Journal of Convex Analysis
Volume 32 (2025), No. 1, 263–290

Derivative of Inverse Mapping and Chain Rule Formulas for Gateaux Derivatives, Theorems and Counter-Examples

Marián Fabian, Jan Kolář
Institute of Mathematics, Czech Academy of Sciences, Praha, Czech Republic
fabian@math.cas.cz, kolar@math.cas.cz

Dedicated to Ralph Tyrrell Rockafellar on the occasion of his 90th birthday.

Received: September 9, 2024
Accepted: October 9, 2024

We investigate what happens if Fréchet differentiability is replaced by Gateaux differentiability in several statements from differential calculus. Our main counter-example shows how the formula for derivative of inverse mapping fails when Gateaux differentiability is assumed both for the mapping and its inverse: We construct an involution $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ with $f(0,0) = (0,0)$, C^∞-smooth off the origin and Gateaux differentiable at $(0,0)$, whose derivative $f'(0,0)$ is singular, and hence $f'(0,0) \circ f'(0,0)$ is not the identity mapping. This mapping simultaneously provides a strong counter-example for the Chain Rule formula with Gateaux derivatives. We also show that the chain rule formula holds true for the Gateaux derivatives in an important special case of mappings between Banach spaces.

Keywords: Chain rule formula, inverse mapping theorem, derivative, Fréchet differentiability, Gateaux differentiability, involution

2020 Mathematics Subject Classification: 26B10, 46G05, 49J50, 58C20.

In many versions of the Inverse Mapping Theorem, the assumption about the Fréchet derivative plays a decisive role, though it sometimes may be disguised under various names, such as (total) differential or *the* derivative in Euclidean spaces, or under generous assumptions like continuity of derivatives or (in Euclidean spaces) Lipschitz condition.

Our focus is on the cases when the differentiability is presumed at a single point. We may assume, without loss of generality, that this point is the origin, and also the image of it is the origin.

When the assumption is relaxed to Gateaux differentiability, counterexamples are easy to find and rather well known. If continuity and (perhaps also) injectivity is demanded from the mapping in question, a moderate difficulty arises. We will touch upon this to make our paper more comprehensive and also to provide some more penetrable material before giving our main results.

We go further and add assumptions about the inverse mapping, most prominently its Gateaux differentiability at 0. That means, instead of expecting this to be a

Both authors acknowledge the support of the Institute of Mathematics of the Czech Academy of Sciences (RVO 67985840).

ISSN 0944-6532 / \$ 2.50

conclusion, we assume the existence of Gateaux derivative of the inverse mapping and ask if it must satisfy the usual formula $f'(0)\circ(f^{-1})'(0) = \text{Identity}$. The question is: *Can the derivative of the inverse mapping exist and differ from what prescribes the well known formula?* Here the interesting phenomenon arises, which we report in this paper: no counter-example exists for a (natural) combination of assumptions, if it includes also the usual requirement of surjectivity of the derivative, see Corollary 6.9. Without the surjectivity assumption, we have Example 5.1.

The formula for the derivative of inverse mapping is related to the Chain Rule formula for derivatives of compound mappings. Here the situation is similar: When we replace the standard Fréchet (or Hadamard) differentiability assumption by Gateaux differentiability, there exist counter-examples whose difficulty is proportional to the sophistication of the questions. When we assume Gateaux differentiability of the compound mapping and some continuity, the interesting phenomenon arises again: If the derivative of the inner function is surjective, the derivatives satisfy the Chain Rule formula. Without the additional assumption, we have counter-examples, see Example 4.3 and Example 5.1.

1. Preliminaries

Let $f : X \longrightarrow Y$ be a mapping between two real normed spaces, and let $x \in X$. We say that f *is Gateaux differentiable* at x, or that f *has a Gateaux derivative* at x, if x is in the algebraic interior of the domain of f, the limit

$$\lim_{t\to 0} \frac{1}{t}(f(x+th) - f(x)) =: Df(x)(h) \in Y$$

exists for every $h \in X$ and the assignment $X \ni h \longmapsto Df(x)(h)$ is linear and continuous. We say that f *is Fréchet differentiable* at x or that f *has a Fréchet derivative* at x, if x is in the interior of the domain of f, if f is Gateaux differentiable at x and the limit above is uniform over all h from the unit ball of X; equivalently, if there is a linear continuous mapping $L : X \longrightarrow Y$ such that.

$$\lim_{h\to 0} \frac{1}{\|h\|}(f(x+h) - f(x) - L(h)) = 0.$$

In both cases, we denote the mapping $Df(x)(\cdot)$ by $f'(x)$ and call it the *Gateaux* or *Fréchet derivative* accordingly. The above definitions can be easily extended to multivalued mappings $f : X \longrightarrow 2^Y$ provided that $f(x)$ is a singleton.

Trivially, Fréchet differentiability implies Gateaux one. If X is finitely dimensional and f is Lipschitzian near an $x \in X$, it is easy to check that Gateaux differentiability of f at x implies Fréchet differentiability. Another simple but important fact says that, if f is Gateaux differentiable in a vicinity of an $x \in X$ and $f'(\cdot)$ is continuous at x, then f is Fréchet differentiable at x.

For $p \in \mathbb{N}$, the *p-th differentiability* and *p-th derivative* $f^{(p)}(\cdot)$ are defined in a natural way by induction. Let $f : X \to Y$ be a mapping whose domain contains an open set $\Omega \subset X$. We say that f is C^p*-smooth* on Ω if the p-th (Fréchet) derivative $f^{(p)}(\cdot)$ exists and is continuous on Ω. We say that f is C^∞*-smooth* on Ω if it is C^p-smooth on Ω for every $p \in \mathbb{N}$. The C^0*-smoothness* will mean just the continuity.

Sometimes, we will also work with (real) *analytic mappings* (that is, C^ω*-smooth mappings*). We recall that every mapping, analytic on an open set, is C^∞-smooth there. For more information we refer for instance to [9, p. 70] and to Chapter 2 in [10].

In what follows, we almost always assume that both X and Y are Euclidean spaces, say $\mathbb{R}^n$ and $\mathbb{R}^m$, provided with the Euclidean norms. There exists a related and historically older concept of Jacobian: Consider a mapping

$$f(\cdot) := \big(f_1(\cdot), \dots, f_m(\cdot)\big) : \mathbb{R}^n \longrightarrow \mathbb{R}^m,$$

a point $x \in \mathbb{R}^n$ lying in the interior of the domain of f, and assume that all partial derivatives $\partial f_i(x)/\partial x_j,\ m = 1, \dots, i,\ j = 1, \dots, n$, exist. Then the *Jacobian* of f at x is defined as the matrix

$$Jf(x) := \begin{pmatrix} \partial f_1(x)/\partial x_1 & \cdots & \partial f_1(x)/\partial x_n \\ \vdots & \ddots & \vdots \\ \partial f_m(x)/\partial x_1 & \cdots & \partial f_m(x)/\partial x_n \end{pmatrix}.$$

Clearly, if f is Gateaux differentiable at x and $e_1, \dots, e_n$ denote the canonical unit basic vectors in $\mathbb{R}^n$, then $f_i'(x)e_j = \partial f_i(x)/\partial x_j$ and $(Jf(x)h^{\mathrm{T}})^{\mathrm{T}} = f'(x)h$ for every $h \in \mathbb{R}^n$ (where the column vector h^{T} is the transpose of the row vector h). Note that it is rather easy to construct a (non-convex) function $f : \mathbb{R}^n \longrightarrow \mathbb{R}^1$ for which $Jf(x)$ can be calculated and yet f is not Gateaux differentiable at x.

Of our frequent use will be the following special case of [11, Theorem 7.1, p. 352]:

Proposition 1.1. *Let $m, n, p \in \mathbb{N}$ and let $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a mapping whose Jacobian $Jf(\cdot)$ exists and is C^{p-1}-smooth on an open set $\Omega \subset \mathbb{R}^n$. Then f is C^p-smooth on Ω.*

If f, g are two mappings, then their composition $f(g(\cdot))$ (if it is possible) is denoted by $f \circ g$.

2. Interior mapping theorem

We start our series of examples by considering a rather simple and known consequence of the deep Brouwer's fixed-point theorem, see for instance [3, Corollary 3.5]:

Theorem 2.1. *Let $0 < m \leq n$ be integers and consider a continuous mapping $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ which is Fréchet differentiable (just) at some $x \in \mathbb{R}^n$. Assume moreover that the rank of the Jacobian $Jf(x)$ is m.*

Then the point $f(x)$ lies in the interior of the range $f(\mathbb{R}^n)$ of f; more precisely, there exists $c > 0$ such that $f(x + t\,\mathbb{B}_n) \supset f(x) + ct\,\mathbb{B}_m$ for all sufficiently small $t > 0$.

One may ask what happens if Fréchet differentiability above is replaced by the Gateaux one. An answer is given below.

Example 2.2. There exists a mapping $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ such that:

(o) $f(0,0) = (0,0)$; (i) f is continuous at $(0,0)$; (ii) f is C^∞-smooth on $\mathbb{R}^2 \backslash \{(0,0)\}$;

(iii) f is Gateaux differentiable at $(0,0)$ with Jacobian $Jf(0,0) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$; and yet

(iv) the range $f(\mathbb{R}^2)$ of f does not contain any neighborhood of $(0,0)$.

Proof. First, for $(x,y) \in \mathbb{R}^2$ define

$$f(x,y) := \begin{cases} (0,0) & \text{if } y = x^2, \text{ and} \\ (x,y) & \text{otherwise.} \end{cases}$$

A straightforward calculation shows that f is Gateaux differentiable at $(0,0)$, with $Jf(0,0) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. Yet $f(\mathbb{R}^2)$ does not contain the parabola

$$P := \{(x,y) \in \mathbb{R}^2 : \ 0 \neq y = x^2\}.$$

Unfortunately, f is not C^∞-smooth, even not continuous at the points of P.

The next construction remedies this non-smoothness. Take a C^∞-smooth function $\eta : \mathbb{R} \longrightarrow [0,1]$ such that

$$\eta(t) := 1 \ \text{ if } \ t \in (-\infty, 1] \cup [4, \infty) \ \text{ and } \ \eta(t) := 0 \ \text{ if } \ t \in [2,3];$$

see, for instance [11, p. 169]. (Note that such an η is not analytic.) For $(x,y) \in \mathbb{R}^2$ define

$$f(x,y) := \begin{cases} \big(\eta(y/x^2)x,\, \eta(y/x^2)^2 y\big) & \text{if } x \neq 0, \ \text{ and} \\ (0,y) & \text{if } \ x = 0. \end{cases} \tag{1}$$

(Note that this new f maps the parabola $\{(t, \kappa t^2) : \, t \in \mathbb{R}\}$ into the same parabola for every $\kappa > 0$.) As $0 \leq \eta(\cdot) \leq 1$, our f is continuous at $(0,0)$.

Let us check the smoothness of f off the origin. So take any $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$. Assume that $x \neq 0$. Then, whenever x' is close enough to x and y' is any number, we have $f(x',y') = \big(\eta(y'/x'^2)x',\, \eta(y'/x'^2)^2 y'\big)$, and therefore our f is C^∞-smooth at each (x,y) with $x \neq 0$. Assume now that $x = 0$, $y \neq 0$. Take any $(x',y') \in \mathbb{R}^2$ such that $|x'| < \frac{1}{3}\sqrt{|y|}$, and $|y' - y| < \frac{1}{3}|y|$. If $x' = 0$, we have $f(x',y') = (x',y')$.

Assume further that $x' \neq 0$. If $y < 0$, then $y'/x'^2 < 0$ (< 1). And if $y > 0$, then $y'/x'^2 > \frac{2}{3}|y|/(\frac{1}{9}|y|) = 6$ (> 4). Therefore $\eta(y'/x'^2) = 1$ and $f(x',y') = (x',y')$. We proved that our f equals the identity in a vicinity of (x,y), and so it is C^∞-smooth at (x,y).

We will prove that f is Gateaux differentiable at $(0,0)$. So fix any $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$. If $x = 0$, we have

$$f(tx,ty) - f(0,0) = (0,ty) = t(x,y)$$

for all $t \in \mathbb{R}$. And, if $x \neq 0$, then for all $0 \neq t \in \mathbb{R}$, sufficiently close to 0, we have

$$f(tx,ty) = \big(\eta(ty/(tx)^2)tx,\, \eta(ty/(tx)^2)^2 ty\big) = t(x,y).$$

Therefore, f is Gateaux differentiable at $(0,0)$, with $Jf(0,0) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

The range $f(\mathbb{R}^2)$ does not meet the points of the set $\{(u,v) \in \mathbb{R}^2 : \ 2u^2 < v < 3u^2\}$. Indeed, take a point (u,v) in it. Then $u \neq 0$ necessarily. Assume that there is $(x,y) \in \mathbb{R}^2$ such that $f(x,y) = (u,v)$. Then $x \neq 0$ necessarily. Thus we have $(u,v) = \big(\eta(y/x^2)x,\, \eta(y/x^2)^2 y\big)$, and so, $\eta(y/x^2) \neq 0$ and

$$(2,3) \ni \frac{v}{u^2} = \frac{\eta(y/x^2)^2 y}{(\eta(y/x^2)x)^2} = \frac{y}{x^2}\,.$$

Hence $\eta(y/x^2) = 0$, a contradiction. $\square$

It is popular to present examples of functions in the form of 'simple' formulas, as for example $\mathbb{R}^2 \ni (x, y) \longmapsto \exp(-(y - x^2)^2/x^4)$ in [8]. Influenced by this, we offer the following

Proof. (An alternative construction of Example 2.2) Define $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ as

$$f(x,y) := \begin{cases} \Big(\big(1-\exp\big(-\frac{(y-x^2)^2}{x^6+y^6}\big)\big)x,\ \big(1-\exp\big(-\frac{(y-x^2)^2}{x^6+y^6}\big)\big)^2 y\Big) & \\ & \text{if } (x,y)\in\mathbb{R}^2\setminus\{(0,0)\}, \\ (0,0) & \text{if } (x,y)=(0,0). \end{cases} \tag{2}$$

A direct calculation reveals that $Df(0,0)(x,y) = (x,y)$ for every $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$. Thus, f is Gateaux differentiable at the origin with Jacobian $Jf(0,0) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

As the argument of exp function in (2) is never positive, we have $\|f(x,y)\| \le \|(x,y)\|$, and so f is continuous at the origin. Also, a brief look at (2) reveals that f is C^∞-smooth, even analytic, on $\mathbb{R}^2 \setminus \{(0,0)\}$.

We observe that $f(\mathbb{R}^2)$ does not contain any point of the parabola $\{(t,t^2) :\ t \neq 0\}$; thus $(0,0)$ will not be in the interior of the range of f. Indeed, assume that there are $0 \neq t \in \mathbb{R}$ and $(x,y) \in \mathbb{R}^2$ such that $f(x,y) = (t,t^2)$. Put

$$\varphi := 1 - \exp(-(y-x^2)^2/(x^6+y^6)).$$

Thus $(\varphi x, \varphi^2 y) = (t, t^2)$, and so $\varphi x \neq 0$. But $1 = t^2/t^2 = \varphi^2 y/(\varphi^2 x^2) = y/x^2$, and so $y = x^2$. By the definition of φ we obtain that $\varphi = 0$, a contradiction. $\square$

Remark 2.3. We note that the constructions in Example 2.2 can be extended to higher dimensions, and also to the space $\ell_2 \times \mathbb{R}$. It is enough to replace, "x^2" by "$\|x\|^2$" in (1) and (2), where $\|\cdot\|$ is the Hilbertian norm.

Remark 2.4. The function $[0,\infty) \ni t \longmapsto \exp(-t)$ in (2) can be replaced by $[0,\infty) \ni t \longmapsto 1/(1+t)$; then our f will be even rational, except at the origin, and we have the formula

$$f(x,y) := \left(\frac{(y-x^2)^2}{x^6+y^6+(y-x^2)^2}\,x,\ \left(\frac{(y-x^2)^2}{x^6+y^6+(y-x^2)^2}\right)^2 y\right), \quad (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}.$$

3. Differentiability of inverse mappings

For instance in [11, p. 361], there is a well known inverse mapping theorem dealing with C^p-smoothness. The proof of it is not dramatically deep; just Banach's contraction principle is needed. Here we state a version with weaker assumptions and conclusions:

Theorem 3.1. *Let $n \in \mathbb{N}$ and consider a continuous mapping $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ which is Fréchet differentiable at some $x \in \mathbb{R}^n$. Assume that the (multivalued) mapping $f^{-1} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is singlevalued at the point $f(x) =: y$ and upper semi-continuous at y and that, moreover, the derivative $f'(x)$ is an isomorphism.*

Then f^{-1} is Fréchet differentiable at y with $(f^{-1})'(y) = f'(x)^{-1}$ and so we have $Jf^{-1}(y) = Jf(x)^{-1}$.

Proof. By Theorem 2.1, $f(x)$ lies in the interior of $f(\mathbb{R}^n)$. Thus we may apply a modification of [1, Theorem 1.5, page 215], or [6, Corollary 3.1]. See also [4, Theorem 1G.1]. $\square$

Our concern is: What happens if the Fréchet differentiability above is replaced by the one of Gateaux? The answer is below.

Example 3.2. There exists a homeomorphism f of $\mathbb{R}^2$ onto $\mathbb{R}^2$ such that:

(o) $f(0,0) = (0,0)$;

(i) f is C^∞-smooth on $\mathbb{R}^2 \setminus \{(0,0)\}$;

(ii) f^{-1} is C^∞-smooth on $\mathbb{R}^2 \setminus \{(0,0)\}$;

(iii) f is Gateaux differentiable at $(0,0)$, with $Jf(0,0) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$; and

(iv) f^{-1} is not Gateaux differentiable at $(0,0)$.

Proof. We realize that, in order to get spoiled Gateaux differentiability at a point, it is enough to do so in one direction, say in the vertical one $(0,1)$. Our mapping f will be of form $f_1 \circ b$, where $b : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is a simple C^∞-diffeomorphism and f_1 is a slight mutation of the identity mapping.

First, we will construct a lower quality version of the mapping $f_1 : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, without the C^∞-smoothness properties. Let $\eta : \mathbb{R} \longrightarrow \mathbb{R}$ be an increasing homeomorphism, with $\eta(0) = 0$, having the derivatives $\eta'(y) > 0$ at every $0 \neq y \in \mathbb{R}$, and such that $\eta'(0)$ and $(\eta^{-1})'(0)$ exist and belong to $[0, +\infty] \setminus \{1\}$; η might possibly be a C^∞-diffeomorphism of $\mathbb{R} \setminus \{0\}$. Our favourite options will be

$$\eta(t) := t^3, \quad \text{or} \quad \eta(t) := \operatorname{sgn}(t)\sqrt[3]{|t|}, \quad \text{or} \quad \eta(t) := 2t, \quad \text{for } t \in \mathbb{R}. \tag{3}$$

(The choice $\eta(t) := t,\ t \in \mathbb{R}$, is not helpful for us.)

For $(x, y) \in \mathbb{R}^2$ we define

$$f_1(x,y) := \begin{cases} (x,y) & \text{if } x \neq 0; \\ (0, \eta(y)) & \text{if } x = 0. \end{cases} \tag{4}$$

Clearly, f_1 is a bijection of $\mathbb{R}^2$ onto $\mathbb{R}^2$, and for every $(x, y) \in \mathbb{R}^2$

$$f_1{}^{-1}(x,y) = \begin{cases} (x,y) & \text{if } x \neq 0; \\ (0, \eta^{-1}(y)) & \text{if } x = 0. \end{cases}$$

We can easily verify that

$$Df_1(0,0)(x,y) = (x,y) = Df_1{}^{-1}(0,0)(x,y) \quad \text{if} \quad (x,y) \in \mathbb{R}^2 \text{ and } x \neq 0,$$

and also that

$$Df_1(0,0)(0,1) = (0, \eta'(0)) \neq (0,1), \quad Df_1{}^{-1}(0,0)(0,1) = (0, (\eta^{-1})'(0)) \neq (0,1).$$

Hence, using the very definition, neither f_1 nor $f_1{}^{-1}$ is Gateaux differentiable at $(0,0)$.

We wish to find a mapping which **is** Gateaux differentiable at $(0,0)$ (and whose inverse **is not** Gateaux differentiable at $(0,0)$.)

We achieve this by composing our f_1, from the right, with the "breeze" mapping $b : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ defined as, say

$$b(x, y) := (x - y^2, y), \quad (x, y) \in \mathbb{R}^2, \tag{5}$$

which is clearly a bijection, even a C^∞-difeomorphism of $\mathbb{R}^2$ onto $\mathbb{R}^2$. (Later it will be helpful to realize that the vertical axis $\{0\} \times \mathbb{R}$ is mapped by the inverse mapping b^{-1} into the parabola $\{(x, y) \in \mathbb{R}^2 : \ x = y^2\}$.) Now, let us put (solemnly)

$$f := f_1 \circ b, \quad \text{that is,} \quad f(x, y) := f_1(b(x, y)), \quad (x, y) \in \mathbb{R}^2. \tag{6}$$

We will show that this f is Gateaux differentiable at $(0, 0)$. So fix any $(0, 0) \neq (x, y) \in \mathbb{R}^2$. Then $tx - (ty)^2 \neq 0$ for all $0 \neq t \in \mathbb{R}$ sufficiently close to 0. Thus for these t's we have

$$\frac{1}{t}\big(f(tx, ty) - f(0, 0)\big) = \frac{1}{t} f_1(tx - t^2y^2, ty) = \frac{1}{t}(tx - t^2y^2, ty) \longrightarrow (x, y) \quad \text{as } t \to 0.$$

Therefore $Df(0, 0)(x, y) = (x, y)$, and so $Jf(0, 0) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$. We verified that f is Gateaux differentiable at the origin.

On the other hand, f^{-1} is not Gateaux differentiable at $(0, 0)$. Indeed, realizing that $f^{-1} = b^{-1} \circ {f_1}^{-1}$, for all $t \in \mathbb{R}$ we have

$$f^{-1}(0, t) - f^{-1}(0, 0) = b^{-1}({f_1}^{-1}(0, t)) = b^{-1}(0, \eta^{-1}(t)) = ((\eta^{-1}(t))^2, \eta^{-1}(t)),$$

and hence the calculation of $Df^{-1}(0, 0)(0, 1)$ leads either to "(∞, ∞)", or to $(0, 0)$, or to $(0, \frac{1}{2})$. Thus, f^{-1} is not Gateaux differentiable if we take $\eta(y) := y^3$. Concerning the other two options, when $\eta(y) := \operatorname{sgn}(y)\sqrt[3]{|y|}$ or $\eta(y) := 2y$, we have to dig further: For every $(x, y) \in \mathbb{R}^2 \setminus (\{0\} \times \mathbb{R})$, we have

$$\begin{aligned} \frac{1}{t}(f^{-1}(tx, ty) - f^{-1}(0, 0)) &= \frac{1}{t} b^{-1}({f_1}^{-1}(tx, ty)) = \frac{1}{t} b^{-1}(tx, ty) \\ &= \frac{1}{t}(tx + (ty)^2, ty) \longrightarrow (x, y) \ \text{ as } 0 \neq t \to 0. \end{aligned}$$

Thus $Df^{-1}(0, 0)(1, 1) - Df^{-1}(0, 0)(1, 0) = (1, 1) - (1, 0) \neq Df^{-1}(0, 0)(0, 1)$.

The linearity of the mapping $\mathbb{R}^2 \ni h \longmapsto Df^{-1}(0, 0)h$ is thus violated. Therefore f^{-1} is again not Gateaux differentiable at $(0, 0)$.

The mapping f_1 is not continuous at the points $(0, y)$, $0 \neq y \in \mathbb{R}$. We may overcome this defect by 'sliding' judiciously the values of the function along the parabolas $K_c = \{(ct^2, t) : \ t \in \mathbb{R}\}$, $c \in [-1, 1]$. In other words, our intention is to define, below, mapping $\widehat{f_1}$ that sends every parabola K_c into itself.

We wish to note that the idea of 'sliding' along curves will be used again in the more advanced formula (17) in Example 5.1.

We consider a C^∞-smooth function $\xi : \mathbb{R} \longrightarrow [0, 1]$ such that $\xi(0) := 0$ and $\xi(t) := 1$ for $t \in (-\infty, -1] \cup [1, \infty)$; see, e.g., [11, pp. 168, 169]. We put

$$\varphi(x, y) := (1 - \xi(x/y^2))\eta(y) + \xi(x/y^2)y\big), \quad \text{for } (x, y) \in \mathbb{R}^2 \setminus (\mathbb{R} \times \{0\})$$

and $\varphi(x, 0) := 0$ for all $x \in \mathbb{R}$.

Clearly, φ is C^∞-smooth on $\mathbb{R}^2\setminus\{(0,0)\}$. The value of $\varphi(x,y)$ is a convex combination of y and $\eta(y)$; this is our simple core tool to remedy the discontinuity. The rest of the expression may be explained as follows. The fraction $c := x/y^2$ identifies to which parabola K_c the point (x,y) belongs. Mapping φ then determines to which point of the parabola we 'slide' the point. Specifically, we choose $\varphi(x,y)$ to determine the second coordinate of the value. The first coordinate then must be calculated as $c\varphi(x,y)^2 = x\varphi(x,y)^2/y^2$, which expression we therefore use in the following definition. We define

$$\widehat{f_1}(x,y) := \big(x\varphi(x,y)^2/y^2, \varphi(x,y)\big), \quad \text{when } (x,y) \in \mathbb{R}^2 \setminus (\mathbb{R}\times\{0\}),$$

and $\widehat{f_1}(x,0) := (0,0)$ for all $x \in \mathbb{R}$. Finally, we define

$$f := \widehat{f_1} \circ b, \quad \text{that is,} \quad f(x,y) := \widehat{f_1}(b(x,y))), \quad (x,y) \in \mathbb{R}^2, \tag{7}$$

where the C^∞-diffeomorphism $b : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ was defined in (5). Some effort is required to prove formally the properties of $\widehat{f_1}$ and f. For details, see Appendix. □

Proof. (An alternative construction of Example 3.2) Consider a slight variation of the mapping from (2): For $(x,y) \in \mathbb{R}^2$ we define

$$f_1(x,y) := \begin{cases} \Big(1 - \frac{1}{2}\exp\big(-\frac{y^2}{x^6+y^6}\big)\Big)(x,y) & \text{if } (x,y) \neq (0,0), \\ (0,0) & \text{if } (x,y) = (0,0). \end{cases} \tag{8}$$

and then we put $$f(x,y) := f_1(x, y-x^2), \quad (x,y) \in \mathbb{R}^2. \tag{9}$$

We can easily observe that $\frac{1}{2}\|(x,y)\| \le \|f_1(x,y)\| \le \|(x,y)\|$ for all $(x,y) \in \mathbb{R}^2$. Hence f and (possibly the multivalued) f^{-1} are continuous at the origin. Clearly, f_1 is analytic on $\mathbb{R}^2 \setminus \{(0,0)\}$, and so is f there.

Let us focus on the Jacobian Jf_1. Fix any $(x,y) \in \mathbb{R}^2 \setminus (\{(0,0)\}$ for a while. Denote

$$\varphi = \varphi(x,y) := 1 - \frac{1}{2}\exp\Big(-\frac{y^2}{x^6+y^6}\Big), \quad \varphi_x := \frac{\partial\varphi(x,y)}{\partial x}, \quad \varphi_y := \frac{\partial\varphi(x,y)}{\partial y}.$$

Then $$Jf_1(x,y) = \begin{pmatrix} \varphi + x\varphi_x & x\varphi_y \\ y\varphi_x & \varphi + y\varphi_y \end{pmatrix}$$

and $$\det Jf_1(x,y) = (\varphi + x\varphi_x)(\varphi + y\varphi_y) - x\varphi_y \cdot y\varphi_x = (\varphi + x\varphi_x + y\varphi_y)\varphi.$$

We focus on

$$\begin{aligned} \varphi + x\varphi_x + y\varphi_y &= 1 - \frac{1}{2}\exp\Big(-\frac{y^2}{x^6+y^6}\Big) - \frac{1}{2}x\frac{\partial}{\partial x}\exp\Big(-\frac{y^2}{x^6+y^6}\Big) \\ &\quad - \frac{1}{2}y\frac{\partial}{\partial y}\exp\Big(-\frac{y^2}{x^6+y^6}\Big) \\ &= 1 - \frac{1}{2}\exp\Big(-\frac{y^2}{x^6+y^6}\Big) - \frac{1}{2}x\exp\Big(-\frac{y^2}{x^6+y^6}\Big)\frac{6y^2x^5}{(x^6+y^6)^2} \\ &\quad - \frac{1}{2}y\exp\Big(-\frac{y^2}{x^6+y^6}\Big)\frac{-2y(x^6+y^6)+6y^2y^5}{(x^6+y^6)^2} \end{aligned}$$

$$= 1 - \frac{1}{2}\exp\Big(-\frac{y^2}{x^6+y^6}\Big) + \frac{1}{2}\exp\Big(-\frac{y^2}{x^6+y^6}\Big)\frac{-6x^6y^2+2y^2x^6+2y^8-6y^8}{(x^6+y^6)^2}$$
$$= 1 - \frac{1}{2}\exp\Big(-\frac{y^2}{x^6+y^6}\Big) + \exp\Big(-\frac{y^2}{x^6+y^6}\Big)\frac{-2y^2}{x^6+y^6}\,.$$

Hence, putting $\dfrac{y^2}{x^6+y^6} := \alpha$, we have that

$$\varphi + x\varphi_x + y\varphi_y = 1 - \tfrac{1}{2}\exp(-\alpha) - 2\alpha\exp(-\alpha).$$

A simple investigation reveals that the function

$$[0,\infty) \ni s \longmapsto 1 - \tfrac{1}{2}\exp(-s) - 2s\exp(-s) =: \rho(s) \tag{10}$$

is bounded below by $1 - 2\exp(-3/4)$ (> 0). Therefore, as $\varphi(\cdot) \geq \frac{1}{2}$, we conclude

$$\det Jf_1(x,y) = (\varphi + x\varphi_x + y\varphi_y)\varphi \geq \frac{1}{2} - \exp(-3/4)\ > 0 \tag{11}$$

for every $(0,0) \neq (x,y) \in \mathbb{R}^2$.

Once having this, the (local) inverse mapping theorem (see [11, Chapter XIV, §1, Theorem 1.2, p. 361]) guarantees that our f_1 is locally C^p-invertible at each point $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$ for every $p \in \mathbb{N}$. In particular, f_1 is locally injective.

Now we aim to prove that f_1 maps $\mathbb{R}^2 \setminus \{(0,0)\}$ injectively onto $\mathbb{R}^2 \setminus \{(0,0)\}$. As f_1 maps the ray $\{(sa,sc) :\ s > 0\}$ into itself, for every $(a,c) \in \mathbb{R}^2 \setminus \{(0,0)\}$, we may consider each ray separately. We fix $(a,c) \in \mathbb{R}^2 \setminus \{(0,0)\}$ and, for $t > 0$, let $\psi(t) := t\varphi(ta,tc)$. Then ψ is a smooth function and $f_1(ta,tc) = \psi(t)(a,c)$. We only need to check that ψ injectively maps $(0,\infty)$ onto the same interval. We now that $M := Jf_1(ta,tc)$ is a regular matrix. Hence

$$(M(a,c)^T)^T = f_1'(ta,tc)(a,c) = \psi'(t)(a,c)$$

is a non-zero vector and we have that $\psi'(t) \neq 0$, obtaining the injectivity of ψ. To obtain the surjectivity we observe that $1 \leq \varphi(\cdot) \leq 2$, $\psi(t) \to 0$ as $t \downarrow 0$ and $\psi(t) \to \infty$ as $t \uparrow \infty$, and use the intermediate value theorem.

So f_1 maps $\mathbb{R}^2$ onto $\mathbb{R}^2$ injectively and the inverse ${f_1}^{-1}$on all of $\mathbb{R}^2$ exist. From the above we have that ${f_1}^{-1}$ is C^p-smooth on all of $\mathbb{R}^2 \setminus \{(0,0)\}$ for every $p \in \mathbb{N}$. Therefore ${f_1}^{-1}$ is also C^∞-smooth therein.

The smoothness properties of f_1 and ${f_1}^{-1}$ off the origin easily transfer to the final mapping f defined by (9) and to its inverse f^{-1} $(= b^{-1} \circ {f_1}^{-1})$. Indeed, we have $b(0,0) = (0,0) = b^{-1}(0,0)$ and both b and b^{-1} are polynomials, hence C^∞-smooth mappings. Thus it is enough to apply [11, Chapter XIII, §6, Theorem 6.5, p. 351].

It remains to study the Gateaux differentiability of the mapping f defined by (9) and of its inverse f^{-1} at the origin. First, we focus on f_1. For every $(a,c) \in \mathbb{R}$, with $c \neq 0$, we have

$$Df_1(0,0)(a,c) = \lim_{t\to 0}\frac{1}{t}f_1(ta,tc)) = \lim_{t\to 0}\frac{1}{t}\Big(1 - \frac{1}{2}\exp\Big(-\frac{t^2c^2}{t^6(a^6+c^6)}\Big)\Big)(ta,tc) = (a,c)$$

while $\quad Df_1(0,0)(1,0) = \lim\limits_{t\to 0}\dfrac{1}{t}f_1(t,0) = \lim\limits_{t\to 0}\dfrac{1}{t}\big(1 - \dfrac{1}{2}\exp(0)\big)(t,0) = \big(\tfrac{1}{2},0\big).$

Thus

$$Df_1(0,0)(0,1) + Df_1(0,0)(1,0) = (0,1) + \left(\tfrac{1}{2},0\right) = \left(\tfrac{1}{2},1\right) \neq (1,1) = Df_1(0,0)(1,1),$$

and so the chance for Gateaux differentiability of f_1 at the origin collapses.

Fortunately, the compound mapping f defined by (9) is Gateaux differentiable at $(0,0)$. (This is caused by composing f_1 with the breeze mapping b (5) from right.) Indeed, for every $(a,c) \in \mathbb{R}^2 \setminus \{(0,0)\}$ we have

$$\frac{1}{t}(f(at,ct) - f(0,0)) = \frac{1}{t}\Big(1 - \frac{1}{2}\exp\Big(-\frac{(ct-a^2t^2)^2}{(a^6t^6 + (ct-a^2t^2)^6)}\Big)\Big)(at, ct - a^2t^2) \longrightarrow (a,c)$$

as $t \to 0$. Hence f is Gateaux differentiable at the origin, with $Jf(0,0) = \left(\begin{smallmatrix}1&0\\0&1\end{smallmatrix}\right)$.

It remains to investigate the differentiability of f_1^{-1} and then of f^{-1} at the origin. For $t \in \mathbb{R}$ denote ${f_1}^{-1}(t,0) =: (u_t, v_t)$. Then $f_1(u_t,v_t) = (t,0)$, that is by (8), $\varphi(u_t,v_t)(u_t,v_t) = (t,0)$. We know that $\frac{1}{2} \leq \varphi(u_t,v_t) \leq 1$. Thus $v_t = 0$, $\varphi(u_t,0)u_t = t$, and

$$\frac{1}{t}({f_1}^{-1}(t,0) - (0,0)) = \left(\frac{u_t}{t}, 0\right) = \left(\frac{1}{\varphi(u_t,0)}, 0\right) = \left(\frac{1}{1 - \frac{1}{2}\exp(0)}, 0\right) = (2,0).$$

Therefore $D{f_1}^{-1}(0,0)(1,0) = (2,0)$.

Now consider any $(a,c) \in \mathbb{R}^2$ with $c \neq 0$. For $t \in \mathbb{R}$ denote ${f_1}^{-1}(ta,tc) =: (u_t,v_t)$. Then $f_1(u_t,v_t) = (ta,tc)$ and so, $\varphi(u_t,v_t)(u_t,v_t) = (ta,tc)$. Thus $|ta| \leq |u_t| \leq 2|ta|$ and $|tc| \leq |v_t| \leq 2|tc|$. We can then easily deduce that

$$\frac{{v_t}^2}{{u_t}^6 + {v_t}^6} \geq \frac{c^2}{2^6(a^6+c^6)t^4}$$

and so $$0 < \exp\left(\frac{-{v_t}^2}{{u_t}^6 + {v_t}^6}\right) \leq \exp\left(\frac{-c^2}{2^6(a^6+c^6)t^4}\right) \longrightarrow \exp(-\infty) = 0$$

as $t \to 0$. Hence $\varphi(u_t,v_t) \to 1$ as $t \to 0$, and so we finally get $D{f_1}^{-1}(0,0)(a,c) = (a,c)$. But then

$$Df_1^{-1}(0,0)(0,1) + Df_1^{-1}(0,0)(1,0) = (0,1) + (2,0) = (2,1) \neq (1,1) = Df_1^{-1}(0,0)(1,1).$$

We disproved the Gateaux differentiability of ${f_1}^{-1}$ at the origin. Hence, neither f^{-1} is Gateaux differentiable at $(0,0)$. Indeed, if so, then $b \circ f^{-1}$ were Gateaux differentiable at $(0,0)$. But $b \circ f^{-1} = {f_1}^{-1}$, a contradiction. □

Remark 3.3. The mapping f in (9), (8) is analytic on $\mathbb{R}^2 \setminus \{(0,0)\}$, and so is f^{-1} by [10, Theorem 2.5.1] and (11).

Remark 3.4. The function $[0,\infty) \ni t \longmapsto \exp(-t)$ in (8) can be replaced by $[0,\infty) \ni t \longmapsto 1/(1+t)$; then our f will be even rational off the origin.

Remark 3.5. We believe that the constructions in Example 3.2 can be extended to higher dimensions, and also to the space $\ell_2 \times \mathbb{R}$. It is perhaps enough to replace " x^6 " by " $\|x\|^6$ " in (8), and " x^2 " by " $\|x\|^2$ " in (9), where $\|\cdot\|$ is the Hilbertian norm.

4. Counter-examples to the Chain Rule

Now we turn our attention to the Chain Rule statement. For the two mappings we choose a mnemonic notation: we use the symbol f_ι for the inner mapping, and f_o for the outer one.

Theorem 4.1. (See, e.g., [5, Corollary 3.2.13, page 123].) *Let X, Y, Z be Banach spaces, let $f_\iota : X \to Y$ be an (inner) mapping defined in a vicinity of an $x \in X$, and $f_o : Y \to Z$ be an (outer) mapping defined in a vicinity of $y := f_\iota(x)$. Assume that f_ι is Gateaux differentiable at x and f_o is Fréchet differentiable at y. Then the compound mapping $f_o \circ f_\iota : X \longrightarrow Z$ is (at least) Gateaux differentiable at x and*

$$(f_o \circ f_\iota)'(x) = f_o{}'(y) \circ f_\iota{}'(x).$$

This can be proved by a small adjustment of the proof of the well known chain rule theorem, say [11, pp. 337–338], where both adjectives 'Gateaux' are replaced by 'Fréchet'.

In Theorem 4.1, the adjective 'Fréchet' may be replaced by 'Hadamard'. This makes a difference only for infinite dimensional spaces.

It is known that Theorem 4.1 is not valid when the mapping f_o is Gateaux differentiable only. Simple counter-examples for this can be found in [1, Remark 1.5, p. 216] and [12, p. 12]. S. Yamamuro considers here the scheme $\mathbb{R} \xrightarrow{f_\iota} \mathbb{R}^2 \xrightarrow{f_o} \mathbb{R}$, where $f_\iota(t) := (t, t^2)$ for $t \in \mathbb{R}$ and

$$f_o(t, u) := \begin{cases} t & \text{if } \ u = t^2 \\ 0 & \text{if } \ u \neq t^2 \end{cases}$$

for $(t, u) \in \mathbb{R}^2$. Clearly, $f_o \circ f_\iota$ is the identity on $\mathbb{R}$, f_ι is (Fréchet) differentiable at 0, f_o is Gateaux differentiable at $(0, 0)$, and

$$Jf_o(0, 0)\, Jf_\iota(0) = (0\ 0) \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix} \right) = (0) \ \neq \ (1) = J(f_o \circ f_\iota)(0).$$

Below, we produce examples in the setting $\mathbb{R}^2 \xrightarrow{f_\iota} \mathbb{R}^2 \xrightarrow{f_o} \mathbb{R}^2$ where both mappings f_ι, f_o are, additionally, homeomorphisms of $\mathbb{R}^2$.

Example 4.2. There exist homeomorphisms $f_\iota : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ and $f_o : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, with $f_\iota(0, 0) = f_o(0, 0) = (0, 0)$, such that f_ι and $f_\iota{}^{-1}$ are quadratic polynomials, f_o and $f_o{}^{-1}$ are analytic on $\mathbb{R}^2 \setminus \{(0, 0)\}$, f_ι is Fréchet differentiable at $(0, 0)$, f_o is Gateaux differentiable at $(0, 0)$, with

$$Jf_\iota(0, 0) = Jf_o(0, 0) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix} \right),$$

and $(f_o \circ f_\iota)$ is not Gateaux differentiable at $(0, 0)$.

Proof. For $(x, y) \in \mathbb{R}^2$ we define

$$f_\iota(x, y) := (x, y + x^2) \quad \text{and} \quad f_o(x, y) := f_1(x, y - x^2),$$

where the mapping f_1 was defined in (8), that is,

$$f_1(x, y) := \begin{cases} \left(1 - \frac{1}{2} \exp\left(-\frac{y^2}{x^6 + y^6}\right)\right)(x, y) & \text{if } \ (x, y) \neq (0, 0), \\ (0, 0) & \text{if } \ (x, y) = (0, 0). \end{cases} \tag{12}$$

By (12), we immediately get that f_o is analytic on $\mathbb{R}^2 \setminus \{(0,0)\}$. In Section 3 (in the alternative construction for Example 3.2) we proved that $f_o\,(=f)$ is a homeomorphism on $\mathbb{R}^2$ and that ${f_o}^{-1}$ is also analytic on $\mathbb{R}^2 \setminus \{(0,0)\}$. We also verified that f_o is Gateaux differentiable at $(0,0)$, with $Jf_o(0,0)$ as claimed. (The f there is our f_o.) Further, we have

$$(f_o \circ f_\iota)(x,y) = f_o(x, y+x^2) = f_1(x,y)$$

for every $(x,y) \in \mathbb{R}^2$, and we proved, in Section 3, that f_1 is not Gateaux differentiable at $(0,0)$. $\square$

Example 4.3. There exist homeomorphisms $f_\iota : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ and $f_o : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, with $f_\iota(0,0) = f_o(0,0) = (0,0)$, such that f_ι is a polynomial, f_o and ${f_o}^{-1}$ are analytic on $\mathbb{R}^2 \setminus \{(0,0)\}$, f_ι and $(f_o \circ f_\iota)$ are Fréchet differentiable at $(0,0)$, f_o is Gateaux differentiable at $(0,0)$, with

$$Jf_\iota(0,0) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right) \text{ (singular)}, \;\; Jf_o(0,0) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \text{ and } J(f_o \circ f_\iota)(0,0) = \left(\begin{smallmatrix} 1/2 & 0 \\ 0 & 0 \end{smallmatrix}\right);$$

and hence the chain rule formula at $(0,0)$ is not valid.

Proof. For $(x,y) \in \mathbb{R}^2$ we define

$$f_\iota(x,y) := (x, x^2 + y^5) \quad \text{and (again)} \quad f_o(x,y) := f_1(x, y - x^2),$$

where f_1 is defined by (12). f_ι is a homeomorphism because it equals the composition of two homeomorphisms $(x,y) \mapsto (x, y^5)$ and $(u,v) \mapsto (u, u^2+v)$. Clearly, f_ι is a polynomial. The properties of f_o were checked in Example 4.2. Let us verify that $(f_o \circ f_\iota)$ is Gateaux differentiable at $(0,0)$. Take any $(a,b) \in \mathbb{R}^2 \setminus \{(0,0)\}$. For $0 \neq t \to 0$ we have

$$\begin{aligned}
\frac{1}{t}(f_o \circ f_\iota)(ta,tb) &= \frac{1}{t} f_o(f_\iota(ta,tb)) = \frac{1}{t} f_o(ta, t^2a^2 + t^5b^b) = \frac{1}{t} f_1(ta, t^5b^5) \\
&= \left(1 - \frac{1}{2}\exp\Big(-\frac{t^{10}b^{10}}{t^6a^6 + t^{30}b^{30}}\Big)\right)(a, t^4b^5) \\
&= \left(1 - \frac{1}{2}\exp\Big(-t^4\frac{b^{10}}{a^6 + t^{24}b^{30}}\Big)\right)(a, t^4b^5) \longrightarrow (\tfrac{1}{2}a, 0),
\end{aligned}$$

no matter if $a = 0$ or $a \neq 0$. Hence $(f_o \circ f_\iota)$ is Gateaux differentiable at $(0,0)$, with Jacobian $\left(\begin{smallmatrix} 1/2 & 0 \\ 0 & 0 \end{smallmatrix}\right)$. Therefore, the chain rule formula collapses.

It remains to check the Fréchet differentiability of $(f_o \circ f_\iota)$ at the origin. Fix for a while any $0 \neq t \in \mathbb{R}$ and any $(0,0) \neq (a,b) \in \mathbb{R}^2$. We already know that

$$\begin{aligned}
&\frac{1}{t}(f_o \circ f_\iota)(ta,tb) - (\frac{a}{2}, 0) \\
&= \left(\Big(1 - \exp\Big(-t^4\frac{b^{10}}{a^6 + t^{24}b^{30}}\Big)\Big)\frac{a}{2}, \Big(1 - \frac{1}{2}\exp\Big(-t^4\frac{b^{10}}{a^6 + t^{24}b^{30}}\Big)\Big)t^4b^5\right) \\
&=: (\xi_1(t,a,b), \xi_2(t,a,b)).
\end{aligned}$$

Hence $\quad \sup\big\{|\xi_2(t,a,b)| : \ (0,0) \neq (a,b) \in [-1,1]^2\big\} \leq t^4 \to 0 \quad \text{as} \quad t \to 0.$

As regards the first coordinate, fix any $0 < \varepsilon < 1$. Consider any $(a,b) \in \mathbb{R}^2$ with $(0,0) \neq (a,b) \in [-1,1]^2$. If $|a| < 2\varepsilon$, then $|\xi_1(t,a,b)| \leq \frac{|a|}{2} < \varepsilon$ for every $0 \neq t \in \mathbb{R}$. Further assume that $2\varepsilon \leq |a| \leq 1$. Then

$$|\xi_1(t,a,b)| \leq \Big(1 - \exp\Big(-\frac{t^4}{2^6\varepsilon^6}\Big)\Big)\frac{1}{2} \longrightarrow 0 \quad \text{as} \quad t \to 0,$$

and hence $$\limsup_{t\to 0} \sup \big\{|\xi_1(t,a,b)| : \ (0,0) \neq (a,b) \in [-1,1]^2\big\} \leq \varepsilon.$$

Here the ε could be arbitrarily small. Therefore

$$\sup \big\{|\xi_1(t,a,b)| : \ (0,0) \neq (a,b) \in [-1,1]^2\big\} \longrightarrow 0 \quad \text{as} \quad t \to 0$$

and so $(f_o \circ f_\iota)$ is Fréchet differentiable at $(0,0)$. □

Remark 4.4. We believe that Remarks 3.4 and 3.5, after some small adjustments, can be applied also to Examples 4.2 and 4.3.

Remark 4.5. It is not possible to construct the mappings f_ι, f_o similarly to Example 4.3 with Jacobian $Jf_\iota(0,0)$ regular, see Theorem 6.6.

The following table presents an overview of selected information pertaining to the Chain Rule. It also encompasses the material presented in the following sections. The properties of the mappings in the examples are abbreviated by G (Gateaux differentiability at the point) and F&G (Fréchet and hence automatically also Gateaux differentiability). With regard to the theorems, we distinguish between the assumptions (F, G) and the conclusions (e.g., "⇒G").

	f_o'	f_ι'	$(f_o \circ f_\iota)'$	Chain Rule formula	f_o and f_ι
Standard C.R.	F	F	⇒F&G	holds	
Theorem 4.1	F	G	⇒G	holds	
[1, Remark 1.5]	G	F&G	¬ G	-	
Example 4.2	G	F&G	¬ G	-	homeomorphisms
Example 5.1 (v)	G	F&G	¬ G	-	homeomorphisms
[12, page 12]	G	F&G	F&G	fails	
Example 4.3	G	F&G	F&G	fails	homeomorphisms
Example 5.1	G	G	F&G	fails	$f_o^{-1} = f_o = f_\iota$
Theorem 6.6	G	G	G	holds conditionally	

5. The main example

In Example 5.1 below, we construct a mapping more specific than what is needed to show that the formula for the derivative of the inverse mapping is not true and the chain rule formula grossly fails when the mappings are only Gateaux differentiable. The reason is that the ideas behind our construction are "symmetric" in the sense that we need to apply them both from the domain side and the range side (i.e., both to the mapping and to its inverse). Therefore, we choose to have an involutive mapping, which will cost us some technical difficulties but saves us from checking the properties of the inverse mapping. (Also, the intuitive understanding of how the inverse side of our ideas works seems to be more difficult. Thus, constructing an involutive mapping as an example might serve as a firm basis – for a hesitating mind.)

Example 5.1. *There exists a continuous mapping $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ such that*

(i) $f(0,0) = (0,0)$,

(ii) *f is an involution, that is, $f(f(x_1,x_2)) = (x_1,x_2)$ for every $(x_1,x_2) \in \mathbb{R}^2$,*

(iii) *f is C^∞-smooth, hence a C^∞-diffeomorphism, on $\mathbb{R}^2 \setminus \{(0,0)\}$,*

(iv) *f is Gateaux differentiable at $(0,0)$, with $Jf(0,0) = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$, and hence*

$$Jf(0,0)\, Jf(0,0) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = J(f \circ f)(0,0),$$

(v) *there is a polynomial homeomorphism $q : \mathbb{R}^2 \to \mathbb{R}^2$ such that $f \circ q$ is not Gateaux differentiable at $(0,0)$.*

Proof. We will proceed in several steps by constructing mappings having more and more properties until we get a mapping having all claimed properties.

Using our experience from constructing previous examples, we try to define for $(x_1,x_2) \in \mathbb{R}^2$

$$f_0(x_1,x_2) := \begin{cases} ({x_2}^9,\ x_1) & \text{if} \quad |x_1| > |x_2|^3, \\ ({x_2}^3,\ \sqrt[3]{x_1}) & \text{if} \quad |x_1| = |x_2|^3, \\ (x_2, \sqrt[9]{x_1}) & \text{if} \quad |x_1| < |x_2|^3. \end{cases} \tag{13}$$

Here and elsewhere in the paper, we use the convention that

$$\sqrt[r]{t} := -\sqrt[r]{-t}$$

when $t < 0$ and $r \in \mathbb{N}$ is an odd number.

A quick look at (13) reveals that $f_0(0,0) = (0,0)$ and that f_0 is continuous at the origin and also at every $(x_1,x_2) \in \mathbb{R}^2$ satisfying $|x_1| \neq |x_2|^3$.

Let us focus on the Gateaux differentiability of f_0 at $(0,0)$. For any $h_2 \in \mathbb{R}$, if $0 \neq t \in \mathbb{R}$ is sufficiently close to 0, we have $|t| > |th_2|^3$, and so

$$\frac{1}{t}(f_0(t,th_2) - f_0(0,0)) = \frac{1}{t}((th_2)^9, t) = (t^8 {h_2}^9, 1) \longrightarrow (0,1) \quad \text{as} \quad t \to 0;$$

we proved that $Df_0(0,0)(1,h_2) = (0,1)$. Further, for all $0 \neq t \in \mathbb{R}$ we have $\frac{1}{t} f_0(0,t) = \frac{1}{t}(t,0) = (1,0)$; thus $Df_0(0,0)(0,1) = (1,0)$. Then

$$Df_0(0,0)(1,1) = (0,1) \neq (1,1) = Df_0(0,0)(1,0) + Df_0(0,0)(0,1),$$

and so f_0 cannot be Gateaux differentiable at $(0,0)$.

But we look for a mapping which is Gateaux differentiable at $(0,0)$. This defect can be remedied by composing our f_0, from right, with a “breeze” mapping

$$b(x_1,x_2) := (x_1 - 8{x_2}^2, x_2), \quad (x_1,x_2) \in \mathbb{R}^2. \tag{14}$$

Clearly, b is a bijection on $\mathbb{R}^2$; it is even a C^∞-diffeomorphism (actually a “polynomial” diffeomorphism). Moreover, for every fixed $(0,0) \neq (h_1,h_2) \in \mathbb{R}^2$ we have $|th_1 - 8(th_2)^2| > |th_2|^3$ for all sufficiently small $0 \neq t \in \mathbb{R}$, and so

$$\begin{aligned}\lim_{t\to 0} \frac{1}{t}(f_0 \circ b)(th_1, th_2) &= \lim_{t\to 0} \frac{1}{t} f_0(th_1 - 8(th_2)^2, th_2) \\ &= \lim_{t\to 0} \frac{1}{t}((th_2)^9, th_1 - 8(th_2)^2) = (0, h_1);\end{aligned}$$

thus $D(f_0 \circ b)(0,0)(h_1,h_2) = (0,h_1)$. We proved that $f_0 \circ b$ is Gateaux differentiable at $(0,0)$ with the Jacobian $J(f_0 \circ b)(0,0) = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ (which is a singular matrix).

Concerning the smoothness of f_0 off the origin, we can immediately check that it is C^∞-smooth, even analytic, at each point $(x_1,x_2) \in \mathbb{R}^2$ for which $|x_1| \neq |x_2|^3$ and $x_1 \neq 0$. However, if $0 \neq x_2 \in \mathbb{R}$, then $Df_0(0,x_2)(1,0) = \infty$, see (13); we thus loose Gateaux differentiability at $(0,x_2)$ with $x_2 \neq 0$. And further, unfortunately, at the points of the “cubic parabola” $\{(x_1,x_2) \in \mathbb{R}^2 : \ |x_1| = |x_2|^3 \neq 0\}$, our f_0 is desperately non-smooth, actually non-continuous.

All the above properties of f_0 are then transferred to the analogous properties of the composed mapping $f_0 \circ b$.

The last (but not least) feature of f_0 is its involutivity. So fix any $(x_1,x_2) \in \mathbb{R}^2$. If $|x_1| = |x_2|^3$, then (13) immediately yields that $f_0(f_2(x_1,x_2)) = (x_1,x_2)$. Further assume that $|x_1| > |x_2|^3$. Denote $f_0(x_1,x_2) =: (y_1,y_2)$. A bit of calculation reveals that $|y_1| < |y_2|^3$. Thus using (13) twice we get that

$$f_0(f_0(x_1,x_2)) = f_0(y_1,y_2) = ({y_2}^9,\ y_1) = (x_1,x_2).$$

And, if $|x_1| < |x_2|^3$, we proceed similarly. We proved that f_0 is an involution.

However we are afraid that $f_0 \circ b$ is not involutive. But this can be remedied by putting

$$f := b^{-1} \circ f_0 \circ b \tag{15}$$

(where $b^{-1}(x_1,x_2) = (x_1 + 8{x_2}^2, x_2)$ for $(x_1,x_2) \in \mathbb{R}^2$) and realizing that composition from left with the (polynomial) mapping b^{-1} will not damage any of the properties of $f_0 \circ b$ proved above. Of course the (valid) chain rule theorem for the case “Fréchet”$\circ$“Gateaux” = “Gateaux” is used here; see [5, p. 123] quoted above.

Now, we can conclude that the mapping f defined by (15) is continuous at the origin and satisfies conditions (i), (ii), (iv), and (v), with $q := b^{-1}$, of the Example 5.1. (The validity of (iii) is not guaranteed yet.)

Our next plan will be to improve the mapping f_0 constructed in (13). More concretely, we look for another mapping which would be analytic also at the points of the axis $\{(0,t) : \ 0 \neq t \in \mathbb{R}\}$. We might observe that there is much freedom in the architecture of (13).

We use it to define a mapping $f_1 : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ for $(x_1, x_2) \in \mathbb{R}^2$ as follows:

$$f_1(x_1,x_2) := \begin{cases} (x_2\sqrt[3]{x_1{}^8},\ x_1) & \text{if} \quad |x_1| > |x_2|^3, \\ (x_2^3,\ \sqrt[3]{x_1}) & \text{if} \quad |x_1| = |x_2|^3, \\ (x_2, x_1/\sqrt[3]{x_2{}^8}) & \text{if} \quad |x_1| < |x_2|^3. \end{cases} \tag{16}$$

Clearly, f_1 is continuous at $(0,0)$. The verification that f_1 is involutive is again straightforward.

The reader may wonder how (13) might have inspired (16). We try to give an explanation here. In the definition of f_0, we see two domains that are separated by the set $\{(x_1,x_2) :\ |x_1| = |x_2|^3\}$; each of the two domains is a union of segments that are orthogonal to and symmetrical about the x_1- or x_2-axes. The image such a segment in f_0 is also a segment. The simplest idea how to remove the singularity at the midpoint of a segment is to replace the function in question by a linear function on the segment. Our wish is that the replacement of the mapping should not change the image of the segment; hence we have to keep the limit values of f_0 at the ends of the segment. This way, (13) transforms into (16). For example for the first line in (13), with fixed $x_1 = a$ (where $a \neq 0$), the variable x_2 runs through interval $I := (-\sqrt[3]{|a|}, \sqrt[3]{|a|})$; the limit values of f_0 at the endpoints of the segment $\{a\} \times I$ are $(-a^3, a) \in \mathbb{R}^2$ and $(a^3, a) \in \mathbb{R}^2$. The linearized function is thus $(a, x_2) \longmapsto (x_2\, a^3/\sqrt[3]{a}, a)$, $x_2 \in I$. Therefore we let $f_1(a, x_2) := (x_2\sqrt[3]{a^8}, a)$ for $x_2 \in I$. This is how we defined f_1 in the first line of (16).

Similarly, considering the restriction of f_0 to the segment $(-|a|, |a|) \times \{\sqrt[3]{a}\}$, the limit values at the endpoints of the segment will be $(\sqrt[3]{a}, \pm\sqrt[9]{a})$. Thus, for $x_1 \in (-|a|, |a|)$, we define

$$f_1(x_1, \sqrt[3]{a}) := \big(\sqrt[3]{a}, \tfrac{x_1}{a}\sqrt[9]{a}\big) = \big(\sqrt[3]{a}, x_1/\sqrt[9]{a^8}\big).$$

This leads us to define $f_1(x_1,x_2) := \big(x_2, x_1/\sqrt[3]{x_2{}^8}\big)$ whenever $(x_1,x_2) \in \mathbb{R}^2$ and $|x_2|^3 > |x_1|$, as in the third line of (16).

Let us focus on the Gateaux differentiability of f_1 defined by (16) at $(0,0)$. For any $h_2 \in \mathbb{R}$, if $0 \neq t \in \mathbb{R}$ is sufficiently close to 0, we have $|t| > |th_2|^3$, and so

$$\frac{1}{t} f_1(t, th_2) = \frac{1}{t}\big(th_2\sqrt[3]{t^8}), t\big) = \big(h_2\sqrt[3]{t^8}), 1\big) \longrightarrow (0,1) \quad \text{as} \quad t \to 0;$$

we proved that $Df_1(0,0)(1,h_2) = (0,1)$. Further, for all $0 \neq t \in \mathbb{R}$ we have $\frac{1}{t} f_1(0,t) = \frac{1}{t}(t,0) = (1,0)$; thus $Df_1(0,0)(0,1) = (1,0)$. Hence

$$Df_1(0,0)(1,1) = (0,1) \neq (1,1) = Df_1(0,0)(0,1) + Df_1(0,0)(1,0)$$

and so neither the f_1 from (16) can be Gateaux differentiable at the origin. Yet, the compound mapping $f_1 \circ b$ where b is from (14) `is` Gateaux differentiable at $(0,0)$. Indeed, for every fixed $(0,0) \neq (h_1,h_2) \in \mathbb{R}^2$ we have $|th_1 - 8(th_2)^2| \geq |th_2|^3$ for all sufficiently small $0 \neq t \in \mathbb{R}$, and so

$$\begin{aligned} \lim_{t\to 0} \frac{1}{t}(f_1 \circ b)(th_1, th_2) &= \lim_{t\to 0} \frac{1}{t} f_1\big(th_1 - 8(th_2)^2, th_2\big) \\ &= \lim_{t\to 0} \frac{1}{t}\big(th_2\sqrt[3]{(th_1 - 8(th_2)^2)^8}, th_1 - 8(th_2)^2\big) = (0, h_1); \end{aligned}$$

thus $D(f_1 \circ b)(0,0)(h_1,h_2) = (0,h_1)$.

We proved that $f_1 \circ b$ is Gateaux differentiable at $(0,0)$ with $J(f_1 \circ b)(0,0) = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$.

We believe (and a suspicious reader is invited to check all the details) that f_1 from (16) has all the properties of f_0 from (13) but one change: f_1 is analytic, hence C^∞-smooth, at every point off the "parabola" $\{(x_1,x_2) \in \mathbb{R}^2 : \ |x_1| = |x_2|^3\}$; in particular this is so at every point of the cross $\{(x_1,x_2) \in \mathbb{R}^2 : \ x_1x_2 = 0\} \setminus \{(0,0)\}$. Indeed, fix any $(x_1,x_2) \in \mathbb{R}^2$ outside the parabola. If $|x_1| > |x_2|^3$, then $x_1 \neq 0$ and $|x_1'| > |x_2'|^3$ for all $(x_1',x_2') \in \mathbb{R}^2$ from a neighborhood of (x_1,x_2), and so the definition formula (16) immediately reveals that f_1 is an analytic mapping at (x_1,x_2). The other case, when $|x_1| < |x_2|^3$, can be treated similarly.

Next, we will try to remedy the discontinuity of f_1 on the (meager) set

$$\{(x_1,x_2) \in \mathbb{R}^2 : \ |x_1| = |x_2|^3\}$$

by redefining it on the (thick) area $\{(x_1,x_2) \in \mathbb{R}^2 : \ \frac{1}{2}|x_2| < |x_1| \leq 8|x_2|^3\}$. After one or two attempts, we got a feeling that the following pattern could work; it will give us a mapping as announced in Example 5.1 (in particular, this new mapping is C^∞-smooth everywhere off the origin). For $(x_1,x_2) \in \mathbb{R}^2$, we put "finally"

$$f_2(x_1,x_2) := \begin{cases} \left(x_2\sqrt[3]{x_1{}^8},\ x_1\right) \quad \left(= f_1(x_1,x_2)\right) \quad \text{if} \quad |x_1| \geq 8\,|x_2|^3, \\[1ex] \left(x_2, x_1/\sqrt[3]{x_2{}^8}\right) \quad \left(= f_1(x_1,x_2)\right) \quad \text{if} \quad |x_1| < \dfrac{1}{2}\,|x_2|^3, \\[1ex] \Big(\operatorname{sgn}(x_2)|x_2|^{3-2(r(c)+s(c))}|x_1|^{8s(c)/3}, \operatorname{sgn}(x_1)|x_1|^{\frac{1}{3}+\frac{2}{3}(r(c)+s(c))}|x_2|^{-8r(c)/3}\Big) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{if} \ \ x_1x_2 \neq 0 \text{ and } c := |x_1|/|x_2|^3. \end{cases} \tag{17}$$

Here we require that the functions $r : [0,\infty) \longrightarrow \mathbb{R}$ and $s : [0,\infty) \longrightarrow \mathbb{R}$ are bounded, continuous, and C^∞-smooth on $(0,\infty)$. (Why do we use two functions r, s, instead of one? Because we did not succeed when considering just one.)

Of course, the three parts of the domain in (17) overlap. The mapping f_2 will be well defined, if we arrange that

$$r(c) = 1, \quad s(c) = 0 \quad \text{if} \quad c \in [0, \tfrac{1}{2}), \tag{18}$$

$$r(c) = 0, \quad s(c) = 1 \quad \text{if} \quad c \in [8, \infty), \tag{19}$$

and possibly that

$$r(c) = s(c) = 0, \qquad \text{if} \quad c \in [\tfrac{2}{3}, \tfrac{3}{2}]. \tag{20}$$

(Of course, we loose here the chance to obtain r and s as analytic functions.) Indeed, consider any $(x_1,x_2) \in \mathbb{R}^2$ with $x_1x_2 \neq 0$, and put $c := |x_1|/|x_2|^3$. If $|x_1| \geq 8|x_2|^3$, then (19) and the third line of (17) yield $f_2(x_1,x_2) = \left(x_2\sqrt[3]{x_1{}^8}, x_1\right)$. If $|x_1| < \frac{1}{2}|x_2|^3$, then (18) and the third line of (17) yield $f_2(x_1,x_2) = \left(x_2, x_1/\sqrt[3]{x_1{}^8}\right)$.

Next, on the following few pages, we impose further requirements on the functions r and s in order to guarantee that our f_2 defined by (17) is an involution. For any $t \in \mathbb{R}$, we have $f_2(0,t) = (t,0)$ and $f_2(t,0) = (0,t)$; thus $f_2(f_2(0,t)) = (0,t)$ and $f_2(f_2(0,t)) = (0,t)$. Now we will, for a longer while, consider any $(x_1,x_2) \in \mathbb{R}^2$ with $x_1x_2 \neq 0$. Let $c := |x_1|/|x_2|^3$.

Then $f_2(f_2(x_1,x_2)) = f_2(\tilde{x}_1,\tilde{x}_2)$, where, by the third line of (17),

$$\tilde{x}_1 := \operatorname{sgn}(x_2)|x_2|^{3-2(r(c)+s(c))}|x_1|^{8s(c)/3}$$

and
$$\tilde{x}_2 := \operatorname{sgn}(x_1)|x_1|^{\frac{1}{3}+\frac{2}{3}(r(c)+s(c))}|x_2|^{-8r(c)/3}. \tag{21}$$

(Note that $\tilde{x}_1\tilde{x}_2 \neq 0$.) For evaluating $f_2(\tilde{x}_1,\tilde{x}_2)$, (17) will be used again, with

$$\begin{aligned}\tilde{c} := |\tilde{x}_1| \,/\, |\tilde{x}_2|^3 &= |x_2|^{3-2(r(c)+s(c))+8r(c)}\,|x_1|^{8s(c)/3-1-2(r(c)+s(c))} \\ &= |x_2|^{3-2(r(c)+s(c))+8r(c)}\,|x_2|^{8s(c)-3-6(r(c)+s(c))}\,c^{8s(c)/3-1-2(r(c)+s(c))} \\ &= c^{2s(c)/3-2r(c)-1}\end{aligned} \tag{22}$$

in place of c. We see that $\tilde{c}$ does not depend on x_1 nor on x_2.

We feel that postulating (if possible)

$$r(\tilde{c}) = s(c) \quad \text{and} \quad s(\tilde{c}) = r(c) \quad \text{for every } c \in (0,\infty) \tag{23}$$

will help below. Assume so! Then we have

$$\begin{aligned}
&f_2(\tilde{x}_1,\tilde{x}_2) \overset{(17)}{=} \Big(\operatorname{sgn}(\tilde{x}_2)|\tilde{x}_2|^{3-2(r(\tilde{c})+s(\tilde{c}))}|\tilde{x}_1|^{8s(\tilde{c})/3}\,,\ \operatorname{sgn}(\tilde{x}_1)|\tilde{x}_1|^{\frac{1}{3}+\frac{2}{3}(r(\tilde{c})+s(\tilde{c}))}|\tilde{x}_2|^{-8r(\tilde{c})/3}\Big) \\
&\overset{(23)}{=} \Big(\operatorname{sgn}(\tilde{x}_2)|\tilde{x}_2|^{3-2(r(c)+s(c))}|\tilde{x}_1|^{8r(c)/3},\ \operatorname{sgn}(\tilde{x}_1)|\tilde{x}_1|^{\frac{1}{3}+\frac{2}{3}(r(c)+s(c))}|\tilde{x}_2|^{-8s(c)/3}\Big) \\
&\overset{(21)}{=} \Big(\operatorname{sgn}(x_1)\Big||x_1|^{\frac{1}{3}+\frac{2}{3}(r(c)+s(c))}|x_2|^{-8r(c)/3}\Big|^{3-2(r(c)+s(c))}\Big||x_2|^{3-2(r(c)+s(c))}|x_1|^{8s(c)/3}\Big|^{8r(c)/3}, \\
&\qquad \operatorname{sgn}(x_2)\Big||x_2|^{3-2(r(c)+s(c))}|x_1|^{8s(c)/3}\Big|^{\frac{1}{3}+\frac{2}{3}(r(c)+s(c))}\Big||x_1|^{\frac{1}{3}+\frac{2}{3}(r(c)+s(c))}|x_2|^{-8r(c)/3}\Big|^{-8s(c)/3}\Big) \\
&= \Big(\operatorname{sgn}(x_1)|x_1|^{(\frac{1}{3}+\frac{2}{3}(r(c)+s(c)))(3-2(r(c)+s(c)))}|x_1|^{64s(c)r(c)/9}, \\
&\qquad \operatorname{sgn}(x_2)|x_2|^{(3-2(r(c)+s(c)))\left(\frac{1}{3}+\frac{2}{3}(r(c)+s(c))\right)}|x_2|^{64r(c)s(c)/9}\Big) = (x_1,x_2),
\end{aligned}$$

where for the validity of the last equality above we need that, additionally,

$$\left(\tfrac{1}{3}+\tfrac{2}{3}(r(c)+s(c))\right)\left(3-2(r(c)+s(c))\right)+64s(c)r(c)/9 = 1 \ \text{ for every } \ c\in(0,\infty). \tag{24}$$

The above reasoning proves:

Lemma 5.2. *If the functions r, $s : [0,\infty) \longrightarrow \mathbb{R}$ satisfy* (24) *and* (23)*, with*

$$\tilde{c} := c^{2s(c)/3-2r(c)-1}, \quad c \in (0,\infty), \tag{25}$$

then the mapping f_2 defined by (17) *is an involution on $\mathbb{R}^2$.*

Now we want to find suitable candidates for the functions $r(\cdot)$, $s(\cdot)$. We strive to ensure (24) now, and later also (23). In order to make our life easier, we consider the transformation (rotation about the origin by the angle $\pi/2$)

$$S(c) = r(c) + s(c), \quad D(c) = r(c) - s(c), \quad c \in [0,\infty). \tag{26}$$

That is, we temporarily (until before Lemma 5.3) assume that

$$r(c) = \frac{S(c)+D(c)}{2}, \quad s(c) = \frac{S(c)-D(c)}{2}, \quad c \in [0,\infty) \tag{27}$$

where D and S are functions defined on $[0,\infty)$. Substituting (26) into (24) and performing a small rearrangement, we get the following requirement:

$$\tfrac{9}{4} + 3S(c) + S(c)^2 = 4D(c)^2 + \tfrac{9}{4}, \quad c \in [0,\infty). \tag{28}$$

When solving this equation, we prefer to take (from the two possibilities)

$$S(c) = \sqrt{4D(c)^2 + \tfrac{9}{4}} - \tfrac{3}{2}, \quad c \in [0,\infty). \tag{29}$$

(The other choice, $-\sqrt{\dots}$ is not suitable since $S(0) = r(0) + s(0) = 1$ is required by (18).) So we may state that if $D(\cdot)$ is any function on $[0,\infty)$ and $S(\cdot)$ is defined by (29), then (28) is satisfied. By (27), this guarantees (24).

It seems reasonable to have a function E with the same values as the exponent in (25) has:

$$E(c) = \tfrac{2}{3}s(c) - 2r(c) - 1, \quad c \in [0,\infty). \tag{30}$$

Substituting here (27) and (29), we have: If the functions r, s are defined by (27), where S is as in (29), then (30) is equivalent to

$$E(c) = -\tfrac{1}{3}\big(\sqrt{16D(c)^2+9} + 4D(c)\big) = G(D(c)), \quad c \in [0,\infty), \tag{31}$$

where $$G(z) := -\tfrac{1}{3}(\sqrt{9+16z^2} + 4z), \quad z \in \mathbb{R}.$$

We can easily verify that the derivative $G'(z)$ exists and is negative for every $z \in \mathbb{R}$; thus G is invertible. Also, by calculating the limits at ∞ and $-\infty$, we conclude that G maps $\mathbb{R}$ onto $(-\infty,0)$; thus G has an inverse $G^{-1} : (-\infty,0) \longrightarrow \mathbb{R}$. A direct calculation reveals that

$$G^{-1}(u) = \frac{3}{8}\Big(\frac{1}{u} - u\Big), \quad u \in (-\infty,0).$$

If E is as in (31), we apply G^{-1} to both sides of equality $E = G \circ D$ to get that $D = G^{-1} \circ E$, which means

$$D(c) = \frac{3}{8}\Big(\frac{1}{E(c)} - E(c)\Big), \quad c \in [0,\infty). \tag{32}$$

We thus have that, if D is defined by (32), then (31) holds. Hence, if the functions r, s are defined by (27) and (29), then both (31), (32) are equivalent to (30).

Inserting (32) into (29) and realizing that $E(\cdot)$ is everywhere negative by (31), a routine gymnastics yields that

$$S(c) = -\frac{3(E(c)+1)^2}{4E(c)}, \quad c \in [0,\infty). \tag{33}$$

Now, substituting (32) and (33) into (27), we get that

$$r(c) = -\frac{9E(c)^2 + 12E(c) + 3}{16E(c)} \quad \text{and} \quad s(c) = -\frac{3E(c)^2 + 12E(c) + 9}{16E(c)}, \quad c \in [0,\infty). \tag{34}$$

We finally arrived at relations (34) that ensure (24), which is one of our two goals. (We did not consider (23) yet.) More formally:

Lemma 5.3. *Assume that* $E:[0,\infty)\to(-\infty,0)$ *is any function and that the functions* $r,s:[0,\infty)\to\mathbb{R}$ *are defined by* (34)*. Then* E *satisfies* (30) *and* (24) *holds true.*

Proof. If we define D by (32) and S by (29), then, as we explained above, S satisfies (33) and r, s satisfy (27); additionally, we explained above that E satisfies (30) and that (24) holds true. □

Now it remains (for the proof of involutivity of our f_2) to find a suitable function $E:[0,\infty)\longrightarrow(-\infty,0)$, possibly C^∞-smooth on $(0,\infty)$, and such that

$$r(c^{E(c)})=s(c),\quad\text{and}\quad s(c^{E(c)})=r(c)\ \text{ for all }\ c\in[0,\infty),\tag{35}$$

(see (23) and (22)). By a direct calculation from (34) (or more easily from (32), (29) and (27)) we see that if

$$E(c)=\begin{cases}-3 & \text{for }\ c\in[0,\frac12],\\ -1 & \text{for }\ c\in[\frac23,\frac32],\\ -\frac13 & \text{for }\ c\in[8,\infty),\end{cases}\tag{36}$$

then the relations (18), (19) and (20) will be satisfied for our r and s from (34).

It will be versatile to put (note that the domain of T is smaller than that of E)

$$T(c):=c^{E(c)},\quad c\in(0,1].\tag{37}$$

Thus, $T(c)=1/c^3$ for $c\in(0,\frac12]$ and $T(c)=1/c$ for $c\in[\frac23,1]$. It would be great if T could be strictly decreasing, since, if so, then T would map the interval $(0,1]$ onto the interval $[1,\infty)$ injectively. If r, s are as in (34), then, for every $c\in[0,1]$, we have

$$r(T(c))\overset{(34)}{=}-\frac{9E(T(c))^2+12E(T(c))+3}{16E(T(c))},\qquad -\frac{-9/E(c)^2+12/E(c)+3}{16/E(c)}\overset{(34)}{=}s(c)\tag{38}$$

and

$$s(T(c))\overset{(34)}{=}-\frac{3E(T(c))^2+12E(T(c))+9}{16E(T(c))},\qquad -\frac{-3/E(c)^2+12/E(c)+9}{16/E(c)}\overset{(34)}{=}r(c)\tag{39}$$

Here we recommend the reader to stop and observe (38) and (39) for a while. In order to guarantee the validity of (35) (and thus getting the involutivity of our f_2), it is enough to find $E(\cdot)$ which, besides (36), satisfies the function equation:

$$E(T(c))=\frac{1}{E(c)},\quad\text{that is,}\quad E\big(c^{E(c)}\big)=\frac{1}{E(c)}\,,\quad c\in(0,1].\tag{40}$$

In what follows we present a logically rigorous construction of the mapping f_2 whose definition we started in (17). (It remains to define the functions r and s.) We will subsequently construct the function $T:(0,1]\longrightarrow[1,\infty)$, then the function $E:[0,\infty)\longrightarrow[-3,-\frac13]$, and finally the functions $r,s:[0,\infty)\longrightarrow\mathbb{R}$ (we will not use the functions S,D,G). We define

$$T(c):=\frac{1}{c^3}\xi(c)+\frac{1}{c}(1-\xi(c)),\quad c\in(0,1],\tag{41}$$

where $\xi:(0,1]\longrightarrow[0,1]$ is a non-increasing C^∞-smooth function with $\xi(c)=1$ for all $c\in(0,\frac12]$ and $\xi(c)=0$ for all $c\in[\frac23,1]$.

For how to find such a function ξ, see e.g. [11, page 168]. Clearly, $T(c) = 1/c^3$ for $c \in (0, \frac{1}{2}]$ and $T(c) = 1/c$ for $c \in [\frac{2}{3}, 1]$. Further, for all $c \in (0,1)$ we have $T'(c) = t_1(c) + t_2(c) < 0$ where $t_1(c) := -3(1/c^4)\xi(c) - (1/c^2)(1 - \xi(c)) < 0$ and $t_2(c) := (1/c^3 - 1/c)\xi'(c) \le 0$. Since (41) represents a convex combination, we have $1/c \le T(c) \le 1/c^3$ for all $c \in (0,1]$. We know that T has negative derivative on all of $(0,1)$, and moreover that the limit of $T(c)$ for $c \downarrow 0$ is ∞ and $T(1) = 1$. Therefore our T maps $(0,1]$ injectively onto $[1,\infty)$.

Now we define the function $E(\cdot)$ as follows:

$$E(0) := -3, \quad E(1) := -1, \tag{42}$$

$$\begin{aligned} & E(c) := \frac{\ln T(c)}{\ln c} && \text{if } c \in (0,1), \\ \text{and} \quad & E(d) := \frac{1}{E(T^{-1}(d)))} \left(= \frac{\ln T^{-1}(d)}{\ln d} \right) && \text{if } d \in (1,\infty). \end{aligned} \tag{43}$$

Thus (40) is satisfied. Also, clearly, $E(c) = -3$ for all $c \in [0, \frac{1}{2}]$; thus our E is continuous at 0. Furthermore, $E(c) = -1$ for all $c \in [\frac{2}{3}, \frac{3}{2}]$. And finally, $E(d) = -1/3$ if $d \in [8,\infty)$. Therefore the function $E(\cdot)$ is C^∞-smooth on all of $(0,\infty)$. Also, recalling that for all $c \in (0,1]$ we have $1/c \le T(c) \le 1/c^3$, we get that

$$-3 \le E(c) \le -\tfrac{1}{3} \quad \text{for every} \quad c \in [0,\infty). \tag{44}$$

Now we are ready to define $r, s : [0,\infty) \longrightarrow \mathbb{R}$ via $E(\cdot)$ by (34). They are clearly C^∞-smooth and (18), (19), and (20) hold. The verification of (35) follows. For $c \in (0,1)$ we have

$$c^{E(c)} = e^{\ln(c)\,E(c)} \overset{(43)}{=} e^{\ln(c)\ln(T(c))/\ln(c)} = e^{\ln(T(c))} = T(c),$$

and so $r(c^{E(c)}) = r(T(c))$, and for $d \in (1,\infty)$ we have

$$d^{\,E(d)} = e^{\ln(d)\,E(d)} \overset{(43)}{=} e^{\ln(d)\;\ln(T^{-1}(d))/\ln(d)} = e^{\ln(T^{-1}(d))} = T^{-1}(d),$$

and so $r(d^{\,E(d)}) = r(T^{-1}(d)$. Thus, for all $c \in (0,1)$ we have

$$r(c^{E(c)}) = r(T(c)) \overset{(34)}{=} -\frac{9E(T(c))^2+12E(T(c))+3}{16E(T(c))} \overset{(40)}{=} -\frac{9/E(c)^2+12/E(c)+3}{16/E(c)} \overset{(34)}{=} s(c),$$

and similarly we get that $r(d^{E(d)}) = s(d)$ for all $d \in (1,\infty)$. We proved the first identity in (35) for all $c \in (0,\infty)$. The second identity in (35) can be obtained analogously. Note that (35) is the same as (23) with (22). By Lemma 5.3, we have (24). Then, by Lemma 5.2 we obtain that f_2 is an involution.

We are a bit concerned about the continuity of f_2 at the origin. So, consider any $(0,0) \ne (x_1,x_2) \in \mathbb{R}^2$ (converging to $(0,0)$). If $|x_1| \ge 8|x_2|^3$ or $|x_1| < \frac{1}{2}|x_2|^3$, then $f_2(x_1,x_2) = f_1(x_1,x_2)$ by (17) and (16), and we can use the continuity of f_1 at $(0,0)$. Assume further that $c := |x_1|/|x_2|^3 \in [\frac{1}{2}, 8)$. Then

$$f_2(x_1,x_2) = \Big(\operatorname{sgn}(x_2)|x_2|^{A(c)}c^{8s(c)/3},\ \operatorname{sgn}(x_1)c^{\frac{1}{3}+\frac{2}{3}(r(c)+s(c))}|x_2|^{A(c)/3}\Big), \tag{45}$$

where $A(c) := 3 - 2(r(c)+s(c)) + 8s(c) = 3 - 2r(c) + 6s(c)$.

By (34) and (44), we obtain that $A(c) = -3/E(c)$, and so $A(c) \in [1, 9]$ for every $c \in [0, \infty)$. Let

$$K := \max_{c\in[1/2,8]} \max\left\{c^{8s(c)/3}, c^{\frac{1}{3}+\frac{2}{3}(r(c)+s(c))}\right\}.$$

This is a finite number since both the functions r and s are continuous on $[\frac{1}{2}, 8]$ by (34) and (44). Thus, for every $(x_1, x_2) \in \mathbb{R}^2$ satisfying $|x_1|/|x_2|^3 \in [\frac{1}{2}, 8)$, we have by (45)

$$\|f_2(x_1, x_2)\| \le K \max_{c\in[\frac{1}{2},8]} \left\{|x_2|^{A(c)}, |x_2|^{A(c)/3}\right\} \le K \max\left\{|x_2|^{1/3}, |x_2|^9\right\},$$

(where $\|\cdot\|$ means the maximum norm on $\mathbb{R}^2$) and thus the continuity of f_2 at the origin follows.

However, f_2 is not Gateaux differentiable at $(0, 0)$ (the proof is almost the same as for f_1). This defect can be remedied (as usual) by composing f_2, from right, with the breeze mapping b from (14), that is,

$$b(x_1, x_2) := (x_1 - 8{x_2}^2, x_2), \quad (x_1, x_2) \in \mathbb{R}^2.$$

We know that b is a bijection on $\mathbb{R}^2$; it is even a C^∞-diffeomorphism (actually a "polynomial" diffeomorphism). A simple picture or a direct calculation reveals that the compound mapping $f_2 \circ b$ `is` Gateaux differentiable at $(0, 0)$. Indeed, for every fixed $(0, 0) \neq (h_1, h_2) \in \mathbb{R}^2$ we have $|th_1 - 8(th_2)^2| \ge 8|th_2|^3$ for all sufficiently small $0 \neq t \in \mathbb{R}$, and so, using (17), (19),

$$\begin{aligned}\lim_{t\to 0} \frac{1}{t}(f_2 \circ b)(th_1, th_2) &= \lim_{t\to 0} \frac{1}{t} f_1\big(th_1 - 8(th_2)^2, th_2\big)\\ &= \lim_{t\to 0} \frac{1}{t}\big(th_2 \sqrt[3]{(th_1 - 8(th_2)^2)^8}, th_1 - 8(th_2)^2\big) = (0, h_1);\end{aligned}$$

thus $D(f_2 \circ b)(0, 0)(h_1, h_2) = (0, h_1)$. We proved that $f_2 \circ b$ is Gateaux differentiable at $(0, 0)$ with $J(f_2 \circ b)(0, 0) = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ (which is a singular matrix).

But the mapping $f_2 \circ b$ is not an involution. However, this can be easily cured by the further composition (by b^{-1} from left)

$$f := b^{-1} \circ f_2 \circ b \tag{46}$$

(and this does not at all damage nor improves the Gateaux differentiability at $(0, 0)$). This final mapping f for sure will satisfy all the properties of Example 5.1 ((v) with $q =: b^{-1}$).

The C^∞-smoothness of f_2 on $\mathbb{R}^2 \setminus \{(0, 0)\}$ follows by the very definition (17) via the C^∞-smoothness of r and s on $(0, \infty)$. In (46), f is defined as the composition of f_2 and polynomial mappings b, b^{-1}, and the preimage in b of the origin is just the origin. Hence f is C^∞-smooth on $\mathbb{R}^2 \setminus \{(0, 0)\}$. □

Remark 5.4. Is it possible to construct f in Example 5.1 analytic off the origin?

Remark 5.5. However, in $\mathbb{R}^n$ with $n \in \mathbb{N}$ even, we can generalize (17) by putting for any $x_1, x_2 \in \mathbb{R}^{n/2}$

$$f_2(x_1,x_2) := \begin{cases} \big(\|x_1\|^{8/3}\, x_2,\ x_1\big) & \text{if } \ \|x_1\| \geq 8\|x_2\|^3, \\ \big(x_2, x_1/\|x_2\|^{8/3}\big) & \text{if } \ \|x_1\| < \frac{1}{2}\|x_2\|^3, \\ \Big(\|x_2\|^{2-2(r(c)+s(c))}\, \|x_1\|^{8s(c)/3}\, x_2\ ,\|x_1\|^{-\frac{2}{3}+\frac{2}{3}(r(c)+s(c))}\, \|x_2\|^{-8r(c)/3}\, x_1\Big) & \\ & \text{if } \ \|x_1\|\cdot\|x_2\| \neq 0 \text{ and } c := \|x_1\|/\|x_2\|^3. \end{cases} \tag{47}$$

An adjustment of (14) is also needed, say,

$$b(x_1,x_2) := (x_1 - 8\|x_2\|^2 w, x_2), \quad x_1, x_2 \in \mathbb{R}^{n/2}, \tag{48}$$

where w is a fixed norm-one vector in $\mathbb{R}^{n/2}$. This way we get Example 5.1 in Euclidean spaces of any even dimension. And perhaps also in 'squares' $X \times X$, where X is any normed space admitting a smooth norm; for instance in any Hilbert space.

Remark 5.6. It is not clear to us how to construct a variant of Example 5.1 for the space $\mathbb{R}^3$.

Remark 5.7. In some of the examples above an easy 'generalization' to spaces $\mathbb{R}^2 \times Z$, where Z is any Banach space, may be done by replacing the mapping f by $\bar{f}((x,y),z) := (f(x,y),z)$. While it is possible to do so, the new mapping is smooth on $(\mathbb{R}^2 \setminus \{(0,0)\}) \times Z$ but not on $(\mathbb{R}^2 \times Z) \setminus \{0\}$.

6. A positive chain rule statement for the Gateaux derivatives

We will need some more concepts. Let X, Y be normed linear spaces, $f : X \to Y$ a mapping, defined in a vicinity of an $x \in X$, let $y \in Y$ and $u \in X$. Then the vector y is called the (*one-sided*) *derivative of f at x in the direction u* if

$$\lim_{t \downarrow 0} \frac{f(x+tu) - f(x)}{t} = y.$$

This is equivalent to saying that, for every sequence $(0,\infty) \ni t_n \to 0$, we have

$$\lim_{n\to\infty} \frac{f(x+t_n u) - f(x)}{t_n} = y. \tag{49}$$

Below, we omit the adjective 'one-sided'. Clearly, if the mapping f is Gateaux differentiable at x, then $f'(x)u$ is the derivative of f at x in the direction u for every $u \in X$.

We will also need some 'Hadamard'-like concepts:

Definition 6.1. Let X, Y be normed linear spaces, $f : X \to Y$ a mapping, defined in a vicinity of an $x \in X$, and let $y \in Y$, $u \in X$ be given. Then the vector y is called the *Hadamard derivative of f at x in the direction u* if

$$\lim_{n\to\infty} \frac{f(x+t_n u_n) - f(x)}{t_n} = y$$

whenever $X \ni u_n \to u$ and $(0,\infty) \ni t_n \to 0$. Having X, Y, f, and x as above, we say that f *is Hadamard differentiable at* x if f is Gateaux differentiable at x and moreover for every $u \in X$ $f'(x)u$ is the Hadamard derivative of f at x in the direction u. □

The following statement says that Hadamard derivatives appear generically for continuous mappings with continuous directional derivatives. An inspiration for the proof of this can be found in [7], and in the references therein. For other results on generic relations between the Gateaux and Hadamard derivative we refer to [13, Proposition 3.8, 3.9].

Proposition 6.2. *Let X be a Banach space, Y a normed linear space, and $f : X \to Y$ a continuous mapping defined in a vicinity of an $x \in X$. Assume that there is a continuous mapping $L : X \to Y$ such that, for every $u \in X$, $L(u)$ is the derivative of f at x in the direction u.*

Then there exists a dense G_δ subset $\Omega \subset X$ such that, for every $u \in \Omega$, $L(u)$ is the Hadamard derivative of f at x in the direction u.

Remark 6.3. Note that if f is continuous and Gateaux differentiable at $x \in X$, then f and $L := f'(x)$ satisfy the assumptions of Proposition 6.2.

Remark 6.4. For the proof of the proposition, the continuity of f is not required at the very point x. For this, (Baire space) $X \setminus \{0\}$ is used in place of X everywhere.

Proof of Proposition 6.2. For $m, k \in \mathbb{N}$, let

$$D_{m,k} := \bigcap_{t \in (0,1/k)} \left\{ u \in X : \left\| \frac{f(x+tu) - f(x)}{t} - L(u) \right\| \le \frac{1}{m} \right\}.$$

Clearly, this is a closed subset of X. For every $m \in \mathbb{N}$, we have $\bigcup_k D_{m,k} = X$. As X is a complete metric space, Baire's theorem guarantees that $\Omega_m := \bigcup_k \operatorname{int} D_{m,k}$ is an (open) dense subset of X. Using the Baire's theorem again, we have that $\Omega := \bigcap_m \Omega_m$ is a dense (and G_δ) set.

Fix any $u \in \Omega$. It remains to prove that $L(u)$ is the Hadamard derivative of f at x in the direction u. To this end, fix an $\varepsilon > 0$ and consider $t_n \in (0,\infty)$, $u_n \in X$ with $t_n \to 0$ and $u_n \to u$. We need to show that

$$\left\| \frac{f(x+t_n u_n) - f(x)}{t_n} - L(u) \right\| < \varepsilon \tag{50}$$

for all $n \in \mathbb{N}$ large enough. So, pick $m \in \mathbb{N}$ with $1/m < \varepsilon/2$. As $u \in \Omega_m$, we have $u \in \operatorname{int} D_{m,k}$ for some $k \in \mathbb{N}$. Hence

$$u_n \in \operatorname{int} D_{m,k} \qquad \text{and} \qquad \|L(u_n) - L(u)\| < \varepsilon/2. \tag{51}$$

for every $n \in \mathbb{N}$ large enough. Then we obtain (50) by (51) and by the definition of $D_{m,k}$. □

As in Section 4, for inner mappings we use the symbol f_ι and for outer mappings we use f_o.

Proposition 6.5. *Let X, Y, Z be normed linear spaces, let $f_\iota : X \to Y$ be a mapping defined in a vicinity of an $x \in X$, let $f_o : Y \to Z$ be a mapping defined in a vicinity of $f_\iota(x) \in Y$, and let $u \in X$, $v \in Y$, $w \in Z$ be given. Assume that v is the derivative of f_ι at x in the direction u and that w is the Hadamard derivative of f_o at the point $f_\iota(x)$ in the direction v.*

Then w is the derivative of the compound mapping $X \ni t \mapsto f_o(f_\iota(t)) \in Z$ at x in the direction u.

Proof. The result is standard. Consider an arbitrary sequence of numbers $t_n > 0$ with $t_n \to 0$. By (49) applied to $f := f_\iota$, we have that

$$u_n := (f_\iota(x + t_n u) - f_\iota(x))/t_n \to v.$$

As w is the Hadamard derivative of f_o at the point $f_\iota(x)$ in the direction of v, we obtain, by the definition of u_n and from Definition 6.1, that

$$\lim_{n\to\infty} \frac{f_o(f_\iota(x + t_n u)) - f_o(f_\iota(x))}{t_n} = \lim_{n\to\infty} \frac{f_o(f_\iota(x) + t_n u_n) - f_o(f_\iota(x))}{t_n} = w.$$

Now, (49) applied for $f := f_o \circ f_\iota$ concludes the argument. □

Theorem 6.6. (Chain rule formula for Gateaux differentiability) *Let X, Z be normed linear spaces, Y a Banach space, $f_\iota : X \to Y$ a mapping defined in a vicinity of an $x \in X$, and $f_o : Y \to Z$ a continuous mapping defined in a vicinity of $f_\iota(x) \in Y$. Assume that f_ι and $f_o \circ f_\iota$ are Gateaux differentiable at x and f_o is Gateaux differentiable at $f_\iota(x)$. Assume further that the derivative $f_\iota'(x) : X \longrightarrow Y$ is an open mapping. Then $(f_o \circ f_\iota)'(x) = f_o'(f_\iota(x)) \circ f_\iota'(x)$.*

Proof. Proposition 6.2 applied to $f := f_o$ and $L := f_o'(f_\iota(x))$ provides a dense subset Ω of Y such that, for every $v \in \Omega$, ${f_o}'(f_\iota(x))v$ is the Hadamard derivative of f_0 at $f_\iota(x)$ in the direction v.

As $f_\iota'(x) : X \to Y$ is an open mapping from X (on)to Y, the set $\Omega_1 := f_\iota'(x)^{-1}(\Omega)$ is dense in X; the verification of this is pretty simple.

Now, fix any $u \in \Omega_1$. We know that $f_\iota'(x)u$ $(= v)$ is the derivative of f_ι at x in the direction u and that $f_\iota'(x)u \in \Omega$. Hence $f_o'(f_\iota(x))\big(f_\iota'(x)u\big)$ $(= w)$ is the Hadamard derivative of f_0 at $f_\iota(x)$. By Proposition 6.5, $f_o'(f_\iota(x))(f_\iota'(x)(u))$ is the derivative of $(f_o \circ f_\iota)$ at x in the direction u, that is, $f_o'(f_\iota(x))(f_\iota'(x)u) = (f_o \circ f_\iota)'(u)$. As this holds for every u from the (dense) set Ω_1, and all the three Gateaux derivatives are continuous mappings, we have $f_o'(f_\iota(x)) \circ f_\iota'(x) = (f_o \circ f_\iota)(x)$. □

Remark 6.7. Example 5.1 shows that the openness of the derivative $f_i'(x)$ in Theorem 6.6 cannot be dropped.

Corollary 6.8. *Let X be a Banach space and $f : X \longrightarrow X$ an involution defined and continuous in a vicinity of an $x \in X$. Assume that f is Gateaux differentiable at x, with the derivative $f'(x)$ surjective.*

Then $f'(x) \circ f'(x)$ is the identity mapping, and hence $f'(x)$ is a square root of the identity.

Corollary 6.9. *In the setting of Theorem* 6.6, *assume moreover that $Z := X$ and that f_ι is a right inverse of f_o, that is, $f_o(f_\iota(x)) = x$ for every $x \in X$.*

Then $f_o'(f_\iota(x)) \circ f_\iota'(x)$ is the identity mapping on X.

Corollary 6.10. *In the setting of Corollary 6.9, assume moreover that both X, Y are Euclidean spaces.*

Then $\dim Y \le \dim X$ and $Jf_o(f_\iota(x))Jf_\iota(x)$ is the identity matrix in X.

Remark 6.11. For a better orientation of a reader we should perhaps recall the following well known facts:

(1) Let X, Y be normed linear spaces and $T : X \longrightarrow Y$ a linear continuous mapping. If T is open, then T is surjective.

(2) If both the spaces X, Y are Banach, then the reverse implication holds; see e.g. [11, Theorem XV.1.3, p. 388].

(3) If A is a matrix with n rows and m columns, then the following assertions are mutually equivalent: (i) the mapping $\mathbb{R}^n \ni h \longmapsto Ah^{\mathrm{T}} \in \mathbb{R}^m$ is open, (ii) this mapping is surjective, (iii) $m \le n$ and the rank of A is m.

7. Appendix

The purpose of this section is to prove that the mapping we defined in (7) satisfies the important properties required by Example 3.2.

We recall that η is as in (3) and we consider a C^∞-smooth function $\xi : \mathbb{R} \longrightarrow [0,1]$ such that $\xi(0) := 0$ and $\xi(t) := 1$ for $t \in (-\infty, -1] \cup [1, \infty)$. We defined

$$\varphi(x,y) := (1 - \xi(x/y^2))\eta(y) + \xi(x/y^2)y\big), \quad \text{when } (x,y) \in \mathbb{R}^2 \setminus (\mathbb{R} \times \{0\})$$

and $\varphi(x, 0) := 0$ for all $x \in \mathbb{R}$ and then

$$\widehat{f_1}(x,y) := \big(x\varphi(x,y)^2/y^2, \varphi(x,y)\big), \quad \text{if } (x,y) \in \mathbb{R}^2 \setminus (\mathbb{R} \times \{0\}),$$

and $\widehat{f_1}(x, 0) := (0, 0)$ if $x \in \mathbb{R}$. As $\xi(t) = 1$ whenever $|t| \ge 1$, it is not difficult to verify that φ is C^∞-smooth on $\mathbb{R}^2 \setminus \{(0,0)\}$ and $\widehat{f_1}$ equals the identity on

$$\Omega := \{(x,y) \in \mathbb{R}^2 : \ x \ge y^2 \text{ or } x \le -y^2\}.$$

Finally, we defined, in (7),

$$f := \widehat{f_1} \circ b, \quad \text{that is,} \quad f(x,y) := \widehat{f_1}(b(x,y))), \quad (x,y) \in \mathbb{R}^2, \tag{52}$$

where the C^∞-diffeomorphism $b : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ was defined in (5).

Let us check the continuity of $\widehat{f_1}$ at the origin. From (3) we easily deduce that φ is continuous at $(0,0)$. Moreover, for any $(x,y) \in \mathbb{R}^2$, with $y \ne 0$, we have

$$\big|x\varphi(x,y)^2/y^2\big| \quad \begin{cases} \le \varphi(x,y)^2 & \text{if } \big|x/y^2\big| \le 1, \\ = xy^2/y^2 & \text{if } \big|x/y^2\big| > 1. \end{cases}$$

Thus $\widehat{f_1}(x,y) \longrightarrow (0,0)$ as $\mathbb{R}^2 \ni (x,y) \longrightarrow (0,0)$.

We can further observe (a wonderful thing) that for every $c \in \mathbb{R}$ the mapping $\widehat{f_1}$ sends the points of the parabola K_c $(= \{(cy^2, y) :\ y \in \mathbb{R}\})$ into K_c. Actually it does so bijectively since the assignment $\mathbb{R} \ni y \longmapsto (1 - \xi(c))\eta(y) + \xi(c)y \in \mathbb{R}$ is strictly monotone and maps $\mathbb{R}$ onto $\mathbb{R}$. Great! Now, since $\bigcup_{c\in\mathbb{R}} K_c = \mathbb{R}^2 \setminus (\mathbb{R} \times \{0\})$, and $K_c \cap K_d$=$\{(0,0)\}$ whenever $c, d \in \mathbb{R}$ are distinct, we can conclude that our $\widehat{f_1}$ is a bijection of $\mathbb{R}^2$ onto $\mathbb{R}^2$. Thus, the continuity of $\widehat{f_1}$ and the invariance of domain theorem guarantee that the inverse $\widehat{f_1}^{-1}$ is everywhere continuous.

From the very definitions of φ and $\widehat{f_1}$, we can immediately derive that these mappings are C^∞-smooth on $\mathbb{R}^2 \setminus \{(0, 0)\}$.

What about the smoothness of $\widehat{f_1}^{-1}$ on $\mathbb{R}^2\setminus\{(0, 0)\}$? On the set $\{(x, y)\in\mathbb{R}^2 : y \neq 0\}$ $\widehat{f_1}$ is the composition of the following three mappings:

$$
\begin{gathered}
(x, y) \longmapsto (x/y^2, y) =: (c, y), \\
(c, y) \longmapsto (c, (1 - \xi(c)) \cdot \eta(y) + \xi(c) \cdot y) =: (c, \phi), \qquad (53) \\
(c, \phi) \longmapsto (c\phi^2, \phi),
\end{gathered}
$$

whose Fréchet derivatives are represented by the Jacobian matrices

$$
\begin{gathered}
\begin{pmatrix} 1/y^2 & -2x/y^3 \\ 0 & 1 \end{pmatrix} \\
\begin{pmatrix} 1 & 0 \\ -\xi'(c) \cdot \eta(y) + \xi'(c) \cdot y & (1 - \xi(c)) \cdot \eta'(y) + \xi(c) \end{pmatrix} \\
\begin{pmatrix} \phi^2 & 2c\phi \\ 0 & 1 \end{pmatrix},
\end{gathered}
$$

respectively; the matrices are regular provided $1/y^2 \neq 0$, $(1-\xi(c))\cdot\eta'(y)+\xi(c) \neq 0$, and $\phi^2 \neq 0$ (where we let $c := x/y^2$ and ϕ is as in (53)). The first condition is obviously satisfied. The ϕ (see (53)) is a convex combination of values $\eta(y)$ and y which have the same sign, so $\phi \neq 0$. The remaining value, $(1 - \xi(c)) \cdot \eta'(y) + \xi(c)$, is a convex combination of $\eta'(y)$ and 1, hence positive provided $\eta'(y) > 0$. We note that all this is satisfied for each of our three options of η and for every $0 \neq y \in \mathbb{R}$.

Thus the Jacobian matrix $J\widehat{f_1}$ is continuous at each point of $\{(x, y) \in \mathbb{R}^2 :\ y \neq 0\}$ and is regular because it equals the product of the three regular matrices above. Hence the inverse mapping $\widehat{f_1}^{-1}$ is also C^∞-smooth away from $\{(x, y) \in \mathbb{R}^2 :\ y = 0\}$; see [2, Theorems 4.2.1 and 5.4.4] or [11, Chapter IV, §1, Theorem 1.2, p. 361]. And since $\widehat{f_1}$ equals the identity on the set Ω, our $\widehat{f_1}^{-1}$ is C^∞-smooth on all of $\mathbb{R}^2\setminus\{(0, 0)\}$.

Also the mapping f defined by (52) and its inverse are C^∞-smooth on $\mathbb{R}^2 \setminus \{(0, 0)\}$.

It seems that everything concerning f was already proved except the Gateaux differentiability statements at the origin. So, fix any $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$.

If $y = 0$, then we easily get that $Df(0,0)(x,y) = (x,y)$. If $y \neq 0$, then we have $|tx - (ty)^2|/(ty)^2 \geq 1$ (that is, $(tx - (ty)^2, ty) \in \Omega$) for all $0 \neq t \in \mathbb{R}$ sufficiently close to 0, and so

$$\frac{1}{t}\big(f(tx,ty) - f(0,0)\big) = \frac{1}{t}\widehat{f}_1(tx - (ty)^2, ty)) = (x - ty^2, y) \longrightarrow (x,y)$$

as $0 \neq t \to 0$, which means that f is Gateaux differentiable at $(0,0)$ and satisfies $Jf(0,0) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$.

That f^{-1} is not Gateaux differentiable at $(0,0)$ can be proved exactly as for the inverse of the former f defined by (6). (Yes, we profit again from the fact that the derivative $(\eta^{-1})'(0)$ is different from 1.)

References

[1] V. I. Averbukh, O. G. Smolyanov: *The theory of differentiation in linear topological spaces*, Russ. Math. Surv. 22 (1967) 201–258.

[2] H. Cartan: *Calcul Différentiel. I: Calcul Différentiel dans les Espaces de Banach*, Herman, Paris (1967).

[3] R. Cibulka, M. Fabian, A. Y. Kruger: *On semiregularity of mappings*, J. Math. Analysis Appl. 473/2 (2019) 811–836.

[4] A. L. Dontchev, R. T. Rockafellar: *Implicit Functions and Solution Mappings. A View from Variational Analysis*, 2nd ed., Springer Series in Operations Research and Financial Engineering, Springer, New York (2014).

[5] P. Drábek, J. Milota: *Methods of Nonlinear Analysis: Applications to Differential Equations*, 2nd ed., Birkhäuser Advanced Texts, Birkhäuser, Basel (2013).

[6] M. Fabian: *Concerning a geometrical characterization of Fréchet differentiability*, Časopis Pěst. Matematiky 104 (1979) 45–64.

[7] M. Fabian: *Differentiability via one sided directional derivatives*, Proc. Am. Math. Soc. 82 (1981) 495–500.

[8] 'gerw' (https://math.stackexchange.com/users/58577/gerw): *Composition between Frechet and Gateaux derivative*, https://math.stackexchange.com/q/4529899 (version: 2022-09-13).

[9] P. Hájek, M. Johanis: *Smooth Analysis in Banach Spaces*, De Gruyter Series in Nonlinear Analysis and Applications Vol. 19, De Gruyter, Berlin (2014).

[10] S. Krantz, H. R. Parks: *A Primer of Real Analytic Functions*, Birkhäuser, Basel (2002).

[11] S. Lang: *Real and Functional Analysis*, 3rd ed., Graduate Texts in Mathematics, Springer, New York (1993).

[12] S. Yamamuro: *Differential Calculus in Topological Linear Spaces*, Lecture Notes in Mathematics Vol. 374, Springer, Berlin (1974).

[13] L. Zajíček: *Hadamard differentiability via Gâteaux differentiability*, Proc. Amer. Math. Soc. 143/1 (2015) 279–288.