' \circ g \cdot g' \cdot g''$.

Also, for any $x_0, y_0 \in \mathbb{R}$ and any sequence $(t_n)_{n \in \mathbb{N}}$ of real numbers, there is $g \in C^\infty(\mathbb{R})$ with $g(x_0) = y_0$ and $g^{(n)}(x_0) = t_n$ for any $n \in \mathbb{N}$, cf. Hörmander [Ho, p. 16]. This may be shown by adding infinitely many small bump functions. Similarly, given $(s_n)_{n \in \mathbb{N}}$, we may choose $f \in C^\infty(\mathbb{R})$ with $f(y_0) = x_0$ and $f^{(n)}(y_0) = s_n$, $n \in \mathbb{N}$.

Therefore, (4.6) implies, for all $x_0, y_0 \in \mathbb{R}$ and all $(s_n), (t_n)$,

$$\begin{aligned} &F(x_0, x_0, s_1t_1, t_1^2s_2 + s_1t_2, t_1^3s_3 + s_1t_3 + \varphi_{31}, \dots, t_1^js_j + s_1t_j + \varphi_{j1}, \dots) \\ &= F(y_0, y_0, s_1t_1, t_1s_2 + s_1^2t_2, t_1s_3 + s_1^3t_3 + \varphi_{32}, \dots, t_1s_j + s_1^jt_j + \varphi_{j2}, \dots), \end{aligned} \quad (4.7)$$

where $\varphi_{j1}, \varphi_{j2} \in \mathbb{R}$ for $j \ge 3$ depend only on the values of $s_1, \dots, s_{j-1}$ and $t_1, \dots, t_{j-1}$, e.g., $\varphi_{31} = 3s_2t_1t_2$, $\varphi_{32} = 3t_2s_1s_2$. The last dots in (4.7) mean that the variables extend up to $j \le k$ if $k \in \mathbb{N}$, or range over all j if $k = \infty$. Given $z_0 \in \mathbb{R}$, the functions g and f may be chosen with respect to (z_0, y_0) instead of (x_0, y_0) for the same sequences (t_n) and (s_n). Then (4.7) is also true with x_0 being replaced by z_0 which means that $F(x_0, x_0, s_1, \dots, s_j, \dots)$ is independent of x_0. We put

$$K(s_1, \dots, s_j, \dots) := F(x_0, x_0, s_1, \dots, s_j, \dots).$$

Assume that s_1, t_1 are such that $s_1t_1 \notin \{0, 1, -1\}$. We claim that for arbitrary values (a_j) and (b_j)

$$K(s_1t_1, a_2, \dots, a_j, \dots) = K(s_1t_1, b_2, \dots, b_j, \dots),$$

i.e., that K only depends on the first variable s_1t_1 if $s_1t_1 \notin \{0,1,-1\}$. To see this, first note that $\det\begin{pmatrix} t_1^j & s_1 \\ t_1 & s_1^j \end{pmatrix} = (s_1t_1)((s_1t_1)^{j-1}-1) \neq 0$ for $j \geq 2$. Hence, we may solve successively and uniquely the sequence of (2×2)-linear equations for $(s_2,t_2),(s_3,t_3),\ldots,(s_j,t_j)$

$$\begin{aligned} t_1^2 s_2 + s_1 t_2 &= a_2, & t_1 s_2 + s_1^2 t_2 &= b_2, \\ t_1^3 s_3 + s_1 t_3 &= a_3 - \varphi_{31}, & t_1 s_3 + s_1^3 t_3 &= b_3 - \varphi_{32}, \\ &\vdots & &\vdots \\ t_1^j s_j + s_1 t_j &= a_j - \varphi_{j1}, & t_1 s_j + s_1^j t_j &= b_j - \varphi_{j2}, \end{aligned}$$

Here the values obtained for (s_2,t_2) are used to determine φ_{31} and φ_{32} according to the Faà di Bruno formula, and the values up to (s_{j-1},t_{j-1}) to determine φ_{j1} and φ_{j2} accordingly. We then conclude from (4.7)

$$K(s_1t_1,a_2,\ldots,a_j,\ldots) = K(s_1t_1,b_2,\ldots,b_j,\ldots).$$

This means that $K(u_1,u_2,\ldots,u_j,\ldots)$ is independent of the variables $u_2,\ldots,u_j,\ldots$, if $u_1 \notin \{0,1,-1\}$. We then put $\widetilde{K}(u_1) := K(u_1,u_2,\ldots,u_j,\ldots)$. If $u_1 = 1$ choose $t_1 = 2$, $s_1 = 1/2$, $u_1 = s_1t_1 = 1$. Then by (4.5) and (4.6), we find that for any $s_2,\ldots,s_j,\ldots,t_2,\ldots,t_j,\ldots$ we have

$$K\left(1, 4s_2 + \tfrac{1}{2}t_2, \ldots, 2^j s_j + \tfrac{1}{2^j}t_j + \varphi_j, \ldots\right) = \widetilde{K}(2)\widetilde{K}\left(\tfrac{1}{2}\right).$$

Given arbitrary real numbers $u_2,\ldots,u_j,\ldots$, we find successively $s_2,t_2,s_3,t_3,\ldots$ such that the left-hand side equals $K(1,u_2,u_3,\ldots,u_j,\ldots)$ and hence $\widetilde{K}(1) = K(1,u_2,\ldots,u_j,\ldots)$ is also independent of u_j for $j \geq 2$. A similar statement is true for $u_1 = -1$. To show that $K(0,u_2,\ldots,u_j,\ldots)$ is independent of the u_j for $j \geq 2$, too, choose $t_1 = a$, $s_1 = 0$ in (4.7) to find

$$K(0,a^2s_2,\ldots,a^js_j+\varphi_{j1},\ldots) = K(0,as_2,\ldots,as_j+\varphi_{j2},\ldots),$$

for all $a \in \mathbb{R}$, which again implies independence of further variables. We now write $K(u_1)$ for $\widetilde{K}(u_1)$. For values $y_0 \neq x_0 = z_0$, we then know by (4.5) that

$$F(x_0,y_0,t_1,t_2,\ldots,t_j,\ldots) = \frac{K(s_1t_1)}{F(y_0,x_0,s_1,s_2,\ldots,s_j,\ldots)}.$$

Since the left-hand side is independent of $s_1,s_2,\ldots,s_j,\ldots$ and the right-hand side is independent of $t_2,\ldots,t_j,\ldots$, this equation has the form

$$F(x_0,y_0,t_1) = \frac{K(t_1)}{F(y_0,x_0,1)}. \tag{4.8}$$

Note that $F(y_0,x_0,1) \neq 0$ since, using Lemma 4.5(i),

$$F(y_0,x_0,1)F(x_0,y_0,1) = K(1) = T(\mathrm{Id})(x_0) = 1.$$

Define $G : \mathbb{R}^2 \to \mathbb{R}_{\neq 0}$ by $G(x_0, y_0) = 1/F(y_0, x_0, 1)$. Then by (4.8)

$$F(x_0, y_0, t_1) = G(x_0, y_0)K(t_1),$$

with $G(x_0, x_0) = 1$. Using the independence of the derivatives of order ≥ 2, (4.5) implies, for all $x_0, y_0, z_0 \in \mathbb{R}$, that

$$\begin{aligned} F(x_0, z_0, 1) &= F(y_0, z_0, 1)F(x_0, y_0, 1), \\ G(x_0, z_0) &= G(y_0, z_0)G(x_0, y_0). \end{aligned}$$

Define $H : \mathbb{R} \to \mathbb{R}_{\neq 0}$ by $H(y) := G(0, y)$. Then

$$G(x, y) = G(x, 0)G(0, y) = G(0, y)/G(0, x) = H(y)/H(x).$$

Again using (4.8), we get

$$F(x_0, y_0, t_1) = \frac{H(y_0)}{H(x_0)} K(t_1), \tag{4.9}$$

and T has the form

$$Tf(x_0) = F\big(x_0, f(x_0), f'(x_0)\big) = \frac{H \circ f(x_0)}{H(x_0)} K(f'(x_0)), \quad f \in C^k(\mathbb{R}). \tag{4.10}$$

(ii) To identify the form of K, note that by (4.5) for $x_0 = y_0 = z_0$,

$$K(s_1 t_1) = K(s_1)K(t_1), \quad s_1, t_1 \in \mathbb{R},$$

i.e., K is multiplicative on $\mathbb{R}$. Let $b \neq 0$. Apply (4.10) to $f(x) = bx$, we get that $Tf(x) = \frac{H(bx)}{H(x)} K(b)$. Note that $K(b) \neq 0$ since otherwise, by multiplicativity, $K \equiv 0$. Since $Tf \in C(\mathbb{R})$, also $\frac{H(bx)}{H(x)}$ defines a continuous function in x which is strictly positive since H is never zero. We may assume that H is positive. Then for any $b \neq 0$, $\varphi(x) := \ln H(x) - \ln H(bx)$ defines a continuous function $\varphi \in C(\mathbb{R})$. By Proposition 2.8(a), $\ln H$ is measurable and hence also H is measurable. Choosing $f(x) = \frac{1}{2}x^2$ in (4.10), we conclude that

$$K(x) = Tf(x) \frac{H(x)}{H\big(\frac{1}{2}x^2\big)}.$$

Since Tf is continuous and H is measurable, also K is measurable. By Proposition 2.3, the multiplicative function K has the form $K(x) = |x|^p$ or $K(x) = |x|^p \operatorname{sgn} x$ for a suitable $p \in \mathbb{R}$, $x \neq 0$. Hence we conclude from (4.10) and the continuity of Tf that $\frac{H \circ f}{H}$ is continuous for any $f \in C^k(\mathbb{R})$ at any point $x \in \mathbb{R}$ such that $f'(x) \neq 0$.

(iii) We now show that H is continuous. For any $c \in \mathbb{R}$, let

$$b(c) := \overline{\lim_{y \to c}}\, H(y), \quad a(c) := \underline{\lim}_{x \to c}\, H(x).$$

We claim that $\frac{b(c)}{H(c)}$ and $\frac{a(c)}{H(c)}$ are constant functions of c. In the case that for some c_0, $b(c_0)$ or $a(c_0)$ are infinite or zero, this should mean that all other values $b(c)$ or $a(c)$ are also infinite or zero. Assume to the contrary that there are c_0 and c_1 such that $\frac{b(c_1)}{H(c_1)} < \frac{b(c_0)}{H(c_0)}$. Choose any maximizing sequence y_n, $\lim_{n\to\infty} y_n = c_0$ with $\lim_{n\to\infty} H(y_n) = b(c_0)$. Since for $f(t) = t + c_1 - c_0$, $\frac{H\circ f}{H}$ is continuous by part (ii), $\lim_{n\to\infty} \frac{H(y_n+c_1-c_0)}{H(y_n)} = \frac{H(c_1)}{H(c_0)}$ exists and using $\overline{\lim}_{n\to\infty} H(y_n + c_1 - c_0) \le b(c_1)$, we arrive at the contradiction

$$\frac{b(c_0)}{H(c_0)} = \frac{H(c_1)}{H(c_0)}\frac{b(c_0)}{H(c_1)} = \lim_{n\to\infty} \frac{H(y_n + c_1 - c_0)}{H(y_n)}\frac{H(y_n)}{H(c_1)}$$
$$\le \overline{\lim_{n\to\infty}} \frac{H(y_n + c_1 - c_0)}{H(c_1)} \le \frac{b(c_1)}{H(c_1)} < \frac{b(c_0)}{H(c_0)}.$$

The argument is also valid assuming $b(c_1) < b(c_0) = \infty$. The proof for $a(c)$ is similar.

If H would be discontinuous at some point, it would be discontinuous anywhere since the functions $\frac{a}{H}$ and $\frac{b}{H}$ and hence $\frac{b}{a}$ are constant, under this assumption with $\frac{b}{a} > 1$. Assume that this is the case, and choose a sequence $(c_n)_{n\in\mathbb{N}}$ of pairwise disjoint numbers with $\lim_{n\to\infty} c_n = 0$. Let $\delta_n := \frac{1}{4}\min\{|c_n - c_m| \mid n \ne m\}$ and choose $0 < \epsilon_n < \delta_n$ such that $\sum_{n\in\mathbb{N}}(\epsilon_n/\delta_n)^k < \infty$ for all $k \in \mathbb{N}$, i.e., $(\epsilon_n)_{n\in\mathbb{N}}$ should decay much faster to zero than δ_n. Since H is discontinuous at any c_n, $\frac{b(c_n)}{a(c_n)} > 1$. By the above argument, this is independent of $n \in \mathbb{N}$, $1 < \frac{b}{a} := \frac{b(c_n)}{a(c_n)}$. By definition of $b(c_n)$ and $a(c_n)$, we may find $y_n, x_n \in \mathbb{R}$ with

$$|y_n - c_n| < \epsilon_n, \quad |x_n - c_n| < \epsilon_n, \quad \frac{H(y_n)}{H(x_n)} > \frac{b+a}{2a} > 1.$$

If $\frac{b}{a} = \infty$, choose them with $\frac{H(y_n)}{H(x_n)} > 2$. Let ψ be a C^∞-cutoff function like $\psi(x) = \exp\left(-\frac{x^2}{1-x^2}\right)$ for $|x| < 1$, and $\psi(x) = 0$ for $|x| \ge 1$, and put $g_n(x) = (y_n - x_n)\psi\left(\frac{x-x_n}{\delta_n}\right)$. The functions $(g_n)_{n\in\mathbb{N}}$ have disjoint support since for any $m \ne n$

$$|x_n - x_m| \ge |c_n - c_m| - 2\epsilon_n \ge 4\delta_n - 2\epsilon_n \ge 2\delta_n.$$

Hence $g_n(x_m) = (y_n - x_n)\delta_{nm}$. Since

$$\sum_{n\in\mathbb{N}} \|g_n^{(k)}\|_\infty \le \sum_{n\in\mathbb{N}} \left(\frac{|y_n - x_n|}{\delta_n}\right)^k \|\psi^{(k)}\|_\infty \le \sum_{n\in\mathbb{N}} \left(\frac{2\epsilon_n}{\delta_n}\right)^k \|\psi^{(k)}\|_\infty < \infty$$

holds for any $k \in \mathbb{N}$,

$$f(x) := x + \sum_{n\in\mathbb{N}} g_n(x), \quad x \in \mathbb{R}$$

defines a C^∞-function f with $f(x_n) = y_n$, $f(0) = 0$ and $f'(0) = 1 \neq 0$. Since $x_n \to 0$, $y_n = f(x_n) \to 0$, the continuity of $\frac{H\circ f}{H}$ yields the contradiction

$$1 = \frac{H(0)}{H(0)} = \lim_{n\to\infty} \frac{H(y_n)}{H(x_n)} > \frac{b+a}{2a} > 1.$$

This proves that H is continuous.

Now (4.10) implies that

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

for any $f \in C^k(\mathbb{R})$, $x \in \mathbb{R}$. By assumption $Tf \in C(\mathbb{R})$ is continuous for any $f \in C^k(\mathbb{R})$. This requires $p \geq 0$, choosing functions f whose derivatives have zeros. In fact, if the term $\operatorname{sgn} f'(x)$ is present, $p > 0$ is needed to guarantee the continuity of all functions in the image of T.

(iv) If $T(2\,\mathrm{Id}) = 2$ is the constant function 2, then $\frac{H(2x)}{H(x)} 2^p = 2$ for all x, which for $x = 0$ yields $p = 1$. For $b = 1/2$, the function φ in part (ii) is constant,

$$\varphi(x) = \ln H(x) - \ln H(x/2) = 0.$$

Hence, the argument in the proof of Proposition 2.8(a) shows that $\ln H(x) = \ln H(1)$, $H(x) = H(1)$, taking $L = \ln H$ in Proposition 2.8(a). Hence, $\frac{H\circ f}{H} = 1$ and $Tf = f'$ or $Tf = |f'|$. If $T(-2\,\mathrm{Id}) = -2$, the only possible solution of Theorem 4.1 is $Tf = f'$.

Clearly, the operators T given by formulas (4.2) and (4.3) satisfy the chain rule (4.1). This proves Theorem 4.1. □

If the image of T consists of smooth functions, we have further restrictions on H and p:

Proposition 4.7. *Let $k \in \mathbb{N}$, $k \geq 2$ and suppose that $T : C^k(\mathbb{R}) \to C^{k-1}(\mathbb{R})$ satisfies the chain rule* (4.1) *with $T|_{C_b^k(\mathbb{R})} \not\equiv 0$. Then there exists $H \in C^{k-1}(\mathbb{R})$, $H > 0$ and p with either*

$$p > k-1 \quad \text{and} \quad Tf = \frac{H\circ f}{H} |f'|^p \{\operatorname{sgn} f'\}$$

or

$$p \in \{0, \dots, k-1\} \quad \text{and} \quad Tf = \frac{H\circ f}{H} (f')^p, \quad f \in C^k(\mathbb{R}).$$

If the chain rule holds for $T : C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ with $T|_{C_b^\infty(\mathbb{R})} \not\equiv 0$, there is $H \in C^\infty(\mathbb{R})$ and $p \in \mathbb{N}_0$ such that

$$Tf = \frac{H\circ f}{H} (f')^p, \quad f \in C^\infty(\mathbb{R}).$$

Proof. By Theorem 4.1, T is of the above form with $H \in C(\mathbb{R})$ and $p \geq 0$. Suppose T maps $C^k(\mathbb{R})$ into $C^{k-1}(\mathbb{R})$. Then the condition on p is needed to guarantee that Tf is in $C^{k-1}(\mathbb{R})$ for functions f whose derivatives have zeros.

We claim that H is smooth, i.e., $H \in C^{k-1}(\mathbb{R})$. Let $L := -\log H$. Obviously $L \in C^{k-1}(\mathbb{R})$ if and only if $H \in C^{k-1}(\mathbb{R})$. Take $f(x) = x/2$. By assumption $Tf \in C^{k-1}(\mathbb{R})$ and, hence,

$$\varphi(x) := L(x) - L(x/2)$$

defines a function $\varphi \in C^{k-1}(\mathbb{R})$. We prove by induction on $k \geq 2$ that $\varphi \in C^{k-1}(\mathbb{R})$ and $L \in C^{k-2}(\mathbb{R})$ imply that $L \in C^{k-1}(\mathbb{R})$.

For $k = 2$, $\varphi \in C^1(\mathbb{R})$ and $L \in C(\mathbb{R})$ since $H \in C(\mathbb{R})$. By Proposition 2.8(b) with $\psi = \varphi$ and $a = 1$, we get $L \in C^1(\mathbb{R})$.

To prove the induction step, assume $k \geq 3$, $\varphi \in C^{k-1}(\mathbb{R})$ and $L^{(k-2)} \in C(\mathbb{R})$. We have to show that $L \in C^{k-1}(\mathbb{R})$. Let $\psi(x) := \varphi^{(k-2)}(x) = L^{(k-2)}(x) - \frac{1}{2^{k-2}} L^{(k-2)}\big(\frac{x}{2}\big)$. Then $\psi \in C^1(\mathbb{R})$ and $L^{(k-2)} \in C(\mathbb{R})$. By Proposition 2.8(b) with $a = \frac{1}{2^{k-2}}$, $L^{(k-2)} \in C^1(\mathbb{R})$, i.e., $L \in C^{k-1}(\mathbb{R})$. This proves that $H \in C^{k-1}(\mathbb{R})$. □

4.2 The chain rule on different domains

In the case of C^1-functions, there is an analogue of Theorem 4.1 for functions $f : \mathbb{R}^n \to \mathbb{R}^n$ on $\mathbb{R}^n$ when $n > 1$. For finite-dimensional Banach spaces X and Y and $k \in \mathbb{N}_0$, let

$$C^k(X, Y) = \{f : X \to Y \mid f \text{ is } k\text{-times continuously Fréchet differentiable}\},$$

with $C(X, Y) = C^0(X, Y)$. Let $L(X, Y) := \{f \in C(X, Y) \mid f \text{ is linear}\}$ and $C_b^k(X, \mathbb{R}^n) := \{f \in C^k(X, \mathbb{R}^n) \mid \mathrm{Im}(f) \subset J \text{ for some open half-space } J \subset \mathbb{R}^n\}$. The derivative D is a map $D : C^1(\mathbb{R}^n, \mathbb{R}^n) \to C(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^n))$ satisfying the chain rule

$$D(f \circ g)(x) = \big((Df) \circ g\big)(x) \cdot (Dg)(x), \quad f, g \in C^1(\mathbb{R}^n, \mathbb{R}^n),\ x \in \mathbb{R}^n.$$

More generally, we consider operators $T : C^1(\mathbb{R}^n, \mathbb{R}^n) \to C(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^n))$ satisfying the *chain rule equation*

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

The multiplication on the right is the non-commutative composition of linear operators on $\mathbb{R}^n$. We do not write it with composition symbol $\circ$ to distinguish it from the composition of the non-linear functions f, g. In fact, in the following we will omit the symbol $\cdot$ for this composition. In stating the analogue of Theorem 4.1 for $n > 1$, we need another assumption on T.

An operator $T : C^1(\mathbb{R}^n, \mathbb{R}^n) \to C(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^n))$ is *locally surjective* provided that there is $x \in \mathbb{R}^n$ so that

$$\big\{(Tf)(x) \;\big|\; f \in C^1(\mathbb{R}^n, \mathbb{R}^n),\ f(x) = x,\ \det f'(x) \neq 0\big\} \supseteq \mathrm{GL}(n, \mathbb{R}).$$

In the following result on the chain rule for maps of this type we use the notation $\det T|_{C_b^1}(\mathbb{R}^n, \mathbb{R}^n) \not\equiv 0$ to mean that there should be a function $f \in C_b^1(\mathbb{R}^n, \mathbb{R}^n)$ and a point $x \in \mathbb{R}^n$ such that $\det(Tf(x)) \neq 0$.

Theorem 4.8 (Multidimensional chain rule)**.** *Let $n \geq 2$, and assume that $T : C^1(\mathbb{R}^n, \mathbb{R}^n) \to C(\mathbb{R}^n, L(\mathbb{R}^n, \mathbb{R}^n))$ satisfies the chain rule equation*

$$T(f \circ g)(x) = \big((Tf) \circ g\big)(x)\ Tg(x), \quad f, g \in C^1(\mathbb{R}^n, \mathbb{R}^n),\ x \in \mathbb{R}^n. \tag{4.11}$$

Assume also that $\det T|_{C_b^1(\mathbb{R}^n, \mathbb{R}^n)} \not\equiv 0$ and that T is locally surjective. Then there are $p \geq 0$ and $H \in C(\mathbb{R}^n, \mathrm{GL}(n, \mathbb{R}))$ such that, if $n \in \mathbb{N}$ is odd, for all $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ and $x \in \mathbb{R}^n$

$$(Tf)(x) = \big|\det f'(x)\big|^p (H \circ f)(x) f'(x) H(x)^{-1}.$$

If $n \in \mathbb{N}$ is even, T either has the same form or

$$Tf(x) = \mathrm{sgn}\big(\det f'(x)\big)\big|\det f'(x)\big|^p (H \circ f)(x) f'(x) H(x)^{-1},$$

the latter with $p > 0$.

Conversely, these formulas define operators T which satisfy the chain rule and are locally surjective.

If additionally to (4.11), $T(2\,\mathrm{Id})(x) = 2\,\mathrm{Id}$ *holds for all $x \in \mathbb{R}^n$, then $H = \mathrm{Id}$ and $Tf = f'$ or, if n is even, possibly $Tf = \mathrm{sgn}(\det f')f'$.*

Remarks. (a) Note that a priori we do not impose any continuity condition on T.

(b) For odd integers $n \in \mathbb{N}$, $p > 0$ and $H \in C(\mathbb{R}^n, \mathrm{GL}(n, \mathbb{R}))$,

$$(Tf)(x) := \mathrm{sgn}\big(\det f'(x)\big)\big|\det f'(x)\big|^p (H \circ f)(x) f'(x) H(x)^{-1}$$

also solves the chain rule equation, but is not locally surjective since in this case $\det((Tf)(x)) \geq 0$ for all $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ with $f(x) = x$.

(c) If T is not assumed to be locally surjective, there are various other solutions of (4.11):

Take any continuous multiplicative homomorphism $\Phi : \mathbb{R} \to L(\mathbb{R}^n, \mathbb{R}^n)$ with $\Phi(0) = 0$ and $\Phi(1) = \mathrm{Id}$ and any continuous function $H \in C(\mathbb{R}^n, \mathrm{GL}(n, \mathbb{R}))$, and define

$$(Tf)(x) = (H \circ f)(x)\Phi\big(\det f'(x)\big) H(x)^{-1},$$

$x \in \mathbb{R}^n$, $f \in C^1(\mathbb{R}^n, \mathbb{R}^n)$. Then T satisfies (4.11). As for specific examples, take as Φ a one-parameter group like $\Phi(t) = \exp(\ln|t| A) = |t|^A$ for some fixed matrix $A \in L(\mathbb{R}^n, \mathbb{R}^n)$ and $t \in \mathbb{R}$. Here $\ln|t|$ might also be replaced by $(\mathrm{sgn}\, t) \cdot \ln|t|$.

(d) As in the case of one variable ($n = 1$), the function H is completely determined by the function $T(2\,\mathrm{Id})$. The inner automorphism defined by H, with additional composition by f, applied to the derivative, essentially yields T up to a character in terms of $\det f'$.

For the proof of Theorem 4.8 we refer to [KM2]. We will not reproduce it here since it is not in line with our main goals. We just mention a few steps of the proof.

The localization step for $n \geq 2$ is similar to the case $n = 1$, yielding

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

for a suitable function $F : \mathbb{R}^n \times \mathbb{R}^n \times L(\mathbb{R}^n, \mathbb{R}^n) \to L(\mathbb{R}^n, \mathbb{R}^n)$. The analysis of this representing function F is different from the case $n = 1$, due to the non-commutativity of the composition of linear maps in $L(\mathbb{R}^n, \mathbb{R}^n)$. However, again one may show that

$$K(v) := F(x, x, v) \in \mathrm{GL}(n, \mathbb{R}), \quad v \in \mathrm{GL}(n, \mathbb{R})$$

is independent of $x \in \mathbb{R}^n$ and multiplicative, $K(uv) = K(u)K(v)$ for all $u, v \in \mathrm{GL}(n, \mathbb{R})$, with $K(\mathrm{Id}) = \mathrm{Id}$, $K(v)^{-1} = K(v^{-1})$. The proof proceeds identifying these automorphisms K of $\mathrm{GL}(n, \mathbb{R})$ as inner automorphisms multiplied by characters in terms of $\det v$, i.e., powers of $|\det v|$, possibly multiplied by $\mathrm{sgn}(\det v)$. This result on the automorphisms of $\mathrm{GL}(n, \mathbb{R})$ replaces (the simpler) Proposition 2.3. Additional arguments are also needed to prove the continuity of H.

We may also consider the chain rule equation on real or complex spaces of polynomials or analytic functions. For $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, let $\mathcal{P} := \mathcal{P}(\mathbb{K})$ denote the space of polynomials with coefficients in $\mathbb{K}$, $\mathcal{E} := \mathcal{E}(\mathbb{K})$ the space of real-analytic functions ($\mathbb{K} = \mathbb{R}$) or entire functions ($\mathbb{K} = \mathbb{C}$) and $C := C(\mathbb{K})$. Moreover, let $\mathcal{P}_n := \mathcal{P}_n(\mathbb{K})$ be the subset of $\mathcal{P}$ consisting of polynomials of degree $\leq n$. There are simple operators $T : \mathcal{P}(\mathbb{K}) \to C(\mathbb{K})$ satisfying the chain rule $T(f \circ g) = (Tf) \circ g \cdot Tg$ which have a different form than the solutions determined so far: For $f \in \mathcal{P}(\mathbb{K})$ and $c \in \mathbb{R}$, let $Tf := (\deg f)^c$, T mapping into the constant functions. Then T satisfies the chain rule on $\mathcal{P}$. More generally, if $\deg f = \prod_{j=1}^r p_j^{l_j}$ is the decomposition of $\deg f$ into prime powers and $c_j \in \mathbb{R}$, $Tf = \prod_{j=1}^r p_j^{c_j l_j}$ will satisfy the chain rule and also

$$Tf = \prod_{j=1}^{r} p_j^{c_j l_j} \frac{H \circ f}{H} |f'|^p \{\mathrm{sgn}\, f'\}^m$$

will define a map $T : \mathcal{P} \to C$ satisfying the chain rule, if $H \in C(\mathbb{K})$, $H \neq 0$, $p \geq 0$, $m \in \mathbb{N}_0$. We do not know whether this yields the general solution of the chain rule equation for $T : \mathcal{P} \to C$. However, we can give the general solution of the chain rule equation for such maps under a mild continuity assumption.

Let $X \in \{\mathcal{P}(\mathbb{K}), \mathcal{E}(\mathbb{K})\}$ and $Y \in \{\mathcal{P}(\mathbb{K}), \mathcal{E}(\mathbb{K}), C(\mathbb{K})\}$. An operator $T : X \to Y$ is *pointwise continuous at* 0 provided that for any sequence $(f_n)_{n \in \mathbb{N}}$ of functions in X converging uniformly on all compact sets of $\mathbb{K}$ to a function $f \in X$, we have pointwise convergence of the images at zero, i.e., $\lim_{n\to\infty}(Tf_n)(0) = (Tf)(0)$. For $\xi \in \mathbb{K} \setminus \{0\}$, denote $\operatorname{sgn}\xi := \xi/|\xi|$. We then have the following two results for the chain rule.

Theorem 4.9. *Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ and suppose that $T : \mathcal{P}(\mathbb{K}) \to C(\mathbb{K})$, $T \neq 0$, satisfies the chain rule equation*

$$T(f \circ g) = (Tf) \circ g \cdot Tg, \quad f, g \in \mathcal{P}(\mathbb{K}) \tag{4.12}$$

and is pointwise continuous at 0. Then there is a nowhere vanishing continuous function $H \in C(\mathbb{K})$ and there are $p \in \mathbb{K}$ with $\operatorname{Re}(p) \geq 0$ and $m \in \mathbb{Z}$ such that

$$Tf = \frac{H \circ f}{H} |f'|^p (\operatorname{sgn} f')^m. \tag{4.13}$$

For $\mathbb{K} = \mathbb{R}$, $m \in \{0, 1\}$ suffices and $H > 0$. For $p = 0$, only $m = 0$ yields a solution with range in $C(\mathbb{K})$. If T maps into the space $\mathcal{P}(\mathbb{K})$, H is constant and $p = m \in \mathbb{N}_0$ so that T has the form $Tf = f'^m$.

The result for entire functions is

Theorem 4.10. *Let $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ and assume that $T : \mathcal{E}(\mathbb{K}) \to \mathcal{E}(\mathbb{K})$, $T \neq 0$, satisfies the chain rule equation*

$$T(f \circ g) = (Tf) \circ g \cdot Tg, \quad f, g \in \mathcal{E}(\mathbb{K})$$

and is pointwise continuous at 0. Then there is a function $h \in \mathcal{E}(\mathbb{K})$ and there is $m \in \mathbb{N}_0$ such that

$$Tf = \exp(h \circ f - f) \cdot f'^m.$$

Proof of Theorem 4.9. (a) Since $T \neq 0$, there are $n_0 \in \mathbb{N}$, $g \in \mathcal{P}_{n_0}(\mathbb{K})$ and $x_1 \in \mathbb{K}$ such that $Tg(x_1) \neq 0$. Let $n \in \mathbb{N}$, $n \geq n_0$. We restrict T to $\mathcal{P}_n(\mathbb{K}) =: \mathcal{P}_n$ and apply (4.12) for $f, g \in \mathcal{P}_n$ with $f \circ g \in \mathcal{P}_n$. For any $x_0 \in \mathbb{K}$, consider the shift $S(x) := x + x_1 - x_0$, $S \in \mathcal{P}_1 \subset \mathcal{P}_n$ and put $f := g \circ S$. Then by (4.12)

$$0 \neq (Tg)(x_1) = T(f \circ S^{-1})(x_1) = (Tf)(x_0) T(S^{-1})(x_1).$$

Hence, $Tf(x_0) \neq 0$. Moreover, $Th = T(h \circ \mathrm{Id}) = Th \cdot T(\mathrm{Id})$ for all $h \in \mathcal{P}_n$. Hence, $T(\mathrm{Id}) = 1$ is the constant function 1. For $x \in \mathbb{K}$, let $S_x(y) := x + y$, $S_x \in \mathcal{P}_1 \subset \mathcal{P}_n$. Again by (4.12)

$$1 = T(\mathrm{Id}) = T(S_{-x} \circ S_x) = T(S_{-x}) \circ S_x \cdot T(S_x).$$

Thus for all $y \in \mathbb{K}$, $T(S_x)(y) \neq 0$. In particular, $T(S_x)(0) \neq 0$ for all $x \in \mathbb{K}$. Again by (4.12)

$$T(f \circ S_x)(0) = (Tf)(x) \cdot T(S_x)(0),$$

so that for any $x \in \mathbb{K}$ and $f \in \mathcal{P}_n$

$$Tf(x) = \frac{T(f \circ S_x)(0)}{T(S_x)(0)}. \tag{4.14}$$

Let $f_j \in \mathcal{P}_n$ be a sequence converging uniformly on compacta to $f \in \mathcal{P}_n$. Then (4.14) and the pointwise continuity assumption at 0 imply that $\lim_{j\to\infty}(Tf_j)(x) = (Tf)(x)$ for all $x \in \mathbb{K}$, and not only for $x = 0$. By (4.14) it suffices to determine the form of $(Tf)(0)$ for any $f \in \mathcal{P}_n$. Since, for any $f \in \mathcal{P}_n$, $f(x) = \sum_{j=0}^{n} \frac{f^{(j)}(0)}{j!} x^j$ is determined by the sequence $(f^{(j)}(0))_{0\le j\le n}$, $(Tf)(0)$ is a function of these values. Hence, there is $F_n : \mathbb{K}^{n+1} \to \mathbb{K}$ such that

$$(Tf)(0) = F_n\big(f(0), f'(0), \dots, f^{(n)}(0)\big), \quad f \in \mathcal{P}_n. \tag{4.15}$$

Since $(f \circ S_x)^{(j)} = f^{(j)} \circ S_x$, (4.14) and (4.15) imply

$$Tf(x) = \frac{F_n\big(f(x), f'(x), \dots, f^{(n)}(x)\big)}{F_n(x, 1, 0, \dots 0)}, \tag{4.16}$$

with $F_n(x, 1, 0, \dots, 0) = T(S_x)(0) \neq 0$ for any $x \in \mathbb{K}$.

(b) We now show that Tf does not depend on the higher derivatives $f^{(j)}$ for $j \ge 2$. Fix $x \in \mathbb{K}$ and define $G_n = G_{n,x} : \mathbb{K}^n \to \mathbb{K}$ by

$$G_n(\xi_1, \dots, \xi_n) := \frac{F_n(x, \xi_1, \dots, \xi_n)}{F_n(x, 1, 0, \dots, 0)}, \quad \xi_i \in \mathbb{K}. \tag{4.17}$$

For any $(\eta_1, \dots, \eta_n) \in \mathbb{K}^n$, there is a polynomial $g \in \mathcal{P}_n$, with $g(x) = x$ and $g^{(j)}(x) = \eta_j$ for $j = 1, \dots, n$. For $\xi_1 \in \mathbb{K}$, define $f \in \mathcal{P}_1 \subset \mathcal{P}_n$ by $f(y) := \xi_1(y - x) + x$. Then $f(x) = x$, $(f \circ g)^{(j)}(x) = \xi_1 \eta_j$ and $(g \circ f)^{(j)}(x) = \xi_1^j \eta_j$. Therefore, by (4.16) and (4.17)

$$\begin{aligned} G_n(\xi_1\eta_1, \dots, \xi_1\eta_n) &= G_n\big((f \circ g)'(x), \dots, (f \circ g)^{(n)}(x)\big) \\ &= T(f \circ g)(x) = (Tf)(x)(Tg)(x) = (Tg)(x)(Tf)(x) \\ &= T(g \circ f)(x) = G_n(\xi_1\eta_1, \dots, \xi_1^n\eta_n). \end{aligned}$$

Given $(t_1, \dots, t_n) \in \mathbb{K}^n$ and $\alpha \in \mathbb{K}$, $\alpha \neq 0$, let $\eta_i = t_i/\alpha$.

Applying the previous equations with $\xi_1 = \alpha$, we conclude

$$G_n(t_1, t_2, \dots, t_n) = G_n(t_1, \alpha t_2, \dots, \alpha^{n-1} t_n). \tag{4.18}$$

Fix $t_1 \in \mathbb{K}$ and define $\widetilde{G}_n : \mathbb{K}^{n-1} \to \mathbb{K}$ by $\widetilde{G}_n(t_2, \dots, t_n) := G_n(t_1, t_2, \dots t_n)$. Then $\widetilde{G}_n$ is continuous at zero: if $t^{(m)} = (t_2^{(m)}, \dots, t_n^{(m)}) \to 0 \in \mathbb{K}^{n-1}$ for $m \to \infty$, choose polynomials $f_m \in \mathcal{P}_n$ with $f_m(x) = x$, $f_m'(x) = t_1$ and $f_m^{(j)}(x) = t_j^{(m)}$ for $2 \le j \le n$. Clearly, f_m converges uniformly on compact sets to f, where

$f(x) = t_1(y-x)+x$. By the assumption of pointwise continuity at 0 of T, (4.16) and (4.17),

$$\widetilde{G}_n(t_2^{(m)},\dots,t_n^{(m)}) = G_n(t_1,t_2^{(m)},\dots,t_n^{(m)}) = (Tf_m)(x) \\ \longrightarrow (Tf)(x) = G_n(t_1,0,\dots,0) = \widetilde{G}_n(0,\dots,0).$$

Hence, $\widetilde{G}_n$ is continuous at 0. Letting $\alpha \to 0$ in (4.18), we find

$$G_n(t_1,t_2,\dots,t_n) = \lim_{\alpha\to 0} G_n(t_1,\alpha t_2,\dots,\alpha^{n-1}t_n) = G_n(t_1,0,\dots,0), \tag{4.19}$$

i.e., $G_n = G_{n,x}$ does not depend on the variables $(t_2,\dots,t_n) \in \mathbb{K}^{n-1}$: Therefore Tf is independent of the higher derivatives of f.

(c) For any $f \in \mathcal{P}_n$ with $f(x) = x$ and $f'(x) = \xi_1$, we now know by (4.16), (4.17) and (4.19) that

$$\begin{aligned}(Tf)(x) &= G_n\big(f'(x),\dots f^{(n)}(x)\big) = G_n(\xi_1,0,\dots,0)\\ &= \frac{F_n(x,\xi_1,0,\dots,0)}{F_n(x,1,0,\dots,0)} =: \phi(x,\xi_1).\end{aligned} \tag{4.20}$$

If $g \in \mathcal{P}_1$ satisfies $g(x) = x$, $g'(x) = \eta_1$, we have by (4.12) and (4.20)

$$\phi(x,\xi_1\eta_1) = T(f\circ g)(x) = (Tf)(x)(Tg)(x) = \phi(x,\eta_1)\phi(x,\eta_1).$$

Therefore, $\phi(x,\cdot) : \mathbb{K} \to \mathbb{K}$ is multiplicative for every fixed $x \in \mathbb{K}$. It is also continuous: for $\xi_1^{(m)} \to \xi_1$ in $\mathbb{K}$, put $f_m(y) = \xi_1^{(m)}(y-x)+x$, $f(y) := \xi_1(y-x)+x$. Then $f_m \to f$ converges uniformly on compacta and hence

$$\phi(x,\xi_1^{(m)}) = (Tf_m)(x) \longrightarrow (Tf)(x) = \phi(x,\xi_1).$$

By Proposition 2.3 ($\mathbb{K} = \mathbb{R}$) and Proposition 2.4 ($\mathbb{K} = \mathbb{C}$) there are $p(x) \in \mathbb{K}$ with $\operatorname{Re}(p(x)) \geq 0$ and $m(x) \in \mathbb{Z}$ such that

$$\phi(x,\xi_1) = |\xi_1|^{p(x)}(\operatorname{sgn}\xi_1)^{m(x)}, \tag{4.21}$$

$\operatorname{sgn}\xi_1 = \xi_1/|\xi_1|$ for $\xi \neq 0$ and $\phi(x,0) = 0$, with $m(x) = 0$ if $\operatorname{Re}(p(x)) = 0$ and $m(x) \in \{0,1\}$ if $\mathbb{K} = \mathbb{R}$.

(d) Let $H(x) = T(S_x)(0) = F_n(x,1,0,\dots,0)$. Then $H(x) \neq 0$ and by (4.16), (4.19), (4.20) and (4.21),

$$\begin{aligned}Tf(x) &= \frac{F_x\big(f(x),f'(x),0,\dots,0\big)}{F_n(x,1,0,\dots,0)} = \frac{H(f(x))}{H(x)}\phi\big(f(x),f'(x)\big)\\ &= \frac{(H\circ f)(x)}{H(x)}|f'(x)|^{p(f(x))}\big(\operatorname{sgn} f'(x)\big)^{m(f(x))}.\end{aligned} \tag{4.22}$$

Choosing $f(x) = 2x$, we find that p is a continuous function since Tf and H are continuous. Actually, p is constant: Choosing arbitrary $x, y, z \in \mathbb{K}$ and functions $f, g \in \mathcal{P}_1$ with $g(x) = y$, $f(y) = z$, we have by (4.12) and (4.22),

$$\begin{aligned}\big|f'(y)g'(x)\big|^{p(yz)}&\big(\operatorname{sgn} f'(y)g'(x)\big)^{m(yz)}\\ &= |f'(y)|^{p(z)}\big(\operatorname{sgn} f'(y)\big)^{m(z)}|g'(x)|^{p(y)}\big(\operatorname{sgn} g'(x)\big)^{m(y)}.\end{aligned}$$

Applying this first to polynomials with $f'(y) > 0$, $g'(x) > 0$, we find that $p(yz) = p(z) = p(y) =: p$ for all $y, z \in \mathbb{K}$, i.e., p is constant. Then, using functions with arbitrary sgn-values in S^1, we find that $m(yz) = m(z) = m(y) = m \in \mathbb{Z}$ may be taken constant. With $p = p(f(x))$ and $m = m(f(x))$, (4.22) gives the general solution for $T : \mathcal{P}_n \to C$, both for $\mathbb{K} = \mathbb{R}$ and $\mathbb{K} = \mathbb{C}$.

(e) Since (4.22) is independent of $n \in \mathbb{N}$, this is also the general solution for $T : \mathcal{P} \to C$. In the case that $T : \mathcal{P} \to \mathcal{P}$, i.e., that the range of T consists only of polynomials, all functions

$$Tf = \frac{H \circ f}{H}|f'|^p(\operatorname{sgn} f')^m, \quad f \in \mathcal{P},$$

have to be polynomials. Here $m \in \mathbb{Z}$, $p \in \mathbb{K}$, $\operatorname{Re}(p) \geq 0$. For $f(x) = \frac{1}{2}x^2$ this means that $\frac{H\left(\frac{1}{2}x^2\right)}{H(x)}|x|^p(\operatorname{sgn} x)^m$ is a polynomial. For $p = 0$ also $m = 0$ and $Tf = \frac{H\circ f}{H}$. For $p > 0$, Tf has a zero of order p in $x_0 = 0$. Since Tf is a polynomial, it follows that $p \in \mathbb{N}$ is a positive integer, and $Tf(x) = x^p g(x)$ with $g \in \mathcal{P}$, $g(0) \neq 0$. This implies that $m \in \mathbb{Z}$ has to be such that $x^p = |x|^p(\operatorname{sgn} x)^m$. Therefore $Tf = \frac{H\circ f}{H} f'^p \in \mathcal{P}$ for all $f \in \mathcal{P}$, with $p \in \mathbb{N}_0$. Applying this to linear functions $f(x) = ax + b$, $f^{-1}(y) = \frac{1}{a}y - \frac{b}{a} = x$, we find that $p(x) = \frac{H(ax+b)}{H(x)}$ and $\frac{H(x)}{H(ax+b)} = \frac{1}{p(x)}$ are polynomials in x. Therefore, $\frac{H(ax+b)}{H(x)} =: c_{a,b}$ is constant in $x \in \mathbb{K}$ for any fixed values $a, b \in \mathbb{K}$. In particular

$$\frac{H(2x)}{H(x)} = \frac{H(0)}{H(0)} = 1 =: c_{2,0}, \qquad \frac{H(x+b)}{H(x)} =: c_{1,b}.$$

We find that

$$H(2x + 2b) = H(x + b) = c_{1,b}H(x) = c_{1,b}H(2x) = H(2x + b)$$

for all $x, b \in \mathbb{K}$. Therefore, $H(y + b) = H(y)$ for all $y, b \in \mathbb{K}$. Hence, H is constant and $\frac{H\circ f}{H} = 1$ for all $f \in \mathcal{P}$. We conclude that $Tf = f'^p$, $p \in \mathbb{N}_0$. □

Proof of Theorem 4.10. Since $\mathcal{P}(\mathbb{K}) \subset \mathcal{E}(\mathbb{K})$, Theorem 4.9 yields that $T|_{\mathcal{P}(\mathbb{K})}$ has the form

$$Tf = \frac{H \circ f}{H}|f'|^p(\operatorname{sgn} f')^m, \quad f \in \mathcal{P}(\mathbb{K}), \tag{4.23}$$

with $m \in \mathbb{Z}$, $p \in \mathbb{K}$, $\operatorname{Re}(p) \geq 0$. We also know that H defined by $H(x) = T(S_x)(0)$ is continuous on $\mathbb{K}$. Let $c \in \mathbb{K}$, $c \neq 0$ be arbitrary. Applying (4.23) to $f(z) = cz$

and using that $Tf \in \mathcal{E}(\mathbb{K})$, we get that $z \mapsto \frac{H(cz)}{H(z)}$ is in $\mathcal{E}(\mathbb{K})$, i.e., real or complex analytic. Since H is nowhere zero, there exists an analytic function $k(c,\cdot\,) \in \mathcal{E}(\mathbb{K})$ such that $\frac{H(cz)}{H(z)} = \exp(k(c,z))$, with $k(c,0)=0$. For $c,d \in \mathbb{K}$ we find

$$\exp\bigl(k(cd,z)\bigr) = \frac{H(cdz)}{H(z)} = \frac{H(cdz)}{H(dz)}\frac{H(dz)}{H(z)} = \exp\bigl(k(c,dz)+k(d,z)\bigr),$$

hence $k(cd,z) = k(c,dz)+k(d,z)$. In particular, for $z=1$, $k(c,d)=k(cd,1)-k(d,1)$. Let $h(d) := k(d,1)$ for $d \neq 0$. Then $k(c,d) = h(cd) - h(d)$, and with d replaced by z, $k(c,z) = h(cz) - h(z)$. Since H is continuous, k is continuous as a function of both variables. Therefore,

$$\lim_{c\to 0} k(c,z) = \lim_{c\to 0} h(cz) - h(z) := h(0) - h(z)$$

exists z-uniformly on compact subsets of $\mathbb{K}$. Since $k(c,\cdot\,) \in \mathcal{E}(\mathbb{K})$ for all $c \in \mathbb{K}$, we conclude that $h \in \mathcal{E}(\mathbb{K})$. For $w,z \in \mathbb{K} \setminus \{0\}$ define $c \in \mathbb{K}$ by $w = cz$. Then

$$\frac{H(w)}{H(z)} = \exp\bigl(k(c,z)\bigr) = \exp\bigl(h(w)-h(z)\bigr).$$

This extends by continuity to $w=0$ or $z=0$. Hence $\frac{H\circ f}{H} = \exp(h\circ f - h)$ for all $f \in \mathcal{P}(\mathbb{K})$. Since Tf, $\frac{H}{H\circ f}$ are in $\mathcal{E}(\mathbb{K})$, also $|f'|^p(\operatorname{sgn} f')^m$ has to be real-analytic ($\mathbb{K}=\mathbb{R}$) or analytic ($\mathbb{K}=\mathbb{C}$) for all polynomials f requiring that $p = m \in \mathbb{N}_0$, taking into account that $m \in \mathbb{Z}$, $\operatorname{Re}(p) \geq 0$. Therefore

$$Tf = \exp(h\circ f - h)f'^m, \quad f \in \mathcal{P}(\mathbb{K}), \tag{4.24}$$

$m \in \mathbb{N}_0$. Given any $f \in \mathcal{E}(\mathbb{K})$, its n-th order Taylor polynomials $p_n(f) \in \mathcal{P}(\mathbb{K})$ converge uniformly on compacta to f. By the assumption of pointwise continuity at 0 of T and (4.14), we have for any $z \in \mathbb{K}$, $Tf(z) = \lim_{n\to\infty} T(p_n(f))(z)$. Moreover, $\lim_{n\to\infty} h\circ p_n(f)(z) = h\circ f(z)$ and $\lim_{n\to\infty} p_n(f)'(z) = f'(z)$. Therefore, (4.24) holds for all $f \in \mathcal{E}(\mathbb{K})$. □

Remark. Imposing the additional initial condition $T(-2\,\mathrm{Id}) = -2$ on T in Theorems 4.9 and 4.10 will imply that $p = m = 1$ and that H and h are constant so that $Tf = f'$, i.e., T is the derivative.

4.3 Notes and References

Theorem 4.1 on the solution of the chain rule operator equation was shown by Artstein-Avidan, König and Milman in [AKM].

The proof of the continuity of the function H in part (iii) of the proof of Theorem 4.1 uses similar arguments as in the proof of Theorem 2.6 and as in Step

12 of the proof of Theorem 2 of Alesker, Artstein-Avidan, Faifman and Milman [AAFM].

If the "compound" product $T(f \circ g) \cdot Tg$ on the right side of the chain rule is replaced by a simple product of Tf and Tg, the resulting equation essentially has only trivial solutions, since the right-hand side does not reflect the effects of the composition. We have the following result, cf. Proposition 8 of [KM3]:

Proposition 4.11. *Let $k \in \mathbb{N}_0$ and suppose that $T : C^k(\mathbb{R}) \to C(\mathbb{R})$ satisfies*

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

Assume also that for any $x \in \mathbb{R}$ and any open interval $J \subset \mathbb{R}$ there is $g \in C^k(\mathbb{R})$ with $Im(g) \subset J$ such that $Tg(x) \neq 0$. Then $Tf = \mathbb{1}$ for all $f \in C^k(\mathbb{R})$.

Theorem 4.1 admits a cohomological interpretation. The semigroup $G = (C^k(\mathbb{R}), \circ\,)$ with composition as operation acts on the abelian semigroup $M = (C(\mathbb{R}), \cdot\,)$ with pointwise multiplication as operation by composition from the right, $G \times M \to M$, $fH := H \circ f$. Thus, M is a module over G. Denote the functions from G^n to M by $F^n(G, M)$ and define the coboundary operators

$$d^n : F^n(G, M) \longrightarrow F^{n+1}(G, M), \quad n \in \mathbb{N}_0,$$

using the *additive* notation $+$ for the operation $\cdot$ on M, by

$$\begin{aligned} d^n\varphi(g_1, \dots, g_{n+1}) &= g_1\varphi(g_2, \dots, g_{n+1}) \\ &\quad + \sum_{i=1}^{n} (-1)^i \varphi(g_1, \dots, g_{i-1}, g_i g_{i+1}, g_{i+2}, \dots, g_{n+1}) + (-1)^{n+1}\varphi(g_1, \dots, g_n), \end{aligned}$$

for $\varphi \in F^n(G, M)$, $g_1, \dots, g_{n+1} \in G$. Theorem 4.1 characterizes the cocycles in $\mathrm{Ker}(d^1)$ for $n = 1$. Then $\varphi = T : G = C^k(\mathbb{R}) \to M = C(\mathbb{R})$ has coboundaries

$$d^1T(g_1, g_2) = g_1T(g_2) - T(g_1g_2) + T(g_2), \quad g_1, g_2 \in G.$$

As for cocycles T, $d^1T = 0$ means in multiplicative notation

$$T(g_2 \circ g_1) = T(g_2) \circ g_1 \cdot Tg_1,$$

and these are just the solutions of the chain rule. For $n = 0$, $\varphi \in F^0(G, M)$ can be identified with $\varphi = H \in M = C(\mathbb{R})$ and we have in multiplicative notation $d^0H(g) = \frac{H \circ g}{H}$ for $g \in G = C^k(\mathbb{R})$.

The cohomology group $H^1(G, M) = \mathrm{Ker}(d^1)/\,\mathrm{Im}(d^0)$ is hence, by Theorem 4.1, represented by the maps $g \mapsto |g'|^p \ \{\mathrm{sgn}\, g'\}$ from G to M.

We are grateful to L. Polterovich and S. Alesker for advising us on this cohomological interpretation of Theorem 4.1.

Theorem 4.8 on the chain rule equation in $\mathbb{R}^n$ was proved by König and Milman in [KM2]. The result on the inner automorphisms of $GL(n, \mathbb{R})$, which replaces Proposition 2.3 in the proof for $n > 1$, is taken from Dieudonné [D] and Hua [H]. We are grateful to J. Bernstein and R. Farnsteiner for discussions concerning the proof of Theorem 4.8.

Theorems 4.9 and 4.10 were shown in [KM11]. We would like to thank P. Domański for helpful discussions concerning these results.

Corollary 4.3 stated that the derivative is the only operator (not vanishing on the half-bounded functions) satisfying the chain rule and the extended Leibniz rule. It is interesting to note that on the complex plane there are different operators satisfying the chain rule and the extended Leibniz rule, though not with image in the continuous functions: By Aczél, Dhombres [AD], Theorem 7 in Chapter 5.2, there is a non-zero additive and multiplicative function $K : \mathbb{C} \to \mathbb{C}$ which is not the identity on $\mathbb{C}$. Let $C^1(\mathbb{C})$ denote the continuously differentiable (i.e., entire) functions from $\mathbb{C}$ to $\mathbb{C}$ and $F(\mathbb{C})$ denote all functions from $\mathbb{C}$ to $\mathbb{C}$. Define operators $T, A : C^1(\mathbb{C}) \to F(\mathbb{C})$ by $Tf := K(f')$ and $Af := K(f)$. Then (T, A) satisfy

$$\begin{aligned} T(f \circ g) &= Tf \circ g \cdot Tg, \\ T(f \cdot g) &= Tf \cdot Ag + Af \cdot Tg \; ; \; f, g \in C^1(\mathbb{C}), \end{aligned}$$

but T is not the derivative and A is not the identity on $C^1(\mathbb{C})$.

The analogue of the chain rule in integration is the substitution formula. Let $c \in \mathbb{R}$ be fixed, $I : C(\mathbb{R}) \to C^1(\mathbb{R})$ denote the operator of definite integration from c to x and $D : C^1(\mathbb{R}) \to C(\mathbb{R})$ be the derivative. Then I is injective and

$$f \circ g - (f \circ g)(c) = I(Df \circ g \cdot Dg)$$

holds for all $f, g \in C^1(\mathbb{R})$. Modeling this, more generally we consider operators $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ and $J : C(\mathbb{R}) \to C^1(\mathbb{R})$ such that for some fixed $c \in \mathbb{R}$ and all $f, g \in C^1(\mathbb{R})$

$$f \circ g - (f \circ g)(c) = J(Tf \circ g \cdot Tg).$$

The natural question then is whether T is closely connected to some derivative and J to some definite integral. Let us call $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ non-degenerate if there is $y \in \mathbb{R}$ such that for all $x \in \mathbb{R}$ there is $f \in C_b^1(\mathbb{R})$ with $f(x) = y$ and $Tf(x) \neq 0$. Also $T(\mathrm{Id})(x) \neq 0$ is assumed for all $x \in \mathbb{R}$. We then have by König, Milman [KM12]:

Proposition 4.12. *Assume that $J : C(\mathbb{R}) \to C^1(\mathbb{R})$ and $T : C^1(\mathbb{R}) \to C(\mathbb{R})$ are operators such that for some fixed $c \in \mathbb{R}$*

$$f \circ g - (f \circ g)(c) = J(Tf \circ g \cdot Tg)$$

holds for all $f, g \in C^1(\mathbb{R})$. Suppose further that T is non-degenerate and that J is injective. Then there are constants $p > 0$, $d \neq 0$ such that for all $f \in C^1(\mathbb{R})$ and

$h \in C(\mathbb{R})$

$$Tf(x) = d\ |f'(x)|^p\ \operatorname{sgn} f'(x), \tag{4.25}$$

$$Jh(x) = d^{-2/p} \int_c^x |h(s)|^{1/p}\ \operatorname{sgn} h(s)\ ds. \tag{4.26}$$

If T additionally satisfies the initial condition $T(2\,\mathrm{Id}) = 2$, we have that $p = d = 1$ and

$$Tf(x) = f'(x), \quad Jf(s) = \int_c^x h(s)\ ds.$$

Hence T in (4.25) is a generalized derivative and J in (4.26) is a generalized definite integral. For the proof we refer to [KM12].