Composition operators with closed range between spaces of smooth functions

Dissertation
zur Erlangung des Grades eines Doktors der Naturwissenschaften

Dem Fachbereich IV der Universität Trier
vorgelegt von

Nicolas François Kenessey

Gutachter:
Prof. Dr. Jochen Wengenroth
Prof. Dr. Leonhard Frerick

Dezember 2012

Acknowledgments

First and foremost I wish to thank my doctoral advisor, Jochen Wengenroth, for his support. The opportunity to be his assistant in the past five years has not only been a pleasant and challenging occupation, but it also has given me the financial support required to concentrate on my research. I also would like to express my deep and heartfelt gratitude to Leonhard Frerick for his guidance and our fruitful conversations. A special thank goes to Thomas Kalmes for his thorough reading of this manuscript and his many useful comments. Furthermore, I want to thank my friends and family, especially my mother, for their constant support throughout my years of studies. Anne, Bastian, Bernhard, and Ulf also deserve a special mention for giving me the needed moral backup. I can only hope that they know how crucial their friendship has been. Last but not least I would like to say that I am grateful to the entire department of mathematics of the University of Trier. The working atmosphere has always been nothing but pleasant and my colleagues always felt like a part of my family. It fills me with great sorrow that the end of my doctoral research forces me to leave them.

The bibliography was produced using BibTeX, the references have been provided by the database MathSciNet of the American mathematical Society.

Contents

Introduction **10**

1 The closure of an algebra of smooth functions in one variable. **17**
1.1 Taylor series and formal compositions of power series 19
1.2 The closure of the composition algebra for an injective smooth curve. . . . 26

2 Characterization of closed composition algebras in one dimension **35**
2.1 Introduction and previous results 35
2.2 Necessity of the conditions 39
2.3 Sufficiency of the conditions 47

3 A multidimensional version of Faà di Bruno's formula and estimates for the higher derivatives of an inverse map **55**
Notation 55
The formula of Faà di Bruno in higher dimensions 56
The lower bound of an operator 64

4 A necessary condition in the case of several variables **67**
The lower distance estimate 77

5 Composition isomorphisms between spaces of flat functions. **83**
5.1 Preliminary definitions 84
5.2 C_ψ as an isomorphism between spaces of flat functions. 89
Main result 94
5.3 Some special cases of closed algebras 99

6 Appendix **115**
6.1 Equivalent seminorms on $\mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ 115
6.2 The theorem of Borel for hyperplanes 117
6.3 Stable diffeomorphisms and norm estimates 118

List of Symbols **128**

Index **129**

Zusammenfassung

Die vorliegende Arbeit beschäftigt sich mit bestimmten mathematischen Eigenschaften von sogenannten Kompositionsalgebren als Teilräume der unendlich oft differenzierbaren Funktionen. Eine Kompositionsalgebra ist dabei die Menge

$$\mathcal{A}(\psi) = \left\{ F \circ \psi : F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R}) \right\},$$

wobei $\mathcal{E}(\mathbb{R}^n, \mathbb{R}^m)$ den Raum der glatten, also unendlich oft differenzierbaren, Funktionen auf $\mathbb{R}^n$ mit Werten in $\mathbb{R}^m$ bezeichnet und $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ selbst eine glatte Abbildung ist. Bekanntlich ist der Raum $\mathcal{E}(\mathbb{R}^n, \mathbb{R}^m)$ versehen mit den Halbnormen

$$\|F\|_{K,n} = \sup\{\|D^\alpha F(x)\| : x \in K, |\alpha| \leq n\},$$

wobei $K \subseteq \mathbb{R}^n$ kompakt und $n \in \mathbb{N}$, ein Fréchetraum.
Wir beschränken uns hierbei im Wesentlichen auf injektive Funktionen ψ und untersuchen wann die oben erwähnte Kompositionsalgebra abgeschlossen ist. Diese Frage wurde bereits von Gläser in [Gla63] sowie Bierstone, Milman und Pawłucki in [BMP96] für reell-analytische Funktionen ψ untersucht.
Im ersten Kapitel greifen wir eine Charakterisierung von Allan, Kakiko, O'Farrell und Watson aus [AKOW98] auf. Diese beschreibt den Abschluss einer Kompositionsalgebra durch formale Potenzreihen, falls ψ eine glatte, injektive Kurve einer einzigen reellen Variable ist. Zuerst befassen wir uns mit der in [AKOW98] erwähnten Komposition von formalen Potenzreihen und versehen diese mit einer sinnvollen mathematischen Definition. Anschließend geben wir alternative, funktionalanalytische Beweise für die Ergebnisse aus der oben erwähnten Arbeit in der Hoffnung, dass diese dazu dienen könnten den Fall $q > 1$ einzuschließen.

Das zweite Kapitel basiert auf der gemeinsamen Veröffentlichung [KW11] mit J. Wengenroth, die teilweise erweitert wurde. Dieser Teil befasst sich ebenfalls mit glatten, injektiven Kurven. Wir geben drei Bedingungen an, die sowohl notwendig, als auch hinreichend für die Abgeschlossenheit von $\mathcal{A}(\psi)$ sind. Das Hauptresultat lautet wie folgt:

Theorem.
Sei $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ eine injektive Kurve. Die Algebra $\mathcal{A}(\psi)$ ist genau dann abgeschlossen, wenn die folgenden drei Bedingungen erfüllt sind:

(i) Die Funktion ψ ist eine eigentliche Abbildung,
(d.h. Urbilder von kompakten Mengen sind kompakt)

(ii) jeder kritische Punkt hat endliche Ordnung
(d.h. für alle $x \in \mathbb{R}$ existiert ein $k \in \mathbb{N}$ mit $\psi^{(k)}(x) \neq 0$),

(iii) das Bild $\psi(\mathbb{R})$ ist eine Whitney-reguläre Menge
(d.h. lokal können zwei Punkte $\psi(x), \psi(y)$ durch eine Kurve in $\psi(\mathbb{R})$ verbunden werden deren Länge $C \cdot \|\psi(x) - \psi(y)\|^\alpha$ nicht überschreitet).

Darüber hinaus stellt sich heraus, dass die drei oben angegebenen Bedingungen äquivalent dazu sind, dass ψ eine lokal Hölder stetige Umkehrabbildung besitzt.

Das dritte Kapitel stellt Rechenmethoden zur Behandlung des Falles einer injektiven Abbildung $\psi : \mathbb{R}^q \to \mathbb{R}^d$ für $q > 1$ bereit. Insbesondere beweisen wir eine explizite Formel (Formel von Faà di Bruno) für die höheren Ableitungen von Kompositionen $F \circ G$ von Funktionen, die von mehreren Variablen abhängen. Darüber hinaus können wir folgende Abschätzung für die Halbnormen der Inversen $\theta = \psi^{-1}$ eines glatten Diffeomorphismus $\psi : U \to V$ angeben, wobei U und V offene Teilmengen des $\mathbb{R}^n$ sein sollen. Die Abschätzung lautet:

$$|||\theta^{(k)}(\psi(x))||| \leq C_k \left(1 + |||\psi|||_{\{x\},k}\right)^{\frac{(k-1)k}{2}} \left(1 + |||\psi'(x)^{-1}|||\right)^{\frac{k(k+1)}{2}} .$$

Hierbei fassen wir die Ableitung $\theta^{(k)} : V \to \mathcal{M}_k(\mathbb{R}^n, \mathbb{R}^n)$ als Abbildung mit Werten im Raum $\mathcal{M}_k(\mathbb{R}^n, \mathbb{R}^n)$ der k-linearen Abbildungen auf dem $\mathbb{R}^n$ auf, welchen wir mit der Norm

$$|||T||| = \sup\{\|T[r_1, ..., r_k]\| : \|r_1\| \leq 1, ..., \|r_k\| \leq 1\}$$

versehen.

Kapitel 4 verallgemeinert die Techniken aus [KW11] unter Ausnutzung der bereitgestellten Werkzeuge, um notwendige Bedingungen für die Abgeschlossenheit von $\mathcal{A}(\psi)$ zu finden. Es stellt sich heraus, dass ψ, wie im eindimensionalen Fall, eine eigentliche Abbildung sein muss. Darüber hinaus muss sie die "untere Distanzabschätzung" erfüllen, das heißt, dass wir für alle kompakten Teilmengen $K \subseteq \mathbb{R}^q$ Konstanten $c, \gamma > 0$ finden können, so dass

$$\|\psi(x) - \psi(y)\| \geq c \cdot \|x - y\| \cdot \max\Big\{ \operatorname{dist}(x, E(\psi)) \ , \ \operatorname{dist}(y, E(\psi))\Big\}^\gamma$$

für alle $x, y \in K$ gilt. Hierbei bezeichnet $E(\psi)$ die kritische Menge $\{z : \psi'(z) \text{ ist nicht injektiv}\}$.

In Kapitel fünf beschäftigen wir uns mit der Menge $\mathcal{I}(E(\psi))$ von flachen Funktionen (i.e. die Funktion samt all ihrer Ableitungen verschwindet) auf der kritischen Menge. Wir zeigen bereits in Kapitel 4, dass diese Menge stets im Abschluss von $\mathcal{A}(\psi)$ enthalten ist. Das Hauptresultat lautet wie folgt:

Theorem.
Sei $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ eine injektive, eigentliche Abbildung mit Whitney-regulärem Bild, die die untere Distanzabschätzung erfüllt. Dann gilt stets $\mathcal{I}(E(\psi)) \subseteq \mathcal{A}(\psi)$.

Im Fall einer diskreten kritischen Menge können wir dadurch außerdem beschreiben, wann $\mathcal{A}(\psi)$ abgeschlossen ist und dadurch unser Ergebnis aus Kapitel 2 verallgemeinern.
Zum Schluß betrachten wir einige Spezialfälle, bei denen die kritische Menge $E(\psi)$ nicht diskret ist. Wir können bei besonderer Struktur von ψ beweisen, dass in diesen Fällen unter den drei oben genannten Bedingungen $\mathcal{A}(\psi)$ abgeschlossen ist.

Introduction

This work will study the closure and closedness of special subalgebras, so-called composition algebras, of the space of smooth functions (in one or several variables). Our main interest will be to give necessary and sufficient conditions for such algebras to be a closed subspace of the space of smooth functions. To do this properly, let us clarify some notation.
We will always write $\mathcal{E}(\mathbb{R}^n, \mathbb{R}^k)$ for the space

$$\mathcal{E}(\mathbb{R}^n, \mathbb{R}^k) = \bigcap_{j=1}^{\infty} C^j(\mathbb{R}^n, \mathbb{R}^k)$$

of smooth (or C^∞) functions on $\mathbb{R}^n$ with values in $\mathbb{R}^k$. We endow this space with the family $\{\|\cdot\|_{K,\ell} : K \subseteq \mathbb{R}^n \text{ compact }, \ell \in \mathbb{N}\}$ of seminorms defined by

$$\|f\|_{K,\ell} = \sup\{\|D^\alpha f(x)\| : x \in K, \alpha \in \mathbb{N}_0^n, |\alpha| \leq \ell\},$$

where $D^\alpha f$ denotes the partial derivative of f with respect to the multi-index $\alpha \in \mathbb{N}_0^n$. The generated locally convex space is Fréchet.

Definition.
For a smooth map $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ we define the composition algebra

$$\mathcal{A}(\psi) = \left\{F \circ \psi : F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})\right\}$$

and consider it as a subspace of $\mathcal{E}(\mathbb{R}^q, \mathbb{R})$ together with the relative topology. We will call ψ the generator of $\mathcal{A}(\psi)$ or simply say that the algebra is generated by ψ.
The composition algebra is the image of the linear map $C_\psi : \mathcal{E}(\mathbb{R}^d, \mathbb{R}) \to \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ defined by

$$C_\psi(F) = F \circ \psi,$$

which is called composition operator with symbol ψ. The composition $F \circ \psi$ is sometimes denoted by $\psi^* F$, for instance by Tougeron in [Tou71], and Bierstone, Milman, and Pawłucki in [BMP96].

We will often have to deal with real-valued smooth functions on both the source $\mathbb{R}^q$ and range $\mathbb{R}^d$ of $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$. In order to facilitate the reading of the (sometimes rather

technical) proofs to come, we will always use capital letters for smooth function $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ defined on the target area of ψ and small letters for functions $f \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ defined on the source of ψ. Note that smooth functions on the domain of ψ are elements of codomain of C_ψ and vice versa.
Let us use this notation to prove the continuity of the composition operator. This is a short application of the closed graph theorem. Indeed to show the continuity explicitly, one would need to find for each compact set $K \subseteq \mathbb{R}^q$ and $n \in \mathbb{N}$ some constants $c > 0, k \in \mathbb{N}$ and a compact subset L of $\mathbb{R}^d$ such that

$$\|F \circ \psi\|_{K,n} \leq c \cdot \|F\|_{L,k}$$

for all $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$. This would require to estimate the partial derivatives of $F \circ \psi$ by means of partial derivatives of F. This can be done, for instance by using the formula of Faà di Bruno (cf. [FdB59], page 3) but would be rather technical.
On the other hand C_ψ is a linear map between the Fréchet spaces $\mathcal{E}(\mathbb{R}^d, \mathbb{R})$ and $\mathcal{E}(\mathbb{R}^q, \mathbb{R})$. The closed graph theorem (cf. [Rud73], 2.15) states that it is continuous if and only if its graph

$$G(C_\psi) = \big\{(F, F \circ \psi) : F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})\big\}$$

is a closed subspace of the product $\mathcal{E}(\mathbb{R}^d, \mathbb{R}) \times \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ endowed with the product Fréchet topology. To prove that $G(C_\psi)$ is closed, consider a sequence $(F_n, F_n \circ \psi)_{n \in \mathbb{N}}$ in $G(C_\psi)$ that converges to some (F, f) in $\mathcal{E}(\mathbb{R}^d, \mathbb{R}) \times \mathcal{E}(\mathbb{R}^q, \mathbb{R})$. This implies that both $F_n \to F$ and $F_n \circ \psi \to f$ pointwise and thus

$$f(x) = \lim_{n \to \infty} (F_n \circ \psi)(x) = \lim_{n \to \infty} F_n(\psi(x)) = F(\psi(x)) = F \circ \psi(x),$$

hence the required $(F, f) \in G(C_\psi)$.
Throughout this work, we will mostly consider *injective* generators $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$. It might seem redundant to the reader that we repeat this assumption in every partial result but we did so on purpose for reasons of completeness. Some results however, such as propositions 1.7 and 2.9, did not require injective maps and might be used to attack related questions for non-injective generators.
The main goal of our work is to give sufficient and necessary conditions when some injective map $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ generates a closed composition algebra. When given a specific kind of generator, this can (sometimes) be decided quite easily. Indeed consider the following example.

Example.
The map defined by $\psi : \mathbb{R}^2 \to \mathbb{R}^2, (x, y) \mapsto (x^3, y^3)$ is smooth and injective. Moreover it is a bijection on $\mathbb{R}^2$ and even a diffeomorphism on $\{(x, y) \in \mathbb{R}^2 : x \neq 0 \text{ and } y \neq 0\}$. It will turn out that this map generates a closed composition algebra.
Let us first give an argument why every flat function on the cross

$$C = \{(x, y) \in \mathbb{R}^2 : x = 0 \text{ or } y = 0\}$$

is already a composition. We will give the general idea rather than a tedious computation. Consider some $f \in \mathcal{E}(\mathbb{R}^2, \mathbb{R})$ that is flat (i.e. all derivatives vanish) on C. The function defined by $F(x^3, y^3) = f(x, y)$ obviously satisfies $F \circ \psi = f$ and it comes down to showing that F is smooth on C. To do this one needs to show that $(x, y) \mapsto f(\sqrt[3]{x}, \sqrt[3]{y})$ is smooth. This can be done by computing the partial derivatives and applying l'Hopital's rule using the fact that f is flat on the cross.
One can show that the algebra generated by ψ is already closed. This however requires a lot of work and results (propositions 5.14, 1.7 and a result involving regularly situated sets, cf. Definition I.5.4 in [Mal67]) as well as a careful decomposition of $f \in \overline{\mathcal{A}(\psi)}$ on its critical set.

Of course the simplicity of the idea behind this example relied heavily on the fact that we had extensive knowledge of the generator ψ and that the geometrical structure of its critical set, the cross C, was easy to deal with. For arbitrary smooth functions, this is generally not the case. Nevertheless we are able to give a necessary condition for the fact that every flat function on $E(\psi)$ is already a composition no matter the shape of the critical set $E(\psi) = \{z : \psi'(z) \text{ is not injective}\}$. This condition reads as follows:

Definition.
We say that a map $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ satisfies the "lower distance estimate" if the following holds: For any given compact set $K \subseteq \mathbb{R}^q$ there are $c, \alpha > 0$ such that

$$\|\psi(x) - \psi(y)\| \geq c \cdot \|x - y\| \cdot \operatorname{dist}(x, E(\psi))^\alpha$$

for all $x, y \in K$.

Basically it requires a certain geometrical property of the critical set. Moreover under the additional assumptions of properness and Whitney-regularity of $\psi(\mathbb{R}^q)$ we also obtain sufficiency.
Let us cite the main result of this work, proposition 5.12.

Theorem.
Let $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ be an injective map that satisfies the following conditions:

(i) ψ is proper,

(ii) $\psi(\mathbb{R}^q)$ is a Whitney-regular set (cf. Definition 2.7),

(iii) for any given compact set $K \subseteq \mathbb{R}^q$ there are $c, \alpha > 0$ such that

$$\|\psi(x) - \psi(y)\| \geq c \cdot \|x - y\| \cdot \operatorname{dist}(x, E(\psi))^\alpha$$

for all $x, y \in K$ and where $E(\psi) = \{z : \psi'(z)$ is not injective$\}$.

In this case every smooth function $f \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ that is flat on $E(\psi)$ is contained in $\mathcal{A}(\psi)$ and hence can be written as $f = F \circ \psi$ where we can choose $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ to be flat on $\psi(E(\psi))$.

Let us give a rough sketch of our work.
The first chapter will revisit the work of Allan, Kakiko, O'Farrell, and Watson on the closure of composition algebras generated by an injective curve. In [AKOW98] and under the assumption of an injective $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ they explicitly described the functions belonging to $\overline{\mathcal{A}(\psi)}$ by means of formal power series. Their characterization involved "formal compositions of formal power series" and it turns out that a smooth function belongs to the closure of $\mathcal{A}(\psi)$ if and only if it has the "right kind" of Taylor series at every point, namely the one of a composition $F \circ \psi$. After giving an exact definition of formal compositions we give slightly modified proofs that rely more on functional analysis in the hope that they might hint at a way to extend the result to the case of several variables.
Section 2 is based on the joined work [KW11] with J. Wengenroth and characterizes those algebras $\mathcal{A}(\psi)$ generated by an injective smooth curve $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ that are closed subspaces of $\mathcal{E}(\mathbb{R}, \mathbb{R})$. The stepping stone to our results is a conjecture by the referee in [AKOW98] stating that $\mathcal{A}(\Psi) = \overline{\mathcal{A}(\Psi)}$ "is probably true for those $\Psi : \mathbb{R} \to \mathbb{R}^r$ that are proper, injective and have only critical points of finite order". This conjecture turns out to be incomplete and we give a simple example for its failure. We prove that ψ generates a closed composition algebra if and only if its inverse map $\psi^{-1} : \psi(\mathbb{R}) \to \mathbb{R}$ is locally Hölder continuous. This means that for every compact subset $L \subseteq \psi(\mathbb{R})$ we can find constants $c_L, \gamma_L > 0$ depending only on ψ and L such that

$$\|\psi^{-1}(z) - \psi^{-1}(w)\| \leq c_L \cdot \|z - w\|^{\gamma_L}.$$

Substituting $\psi(x) = z$ and $\psi(y) = w$ this is equivalent to

$$\|x - y\| \leq c_L \cdot \|\psi(x) - \psi(y)\|^{\gamma_L}$$

and could be interpreted as some sort of geometric stability of ψ.
The third chapter is a preparation for attacking the case of injective maps $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ of several variables. We discuss the notation of differential calculus of smooth functions on $\mathbb{R}^n$ which will be required until the end of our work. We also use the notation to give a proof of the multidimensional version of Faà di Bruno's formula which gives an explicit way to compute the higher order derivative of a composition $F \circ G$ of two smooth functions F and G of several variables. We will also discuss some implications of this formula for the norm-estimates of the inverse map of a diffeomorphism. This estimate reads as follows: For a given diffeomorphism $\psi : U \to V$ with inverse $\theta = \psi^{-1}$, we can find constants C_n depending only on n such that

$$|||\theta^{(n)}(\psi(x))||| \leq C_n \cdot \left(1 + |||\psi|||_{\{x\},n}\right)^{\frac{(n-1)n}{2}} \cdot \left(1 + |||\psi'(x)^{-1}|||\right)^{\frac{n(n+1)}{2}}.$$

Once the symbolism of section 3 has been established we use it to analyze our approach of the one-dimensional case. There, it turns out that only the behavior of ψ on the critical set $E(\psi)$ where the derivative $\psi'(x)$ as a linear map is not injective, is of importance to decide whether or not f belongs to $\overline{\mathcal{A}(\psi)}$. We generalize some proofs of section 1 to give easy conditions for a function f to belong to $\overline{\mathcal{A}(\psi)}$. As in the one-dimensional case it turns out that functions that are flat (i.e. all derivatives vanish) on the critical set $E(\psi)$ are automatically contained in $\overline{\mathcal{A}(\psi)}$. We then deduce a rather simple necessary condition for closed composition algebras, namely the estimate

$$\|\psi(x) - \psi(y)\| \geq c_K \cdot \|x - y\| \cdot \max\{\mathrm{dist}(x, E(\psi)), \mathrm{dist}(y, E(\psi)\}^\alpha$$

on every compact set, which we have baptized "lower distance estimate".
Section 5 will turn our attention to the set $\mathcal{I}(E(\psi)$ of flat functions on $E(\psi)$ and show that the necessary lower distance estimate together with some geometrical property of the image $\psi(\mathbb{R}^q)$ are sufficient to obtain that every flat function on the critical set is not only contained in the closure of $\mathcal{A}(\psi)$, but is already a smooth composition $F \circ \psi$ itself. More precisely, our main result reads as follows.

Theorem.
If $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ is an injective map that is proper, has a Whitney-regular image, and satisfies the lower distance estimate, then $\mathcal{I}(E(\psi))$ is contained in $\mathcal{A}(\psi)$.

If the critical set $E(\psi)$ is discrete these conditions are even sufficient to obtain a closed composition algebra. In this case we even obtain a characterization of closed composition algebras that can be viewed as an extension of our result concerning smooth injective curves from section 2.

Theorem.
An injective map $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ with a discrete critical set generates a closed composition algebra if and only if it is a proper map that satisfies the lower distance estimate.

Finally we study some special cases where the critical set is not discrete. Depending on the structure of ψ we obtain simple examples where $\mathcal{A}(\psi)$ is closed.

Chapter 1

The closure of an algebra of smooth functions in one variable.

As mentioned in the introduction, our interest is to characterize closed composition algebras. To be able to recognize whether $\mathcal{A}(\psi) = \{F \circ \psi : F \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)\}$ is a closed subspace of $\mathcal{E}(\mathbb{R}^q, \mathbb{R})$, the space of smooth functions, it is undoubtedly useful to characterize which functions actually belong to its closure $\overline{\mathcal{A}(\psi)}$. Let us mention some related results.

In [Mal67] (II, theorem 1.3 and corollary 1.7), [Tou72] (V, théorème 1.3 and corollaire 1.6) and [Whi48] the authors characterized the closure of a sub-module M in the space $\mathcal{E}^m(L, E)$ resp. $\mathcal{E}(L, E)$ of C^m- resp. C^∞- (or smooth) germs on L with values in E. The fact that they considered the general case of germ-spaces (i.e. quotient spaces of $\mathcal{E}(\mathbb{R}^n, \mathbb{R})$ by certain subspaces of "flat" functions) is not really important to understand their idea. For instance, their result still holds when considering $L = \mathbb{R}^q$ and $E = \mathbb{R}$ in which case $\mathcal{E}(L, E) = \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ is actually the space of smooth real-valued functions. Their characterization of the closure $\overline{M}$ in $\mathcal{E}(\mathbb{R}^q, \mathbb{R})$ was given by the space

$$\widehat{M} = \{f \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}) : \forall x \in \mathbb{R}^q \ \exists g_x \in M \text{ s.t. } D^\alpha f(x) = D^\alpha g_x(x) \ \forall \alpha\}$$

of all those smooth functions having the right derivatives on L.

In [Tou71] Theorem 1.1, Tougeron was able to extend this result to the closure $\overline{\Phi(M)}$ of the image of a sub-module M under a certain kind of group homomorphism Φ. This result is more general, but applying it to $M = \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ and $\Phi = C_\varphi : \mathcal{E}(\mathbb{R}^d, \mathbb{R}) \to \mathcal{E}(\mathbb{R}^q, \mathbb{R})$, $F \mapsto F \circ \varphi$ it turns out that for certain $\varphi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ the closure of $\mathcal{A}(\varphi)$ is also given by the space

$$\widehat{\mathcal{A}(\varphi)} = \left\{f \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}) : \forall x \in \mathbb{R}^q \, \exists F_x \in \mathcal{E}(\mathbb{R}^d, \mathbb{R}) \text{ s.t. } D^\alpha f(x) = D^\alpha (F_x \circ \varphi)(x) \forall \alpha \in \mathbb{N}_0^q\right\}$$

of functions with pointwise admissible derivatives. Tougeron considered only those smooth maps $\varphi : \mathbb{R}^q \to \mathbb{R}^d$ such that for all compact subsets $K \subseteq \mathbb{R}^q$ and $L \subseteq \mathbb{R}^d$ there is some $\alpha > 0$ such that

$$\Gamma(y) = \sup \left\{ \frac{\operatorname{dist}(x, \varphi^{-1}(y))^\alpha}{\|\varphi(x) - y\|} : x \in K \setminus \varphi^{-1}(\{y\}) \right\} < \infty.$$

This condition is restrictive as for instance $\varphi(x) = x \cdot \exp(-1/x^2)$ fails this property whenever $y = 0$ and K contains a sufficiently small interval $[0, \varepsilon]$, however it holds for smooth injective maps with a local Hölder continuous inverse.
In [AKOW98] the authors were able to show that such a description of the closure of $\mathcal{A}(\psi)$ is always true whenever $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ is an injective smooth curve. As a starting point to our work we will give slightly modified proofs of their result. Though the characterization itself is rather simple, basically it tells that a smooth function $f \in \mathcal{E}(\mathbb{R}, \mathbb{R})$ belongs to the closure of $\mathcal{A}(\psi)$ if and only if it has the Taylor series of a composition $F_x \circ \psi$ at every point $x \in \mathbb{R}$, the proofs however are rather technical.
Unless mentioned otherwise we will always deal with a smooth injective map $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ of one variable. We remind the reader that, as discussed in the introduction, we will always write capital letters for functions $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ and small letters for functions $f \in \mathcal{E}(\mathbb{R}, \mathbb{R})$.
Let us begin by discussing the main theorem of [AKOW98] in more detail and deduce the pointwise Taylor condition from the rather complicated formulation of "formal composition" of formal power series that they used. Their main result reads as follows:

Theorem 1.1.
For an injective map $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ and $f \in \mathcal{E}(\mathbb{R}, \mathbb{R})$ the following are equivalent.

(1) f is contained in the closure of the algebra $\mathbb{P}(\psi) = \{p \circ \psi : p \text{ polynomial in } \mathbb{R}^d\}$.

(2) For all $x \in \mathbb{R}$ the formal Taylor series $T_x^\infty f$ of f about x is a formal composition of some power series S with the Taylor series $T_x^\infty \psi - \psi(x)$.

(3) For all $e \in \mathbb{R}$ satisfying $\psi'(e) = 0$ the formal Taylor series $T_e^\infty f$ of f about e is a formal composition of some power series S with the Taylor series $T_e^\infty \psi - \psi(e)$.

Here, the symbol T_x^∞ is used to describe the map assigning to a function its Taylor series at x.
Allan, Kakiko, O'Farrell, and Watson initially considered the algebra $\mathbb{P}(\psi)$ of polynomials in ψ rather than $\mathcal{A}(\psi) = \{F \circ \psi : F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})\}$. This is however irrelevant as the following argument shows that both sets have the same closure. Indeed, the space $\mathbb{P}(\mathbb{R}^d)$ of polynomials in d variables is dense in $\mathcal{E}(\mathbb{R}^d, \mathbb{R})$ and the continuity of the composition operator shown in the introduction implies

$$\mathcal{A}(\psi) = C_\psi(\mathcal{E}(\mathbb{R}^d, \mathbb{R})) = C_\psi\left(\overline{\mathbb{P}(\mathbb{R}^d)}\right) \subseteq \overline{C_\psi(\mathbb{P}(\mathbb{R}^d))} = \overline{\mathbb{P}(\psi)} \subseteq \overline{\mathcal{A}(\psi)},$$

hence both closures must coincide. This was already mentioned in [AKOW01], where Allan, Kakiko, O'Farrell, and Watson used the symbol $\mathcal{A}(\psi)$ for the closure of the composition algebra which they denoted by $C^\infty(\psi)$.
The conditions (2) and (3) look rather technical and formal compositions of power series cannot be considered common mathematical knowledge. We will therefore give exact definitions and explain what the authors meant. In [AKOW98] the authors described the

closure of the algebra as the space of those functions having the “right kind” of Taylor series at each point stating that

> "The “right kind” is of the form $q \circ (T_a^\infty \psi_1 - \psi_1(a), ..., T_a^\infty \psi_r - \psi_r(a))$, where q is a power series in r variables and $T_a^\infty \psi_i$ denotes the Taylor series of ψ_i about a."

It turns out that this is just a rather complicated way to express the simple fact that $T_x^\infty f = T_x^\infty (F_x \circ \psi)$. The analysis of those formal compositions will lead to the formula

$$T_x^\infty (F \circ G) = (T_{G(x)}^\infty F) \bullet (T_x^\infty G),$$

where $\bullet$ denotes the formal composition of power series.

1.1 Taylor series and formal compositions of power series

First we need to have a solid understanding of formal power series in several variables and we will use bold letters $(\mathbf{S}, \mathbf{g}, ...)$ for formal power series in order to distinguish them from the Taylor series of functions we are about to use. There is more than one approach to defining this space. For instance one could view the space of formal power series in one point to be the projective limit of the spaces $\mathbb{P}_n(\mathbb{R}^d)$ of real-valued polynomials in $\mathbb{R}^d$ with degree inferior or equal to n as did [AKOW98]. For $k \leq n$ the projections would then be given by the truncation maps $\varrho_n^k : \mathbb{P}_n(\mathbb{R}^d) \to \mathbb{P}_k(\mathbb{R}^d)$ defined by

$$\varrho_n^k \left(\sum_{|\alpha| \leq n} c_\alpha x^\alpha \right) = \sum_{|\alpha| \leq k} c_\alpha x^\alpha.$$

We can also consider the case $n = \infty$ to obtain a projection ϱ_∞^k from the space of formal power series to $\mathbb{P}_k(\mathbb{R}^d)$.
This appears very natural together with the approach used in [AKOW98] but requires rather technical computations when defining the formal composition of power series. Indeed considering a formal power series $\mathbf{S}$ in q variables as well as different power series $\mathbf{R}_1, .., \mathbf{R}_q$ with $\mathbf{R}_j(0) = 0$ in d variables, one can view $\mathbf{S}$ and $\mathbf{R}_j$ as the respective sequences $(S^n)_{n \in \mathbb{N}_0}$ and $(R_j^n)_{n \in \mathbb{N}_0}$ of polynomials $S^n = \varrho_\infty^n \mathbf{S} \in \mathbb{P}_n(\mathbb{R}^q)$ and $R_j^n = \varrho_\infty^n \mathbf{R}_j \in \mathbb{P}_n(\mathbb{R}^d)$. The intuitive way to define the formal composition $\mathbf{S} \bullet (\mathbf{R}_1, ..., \mathbf{S}_q)$ would be by the sequence of truncations of the polynomials given by the adequate stepwise composition, namely

$$C^n = \varrho_\infty^n (\mathbf{S} \bullet (\mathbf{R}_1, ..., \mathbf{R}_q)) = \varrho_\infty^n \Big((\varrho_\infty^n \mathbf{S}) \circ (\varrho_\infty^n \mathbf{R}_1, ..., \varrho_\infty^n \mathbf{R}_q) \Big).$$

Note that $\varrho_\infty^n(\mathbf{S})$ and $\varrho_\infty^n(\mathbf{R}_j)$ are polynomials and hence their composition is unproblematic. This seems to be the right definition in the projective limit setting but at the first glance, it is not clear that this sequence actually defines an element of the projective limit, which would require to prove $\varrho_{n+1}^n C^{n+1} = C^n$. To this end, we absolutely need $\mathbf{R}_j(0) = 0$ as the following example shows:

Example 1.2.
Consider the two polynomials $p(x) = x + x^2$ and $q(x) = x + 1$ of one real variable x. The sequence defined above yields $(p \circ q)^0 = 0$ since $\varrho^0_\infty p = 0$. Moreover we have

$$(p \circ q)^1 = \varrho^1_\infty(\varrho^1_\infty p \circ \varrho^1_\infty q) = \varrho^1_\infty(\mathrm{id} \circ (\mathrm{id} + 1)) = \mathrm{id} + 1,$$

as well as

$$(p \circ q)^2 = \varrho^2_\infty(\varrho^2_\infty p \circ \varrho^2_\infty q) = \varrho^2_\infty(p \circ q) = 2 + 3\,\mathrm{id} + \mathrm{id}^2,$$

hence we do not have the required property $\varrho^k_n(p \circ q)^n = (p \circ q)^k$ and the sequence of compositions does not define an element in the projective spectrum.

A different approach comes into mind when considering the fact that a polynomial is uniquely defined by its partial derivatives in $x = 0$. One can view a formal power series, i.e. a coherent sequence of polynomials, as a coherent sequence of numbers. By coherent we mean $\varrho^k_n C^n = C^k$ for $k \leq n$. We can therefore identify a power series $\mathbf{S} = \sum_{\alpha \in \mathbb{N}_0^q} s_\alpha x^\alpha$ with its uniquely determined sequence of coefficients $(s_\alpha)_{\alpha \in \mathbb{N}_0^q}$. This leads to the following definition:

Definition 1.3.
We call $X_d = \mathbb{R}^{\mathbb{N}_0^d}$ endowed with the product topology the space of formal power series in d variables and consider a formal power series simply as a sequence of real numbers.
For $x \in \mathbb{R}^d$ we define the formal Taylor map $T^\infty_x : \mathcal{E}(\mathbb{R}^d, \mathbb{R}) \to X_d$ in x by

$$T^\infty_x F = \left(\frac{D^\alpha F(x)}{\alpha!} \right)_{\alpha \in \mathbb{N}_0^d}$$

as well as the steps $T^k_x : \mathcal{E}(\mathbb{R}^d, \mathbb{R}) \to X^k_d = \mathbb{R}^{\{\alpha \in \mathbb{N}_0^d : |\alpha| \leq k\}}$ for $k \in \mathbb{N}_0$ by

$$T^k_x f = \left(\frac{D^\alpha F(x)}{\alpha!} \right)_{\alpha \in \mathbb{N}_0^d, |\alpha| \leq k},$$

which we can identify with the appropriate Taylor polynomials $\displaystyle\sum_{|\alpha| \leq k} \frac{D^\alpha F(x)}{\alpha!}(y - x)^\alpha$.

The usual notation for the space X_d of formal power series in d variables is given by $\mathbb{R}[[x_1, ..., x_d]]$ which is consistent with its interpretation as the “limit” or completion of the spaces $\mathbb{R}[x_1, ..., x_d]$ of real-valued polynomials in d variables. We have chosen a different symbol to shorten up the notation. Let us remark that from our point of view, the space of formal power series is nothing more than a quotient of the space of smooth functions by a subspace of flat functions. To elaborate this, let us define what we mean by flat function:

Definition 1.4.
We call a function $F \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ flat in $x \in \mathbb{R}^q$ if its Taylor series is zero, meaning $T_x^\infty F = (T_x^\infty F_1, ..., T_x^\infty F_d) = 0 \in X_q^d$ and we will write $\mathcal{I}(\{x\})$ for the set of functions which are flat in x.

Let us note that T_x^∞ is a continuous linear map and also surjective by Borel's theorem. Indeed, using for instance the version found in the book of Trèves ([Trè67], Theorem 38.1) we know that every formal power series in n variables is the Taylor series of a smooth map φ at 0 and considering the translation $\tau_x(F)(z) = F(z - x)$ this obviously remains true for Taylor series at x. The kernel of T_x^∞ is obviously the set $\mathcal{I}(\{x\})$ and this implies that $\tilde{T}_x^\infty : \mathcal{E}(\mathbb{R}^d, \mathbb{R})/\mathcal{I}(\{x\}) \to X_d,\ f + \mathcal{I}(\{x\}) \mapsto T_x^\infty F$ is an isomorphism.
The advantage of our view of X_d as $\mathcal{E}(\mathbb{R}^d, \mathbb{R})/\mathcal{I}(\{x\})$ is the fact that the formal composition of power series is much simpler and more natural to define. As mentioned before, it will turn out that

$$\mathbf{F} \bullet \mathbf{g} = T_0^\infty (F \circ g)$$

is a suitable definition, whenever the functions F and g satisfy $T_0^\infty F = \mathbf{F}$ and $T_0^\infty g = \mathbf{g}$. This means that the composition of power series is nothing more than the composition of functions. To prove that this definition is reasonable we need to verify two things:

(i) First, we require for each $\mathbf{F} \in X_d$ some smooth function $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ satisfying $T_0^\infty F = \mathbf{F}$ but this is already ensured by Borel's theorem ([Trè67], Theorem 38.1). At this point, we can already see why the condition $\mathbf{g}_j^0 = 0$ for all j is so important. Indeed by applying Borel's theorem we only have control over the derivatives of F at $0 \in \mathbb{R}^q$ and the condition $\mathbf{g}_j^0 = 0$ just means that $g(0) = 0$. Having in mind the formula $(F \circ g)'(0) = F'(g(0)) \cdot g'(0)$ it is plausible to assume that the derivatives of $F \circ g$ in 0 can be computed using only the derivatives $D^\alpha F$ in $g(0)$ and $D^\beta g(0)$ (which is actually a generalization of Faà di Bruno's formula and will be proven in a more general setting in section 3). Since 0 is the only point where we have control over the derivatives of F, the condition $\mathbf{g}_j^0 = g_j(0) = 0$ does not seem so arbitrary anymore. In fact we can even use Borel's theorem to circumvent this problem and give a more general formula for the composition by considering some G having the Taylor series $\mathbf{F}$ in $\mathbf{g}^0 = (\mathbf{g}_1^0, .., \mathbf{g}_d^0)$ no matter its value. The formula then reads

$$(\mathbf{F} \bullet (\mathbf{g}_1, ..., \mathbf{g}_d)) = T_0^\infty (F \circ (g_1, ..., g_d))(0),$$

whenever $T_0^\infty g = \mathbf{g}$ and $T_{g(0)}^\infty F = \mathbf{F}$.

(ii) The second thing we need to prove is that the Taylor series $T_x^\infty (F \circ g)$ of the composition $F \circ g$ does not depend on the choice of the representation of $\mathbf{F}$ and $\mathbf{g}$ by functions F and g. This will be done in the following proposition:

Proposition 1.5.
Consider $F, G \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ and $f, g \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$. If $f - g$ is flat in $x \in \mathbb{R}^q$ and $F - G$ is

flat in $y = f(x) \in \mathbb{R}^d$ *the function* $(F \circ f) - (G \circ g)$ *is also flat in* x*, i.e.*

$$T_x^\infty(F \circ f) = T_x^\infty(G \circ g).$$

As a consequence the formal composition of power series $\mathbf{F} \in X_n$ *and* $\mathbf{f}_1, ..., \mathbf{f}_n \in X_q$*, namely*

$$\mathbf{F} \bullet (\mathbf{f}_1, ..., \mathbf{f}_n) = T_x^\infty(F \circ f),$$

whenever some $f = (f_1, ..., f_n) \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ *and* $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ *satisfy both* $T_x^\infty f_j = \mathbf{f}_j$ *and* $T_{f(x)}^\infty F = \mathbf{F}$*, is well-defined.*

Proof. We need to prove that the difference $(F \circ f) - (G \circ g)$ is flat in x which will be done in two steps. First we will show that $H \circ f$ is flat in $f(x)$ whenever $H \in \mathcal{I}(\{f(x)\})$. We will then continue by proving that $G \circ f - G \circ g$ is also flat in x when $f - g \in \mathcal{I}(\{x\})$. Before we start, let us mention why this is already sufficient. Fix some $\alpha \in \mathbb{N}_0^q$ and suppose that $(F - G) \circ f$ and $(G \circ f) - (G \circ g)$ are both flat in x. Then we obtain

$$D^\alpha(F \circ f)(x) = D^\alpha((F - G) \circ f)(x) + D^\alpha(G \circ f)(x) = D^\alpha(G \circ f)(x) = D^\alpha(G \circ g)(x),$$

hence the difference $(F \circ f) - (G \circ g)$ is flat in x.
Let us now proceed with the two steps mentioned above:
(i) The composition $H \circ f$ is flat in $x \in \mathbb{R}^q$ whenever $H \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ is flat in $f(x) \in \mathbb{R}^d$. To prove this, we will show that every evaluation $P \mapsto D^\alpha(P \circ f)(x) = 0$ vanishes on the subset $\mathcal{D}(\{f(x)\}^c)$. The continuity of these evaluations together with the fact that $\mathcal{D}(\{f(x)\}^c)$ is dense in $\mathcal{I}(\{f(x)\})$ will then imply $D^\alpha(H \circ f)(x) = 0$ for all functions $H \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ that are flat in $f(x)$.
For a function $\Phi \in \mathcal{D}(\{g(x)\}^c)$ we can find an open neighborhood U of $f(x)$ such that $U \cap \operatorname{supp}(\Phi) = \emptyset$ and hence $\Phi = 0$ on U. Therefore the composition $\Phi \circ f$ vanishes on the open neighborhood $f^{-1}(U)$ of x and every derivative of $\Phi \circ f$ must vanish in $f(x)$.
Regarding the continuity, we have seen in the introduction that the composition operator $C_f : \mathcal{E}(\mathbb{R}^d, \mathbb{R}) \to \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ defined by $C_f(P) = P \circ f$ is continuous. Since the evaluation map $h \mapsto D^\alpha h(x)$ is also continuous, we obtain the required continuity of $P \mapsto D^\alpha(P \circ f)(x)$.
It remains to prove that $\mathcal{D}(\{f(x)\}^c)$ is dense in the set $\mathcal{I}(\{f(x))\})$ of flat functions in $f(x)$, which we will do by using the theorem of bipolars. We require to show that $u(H) = 0$ for all functions H that are flat in $f(x)$ whenever $u \in \mathcal{D}(\{f(x)\})^\circ$. Fix such a distribution u that vanishes on $\mathcal{D}(\{f(x)\}^c)$, then its support must be contained in $\{f(x)\}$. Theorem 2.3.4 from [Hör03] states that u must be a finite sum of evaluations of derivatives in $f(x)$. If $H \in \mathcal{I}(\{f(x)\})$ every such evaluation $D^\alpha H(f(x)$ is zero, hence $u(H) = 0$. The theorem of bipolars then states that H is contained in the closure of $\mathcal{D}(\{f(x)\})$.
(ii) The difference $(G \circ f) - (G \circ g)$ is flat in x whenever $f - g$ is flat in $x \in \mathbb{R}^d$.
The key argument in part (i) of the proof was the continuity of the maps $H \mapsto D^\alpha(H \circ f)(x)$, which we deduced from the continuity of the composition $H \mapsto H \circ f$. For this trick to work now, we require the continuity of $h \mapsto G \circ h$. Unfortunately, this time the closed graph theorem is not an option as the composition R_G defined by $h \mapsto G \circ h$ is not linear.

Nevertheless we are still dealing with Fréchet hence metrizable spaces and it is therefore sufficient to prove that $G \circ h_n$ converges to $G \circ h$ in $\mathcal{E}(\mathbb{R}^q, \mathbb{R})$ whenever h_n converges to h in $\mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$. Before we proceed with the proof of the continuity, let us explain why it is sufficient for our needs. Suppose the map $h \mapsto D^\alpha(G \circ h)(x)$ is continuous. The function $f - g$ is flat in x hence $h = f - g \in \mathcal{I}(\{x\})$. As seen in (i) the set $\mathcal{D}(\{x\}^c) \subseteq \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ of functions with compact support outside $\{x\}$ is dense in $\mathcal{I}(\{x\})$. We can therefore find a sequence $(h_n)_{n\in\mathbb{N}}$ in $\mathcal{D}(\{x\}^c)$ converging to h in $\mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$. Since $x \notin \operatorname{supp}(h_n)$ the function f and $f - h_n$ coincide on a neighborhood U_n of x. This implies $G \circ f = G \circ (f - h_n)$ on the open set U_n containing x, hence $D^\alpha(G \circ f)(x) = D^\alpha(G \circ (f - h_n))(x)$ for all $\alpha \in \mathbb{N}_0^d$. The continuity of those evaluations then implies the required

$$D^\alpha(G \circ f)(x) = D^\alpha(G \circ (f - h_n))(x) \to D^\alpha(G \circ (f - h))(x) = D^\alpha(G \circ g)(x).$$

Let us proceed with the actual proof of the continuity. We need to prove that the convergence $h_n \mapsto h$ in $\mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ implies that every higher order partial derivative $D^\alpha(G \circ h_n)$ converges uniformly to the appropriate derivative $D^\alpha(G \circ h)$ on every compact subset K of $\mathbb{R}^q$. This can be done by induction.

The case $|\alpha| = 0$ will follow form the mean value inequality. The sequence h_n converges to h uniformly on any compact set K by assumption. This implies that the union of all $h_n(K)$ is bounded as all but a finite number of $h_n(K)$ are contained in the bounded neighborhood $h(K) + [-1, 1]^d$ of the compact set $h(K)$. Thus we can find a convex and compact set $L \subseteq \mathbb{R}^q$ containing every $h_n(K)$ as well as $h(K)$. The mean value inequality together with the fact that the gradient ∇G is bounded on L implies that for all $x \in K$ we have

$$|G \circ h_n(x) - G \circ h(x)| \le \sup\{\|\nabla G(t)\| : t \in L\} \cdot |h_n(x) - h(x)| \le \|G\|_{L,1} \cdot \|h_n - h\|_K \to 0.$$

This implies that $G \circ h_n$ converges to $G \circ h$ uniformly on K as n tends to infinity.

The induction step is given by the chain rule and the continuity of both the addition and the multiplication in $\mathcal{E}(\mathbb{R}^d, \mathbb{R})$. Suppose that for all $\Phi \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ the composition $\Phi \circ h_n$ converges to $\Phi \circ h$ in the space $C^k(\mathbb{R}^q, \mathbb{R})$ of k-times continuously differentiable functions (endowed with the seminorms $\|\cdot\|_{K,k}$, K compact). For $|\alpha| = k+1$ we can write $\alpha = \beta + e_j$ with $|\beta| = k$ where e_j denotes the j-th unit vector. Computing the partial derivatives one obtains that

$$D^\alpha(G \circ h_n) = D^{\beta+e_j}(G \circ h_n) = D^\beta(D^{e_j}(G \circ h_n)) = D^\beta\langle \nabla G \circ h_n, D^{e_j} h_n\rangle.$$

The vector valued functions $\nabla G \circ h_n$ converge to $\nabla G \circ h$ in $C^k(\mathbb{R}^q, \mathbb{R}^d)$ since every coordinate $D_j F \circ h_n$ converges to $D_j F \circ g$ by the induction assumption. The addition and multiplication are continuous in C^k. The scalar products $\langle \nabla G \circ h_n, D^{e_j} h_n\rangle$ must therefore converge to $\langle \nabla G \circ h, D^{e_j} h\rangle$ in $C^k(\mathbb{R}^q, \mathbb{R})$. Applying the chain rule once more one obtains that $D^\alpha(f \circ h_n) \to D^\beta\langle \nabla G \circ h, D^{e_j} h\rangle = D^\alpha(G \circ h)$ locally uniformly. □

Unfortunately, the downside of our definition of the composition of formal power series is the loss of the ability to construct the formal composition explicitly. Obviously, there are many possible choices for functions F and g_j having the right Taylor series $\mathbf{F}$ and $\mathbf{g}_j$. We are nonetheless able to give a simple and consistent definition of compositions of formal power series.

Definition 1.6 (Composition of formal power series)**.**
For power series $\mathbf{F} \in X_d$ and $\mathbf{g}_1, ..., \mathbf{g}_d \in X_q$ we define the formal composition $\bullet$ by

$$\mathbf{F} \bullet (\mathbf{g}_1, ..., \mathbf{g}_d) = T_x^\infty (F \circ g),$$

whenever $T_x^\infty g_j = \mathbf{g}_j$ and $T_{g(x)}^\infty F = \mathbf{F}$.
This way, we obtain the formula described in [AKOW98] as the "higher order version of the chain rule", namely

$$(T_{g(x)}^\infty F) \bullet (T_x^\infty g) = T_x^\infty (F \circ g).$$

Now that we have a solid understanding of the space $X_d = \mathcal{E}(\mathbb{R}^d, \mathbb{R})/\mathcal{I}(\{x\})$ of formal power series, we go for the first step in the proof of theorem 1.1, namely to prove that the Taylor series condition

$$T_x^\infty f \in T_x^\infty \mathcal{A}(\psi)$$

is necessary for some f to belong to $\overline{\mathcal{A}(\psi)}$. Since the Taylor map $T_x^\infty : \mathcal{E}(\mathbb{R}^d, \mathbb{R}) \to X_d$ is continuous (a simple consequence of the closed graph theorem), this comes down to showing that $T_x^\infty \mathcal{A}(\psi)$ is a closed subspace of X_d.

Proposition 1.7.
For $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ and $x \in \mathbb{R}^q$ the subspace $T_x^\infty \mathcal{A}(\psi) = \{T_x^\infty (F \circ \psi) : F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})\}$ of X_q is closed. This implies that for every f contained in the closure of $\mathcal{A}(\psi)$ and every $x \in \mathbb{R}^q$ we can find some $F_x \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ such that

$$T_x^\infty f = T_x^\infty (F \circ \psi) = (T_{\psi(x)}^\infty F) \bullet (T_x^\infty \psi).$$

Proof. Consider the operator $\tau_x : \mathcal{E}(\mathbb{R}^d, \mathbb{R}^q) \to X_q$ defined by $\tau_x F = T_x^\infty (F \circ \psi)$. We aim at proving that $T_x^\infty \mathcal{A}(\psi) = \tau_x^\infty (\mathcal{E}(\mathbb{R}^d, \mathbb{R}^q))$ is a closed subspace of X_q. We will do so by using the closed range theorem (26.3 in [MV97]), which states that the range of a continuous linear map $S : Y \to Z$ between Fréchet spaces is closed if and only if for any zero neighborhood U in Y the set $U^\circ \cap S^t(Z')$ is a Banach disk. We recall that a Banach disk is an absolute convex set B such that $\bigcup_{t>0} tB$ together with the Minkowski functional $p_B(x) = \inf\{t > 0 : x \in tB\}$ of B is a Banach space. In the concrete case of $S = \tau_x^\infty$ we will deduce this from the fact that $U^\circ \cap \tau_x^t (X_q')$ is contained in a finite-dimensional subspace of $\mathcal{E}(\mathbb{R}^d, \mathbb{R}^q)'$.
Fix some zero neighborhood U in $\mathcal{E}(\mathbb{R}^d, \mathbb{R})$.
First let us note that the range of τ_x^t contains the kernel of $T_{\psi(x)}^\infty$. For $T_{\psi(x)}^\infty F = 0$ the composition $F \circ \psi$ must be flat in x, hence $F \in \mathrm{Ker}(\tau_x)$. This means that every $u \in \tau_x^t (X_q')$

vanishes on $\mathcal{I}(\{\psi(x)\}$, which implies that its support must be contained in $\{\psi(x)\}$. We can therefore write $u \in U^\circ \cap \tau_x^t(X_q')$ as

$$u(F) = \sum_{\alpha \in \mathbb{N}_0^d, |\alpha| \leq n} u_\alpha F^\alpha(\psi(x))$$

and it remains to show that the order n of u does not exceed some upper bound N depending only on U. This requires more refined knowledge of U.
Since it is a zero neighborhood by assumption, we can find some compact set $K \subseteq \mathbb{R}^d$ and $N \in \mathbb{N}$ such that $B_{K,N}(0,\varepsilon) = \{F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R}) : \|F\|_{K,N} \leq \varepsilon\}$ is contained in U. If $\psi(x) \in K$, the order of any $u \in U^\circ$ with support in $\{\psi(x)\}$ can not exceed N. If $\psi(x) \notin K$, we can find some function Φ that is constant zero near K and 1 near $\psi(x)$. For $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ the difference $F - F \cdot \Phi$ is flat in $\psi(x)$ hence

$$|u(F)| = |u(F \cdot \Phi)| \leq c \cdot \|F \cdot \Phi\|_{K,n} = 0,$$

and we obtain $U^\circ \cap \tau_x^t(X_q') = \{0\}$.
In any case we obtain that $L = U^\circ \cap \tau_x^t(X_q') \subseteq \operatorname{span}\{\delta^\alpha_{\psi(x)} : \alpha \in \mathbb{N}_0^d, |\alpha| \leq N\}$ and we proceed to show that this is sufficient to deduce that this set is a Banach disk. Obviously, the span of L is a finite dimensional subspace of $\operatorname{span}\{\delta^\alpha_{\psi(x)} : \alpha \in \mathbb{N}_0^d, |\alpha| \leq N\}$. It remains to show that the Minkowski functional p_L generates a complete norm. This is however trivial as soon as we know that p_L actually is a norm, as all norms on finite dimensional spaces are equivalent and thus complete. The inclusion $L \subseteq U^\circ$ implies $p_{U^\circ} \leq p_L$ on the $\operatorname{span}(L)$. Since U° is a Banach disk itself, p_{U° is a norm on $\operatorname{span}(U^\circ)$ hence $0 < p_{U^\circ}(u) \leq p_L(u)$ for all $0 \neq u \in \operatorname{span}(L)$.

□

Let us close this preliminary discussion of formal Taylor series by showing that the Taylor map $T_x^\infty : \mathcal{E}(\mathbb{R}^d, \mathbb{R}) \to X_d$ is even an algebra homomorphism. We already know that the space of formal power series $X_d = \mathbb{R}[[x_1, ..., x_d]]$ together with the usual pointwise addition of coefficients and the multiplication $*$, defined by

$$\left(\sum_{\alpha \in \mathbb{N}_0^d} a_\alpha \right) * \left(\sum_{\alpha \in \mathbb{N}_0^d} b_\alpha \right) = \sum_{\alpha \in \mathbb{N}_0^d} \left(\sum_{\beta + \gamma = \alpha} a_\beta b_\gamma \right) x^\alpha,$$

is an algebra. The Taylor map however allows us to define the same algebra structure simply by applying the isomorphism T_x^∞ to the algebra structure of $\mathcal{E}(\mathbb{R}^d, \mathbb{R})$. Indeed, we simply equip X_d with the usual addition and the following multiplication $* : X_d \times X_d \to X_d$, defined by

$$\mathbf{F} * \mathbf{G} = T_x^\infty(F \cdot G),$$

where F and $G \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ satisfy $T_x^\infty F = \mathbf{F}$ and $T_x^\infty G = \mathbf{G}$. Again the surjectivity of T_x^∞, given by Borel's theorem ([Trè67], Theorem 38.1), implies that we can always find such functions F and G. To show that it is well defined we need to show that $T_x^\infty(F \cdot G) = 0$ whenever F or G is flat in x. Using the commutativity of the multiplication in $\mathcal{E}(\mathbb{R}^d, \mathbb{R})$, we can restrict ourselves to the case of F being flat in x. The flatness of $F \cdot G$ is then simply given by the Leibniz rule

$$D^\alpha(f \cdot g)(x) = \sum_{\substack{\beta \in \mathbb{N}_0^d \\ \beta \leq \alpha}} \binom{\alpha}{\beta} D^\beta f(x) \cdot D^{\alpha-\beta} g(x) = 0.$$

The commutativity, distributivity and associativity of the multiplication in $\mathcal{E}(\mathbb{R}^d, \mathbb{R})$ directly extend to the multiplication in X_d, giving it an algebra structure. Therefore T_x^∞ is indeed a continuous algebra homomorphism. The neutral element with respect to $*$ is clearly given by the Taylor series $T_x^\infty 1$ of the constant function $t \mapsto 1$.

1.2 The closure of the composition algebra for an injective smooth curve.

The results gained in the previous section allow us to reformulate the main theorem of [AKOW98] cited above in theorem 1.1 in a more practical way at least from a functional analytical point of view. As we can see in the part (iii) of theorem 1.1 the set of zeros of ψ' is of particular importance. Let us therefore give this object its own symbol.

Definition 1.8.
Given a smooth map $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$, we denote by

$$E(\psi) = \{x \in \mathbb{R} : \psi'(x) = 0\}$$

the set of critical points (or critical set) of ψ .

This critical set of ψ plays a crucial role in deciding whether a function belongs to the closure of the composition algebra or not. As a preparation for the actual proof of (a slightly altered version of) theorem 1.1, let us explain why only the behavior of a function on the critical set of ψ is important.

Theorem 1.9.
For an injective map $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ we have $\mathcal{D}(E(\psi)^c) \subseteq \overline{\mathcal{A}(\psi)}$. Moreover the set

$$\mathcal{I}(E(\psi)) = \{f \in \mathcal{E}(\mathbb{R}, \mathbb{R}) : f^{(n)}(e) = 0 \textit{ for all } e \in E(\psi), n \in \mathbb{N}_0\}$$

of flat functions on $E(\psi)$ is also contained in the closure of $\mathcal{A}(\psi)$.

Proof. First let us prove that for every $x \notin E(\psi)$ we can find an open neighborhood I_x such that $\mathcal{D}(I_x) \subseteq \overline{\mathcal{A}(\psi)}$.

Fix an x outside $E(\psi)$. By definition one can find some coordinate $j \leq d$ such that $\psi_j'(x) \neq 0$, hence there is an open interval I_x such that $\psi_j : I_x \to J_x := \psi_j(I_x)$ is a diffeomorphism between open intervals.
For $f \in \mathcal{D}(I_x)$ let us write $K = \operatorname{supp}(f)$. We will show that for any $L \subseteq \mathbb{R}$ compact we can construct a function $F_L \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ such that $f - F_L \circ \psi = 0$ on L. For $z = \psi(y) \in \psi(I_x)$ we know how to define F, namely through the identity $F(z) = F \circ \psi(y) = f(y)$, hence

$$F(\psi(y)) = f(y) = f \circ \psi_j^{-1} \circ \psi_j(y) = f \circ \psi_j^{-1} \circ \pi_j(\psi(y)),$$

where $\pi_j : \mathbb{R}^d \to \mathbb{R}$ denotes the projection onto the j-th coordinate. We obtain

$$F = f \circ \psi_j^{-1} \circ \pi_j$$

on $\psi(I_x)$, but the composition with ψ_j^{-1} is only defined on $J_x = \psi_j(I_x)$. To smoothly extend F to some $F_L \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ we need to multiply it with some smooth cutoff function φ_L that is zero outside $\pi_j^{-1}(J_x)$. We need to be even more careful to obtain the actual identity $F_L \circ \psi = f$ on L, since it is possible that there are $\psi(z) \in \pi_j^{-1}(J_x)$ for $z \notin I_x$. The function φ_L would therefore need to be constant one near $\psi(\operatorname{supp}(f))$ and constant zero near every other point $\psi(z) \in \pi_j^{-1}(J_x) \setminus \psi(I_x)$ (or at least for any point $z \in L$). Let us proceed with the construction of such a decent cutoff function.
The set $\pi_j^{-1}(J_x^c)$ is closed as preimage of the closed set J_x^c under the continuous projection π_j. Moreover it does not contain any point of $\psi(K)$, since $\psi(K) \subseteq \pi_j^{-1}(\psi_j(K)) \subseteq \pi_j^{-1}(J_x)$. As ψ is injective and $\psi(K) \subseteq \psi(I_x)$, it follows that the intersection $\psi(K) \cap \psi(L \cap I_x^c)$ is also empty. Therefore the set $\psi(L \cap I_x^c) \cup \pi_j^{-1}(J_x^c)$ is closed as finite union of closed sets and does not intersect $\psi(K)$. Using for instance corollary 1.4.11 from Hörmander's book [Hör03], we can find some smooth function φ_L such that $\varphi_L = 1$ on an open neighborhood V_K of $\psi(K)$ and $\varphi_L = 0$ on a neighborhood V_L of $\psi(L \cap I_x^c) \cup \pi_j^{-1}(J_x^c)$.
Define $F_L = \varphi_L \cdot (f \circ \psi_j^{-1} \circ \pi_j)$. On the one hand, this function is obviously smooth on the open set V_L, which contains $\pi_j^{-1}(J_x^c)$ (and $\psi(L \cap I_x^c)$), since φ_L is constant zero there. On the other hand $\psi_j^{-1} : J_x \to I_x$ is smooth, hence the functions $\psi_j^{-1} \circ \pi_j$ and F_L are also smooth on $\pi_j^{-1}(J_x)$. The two open sets V_L and $\pi_j^{-1}(J_x)$ form a cover of $\mathbb{R}^d$, since their union contains both $\pi_j^{-1}(J_x)$ and $\pi_j^{-1}(J_x^c)$, hence $F_L \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$.
Now let us show that $F_L \circ \psi = f$ on L.

- For $y \in L \setminus I_x$ we have $\psi(y) \in \psi(L \cap I_x^c)$, hence $\varphi_L \circ \psi(y) = 0$. This implies the identity $F_L \circ \psi(y) = 0 = f(y)$ for all $y \notin \operatorname{supp}(f)$.

- For $y \in L \cap I_x$ we have $y = \psi_j^{-1}(\psi_j(y))$, hence $F_L \circ \psi(y) = (\varphi_L \circ \psi)(y) \cdot f(y)$. If $y \in K$ we have $\varphi_L \circ \psi(y) = 1$ hence $F_L \circ \psi(y) = f(y)$, and if $y \notin K = \operatorname{supp}(f)$ we have $F_L \circ \psi(y) = (\varphi_L \circ \psi)(y) \cdot f(y) = 0 = f(y)$ since $f(y) = 0$.

Consider the sets $L_n = [-(n+1), n+1]$ and $F_n \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ satisfying $F_n \circ \psi = f$ on L_n. We obtain that $F_n \circ \psi - f = 0$ on L_n, thus all its derivatives vanish on the interior of L_n,

which contains $[-n, n]$. Since those sets absorb every compact subset of $\mathbb{R}$, the sequence $F_n \circ \psi$ converges to f in $\mathcal{E}(\mathbb{R}, \mathbb{R})$ with respect to the usual topology.
The closure of $\mathcal{A}(\psi)$ is a subspace of $\mathcal{E}(\mathbb{R}^q, \mathbb{R})$ and contains all $\mathcal{D}(I_x)$ for $x \notin E(\psi)$. The linear hull $L = \operatorname{span}\{\mathcal{D}(I_x) : x \notin E(\psi)\}$ and its closure $\overline{L}$ must therefore also be contained in $\overline{\mathcal{A}(\psi)}$. To obtain $\mathcal{I}(E(\psi)) \subseteq \overline{\mathcal{A}(\psi)}$, it remains to show that L is dense in the set $\mathcal{I}(E(\psi))$ of flat functions on $E(\psi)$. To do this we will use the theorem of bipolars. We require to prove $u(f) = 0$, whenever $f \in \mathcal{I}(E(\psi))$ and $u \in L^\circ$. Obviously every u that vanishes on L also vanishes on $\mathcal{D}(I_x)$, hence its support does not contain any $x \notin E(\psi)$. We obtain $\operatorname{supp}(u) \subseteq E(\psi)$ and theorem 2.3.3 from [Hör03] implies that $u(f) = 0$, since every derivative of f vanishes on $E(\psi)$. □

Let us emphasize that we have only proven $\mathcal{D}(E(\psi)^c) \subseteq \overline{\mathcal{A}(\psi)}$, and that in general it is not contained in $\mathcal{A}(\psi)$ itself, as the following counterexample shows. This stronger inclusion requires an additional property of ψ.

Example 1.10.
Suppose one can find $x \notin E(\psi)$ as well as an unbounded sequence $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ such that $\psi(x_n) \to \psi(x)$. In this case $\mathcal{D}(E(\psi)^c)$ is not contained in $\mathcal{A}(\psi)$. Indeed, choose some $f \in \mathcal{D}(E(\psi)^c)$ with compact support and $f(x) = 1$ and suppose one could write $f = F \circ \psi$ with some $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$. The sequence $(x_n)_{n\in\mathbb{N}}$ is unbounded, hence we can find a subsequence $(n_k)_{k\in\mathbb{N}}$ that does not touch the support of f. Since $F(\psi(x)) = 1$ and F is continuous we obtain $F(\psi(x_{n_k})) \to 1$ which leads to the contradiction

$$0 = \lim_{k\to\infty} f(x_{n_k}) = \lim_{k\to\infty} F \circ \psi(x_{n_k}) = 1.$$

The belief that this is not of importance if the critical set $E(\psi)$ is empty is erroneous. The simple example $\psi(x) = (\sin(x), \exp(x))$ satisfies $E(\psi) = \emptyset$, which implies $\overline{\mathcal{A}(\psi)} = \mathcal{E}(\mathbb{R}, \mathbb{R})$ by theorem 1.1. We will see in proposition 2.9 that $\mathcal{A}(\psi)$ can not be closed

The property prohibiting the existence of such an unbounded sequence with bounded image is exactly the properness of ψ. As we will see in proposition 2.16, it turns out that the properness of ψ is actually sufficient to obtain $\mathcal{D}(E(\psi)^c) \subseteq \mathcal{A}(\psi)$ in the injective case, though this does by far not mean that every flat function on $E(\psi)$ is contained in the composition algebra.
As a corollary to theorem 1.9, we can even replace the flatness of f by the flatness of its derivative.

Corollary 1.11.
Consider some injective $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$. Every function $g \in \mathcal{E}(\mathbb{R}, \mathbb{R})$ satisfying $g' \in \mathcal{I}(E(\psi))$ is contained in the closure of $\mathcal{A}(\psi)$.

Proof. Again, fix some $x \notin E(\psi)$ and some open interval $I_x \subseteq E(\psi)^c$ containing x. We will show that every $g \in \mathcal{E}(\mathbb{R}, \mathbb{R})$ satisfying $g' \in \mathcal{D}(I_x)$ is contained in $\overline{\mathcal{A}(\psi)}$. Consider such

a g with $g' \in \mathcal{D}(I_x)$. The function g is constant on the sets $E^-(\psi) = \{e \in E(\psi) : e < x\}$ and $E^+(\psi) = \{e \in E(\psi) : e > x\}$. Using the theorem of bipolars it is sufficient to show $u(g) = 0$, whenever u vanishes on $\mathcal{A}(\psi)$.
For an arbitrary $u \in \mathcal{A}(\psi)^\circ$ we know by theorem 1.9 that $K = \operatorname{supp}(u) \subseteq E(\psi)$. Note that $x \notin E(\psi)$ and thus $E(\psi) = E^-(\psi) \cup E^+(\psi)$. The sets $K^- = K \cap E^-(\psi)$ and $K^+ = K \cap E^+(\psi)$ are disjoint and compact. Their image under ψ must therefore be disjoint and compact due to the injectivity and continuity of ψ. We can now separate those sets using some $\varphi \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ that is constant one near $\psi(K^-)$ and constant zero near $\psi(K^+)$. Define by g^- and g^+ the values of g on $E^-(\psi)$ and $E^+(\psi)$, then $g - g^- \cdot (\varphi \circ \psi) - g^+ \cdot ((1-\varphi) \circ \psi)$ is flat on $K = K^- \cup K^+$. By theorem 2.3.3 from Hörmander's book [Hör03] we obtain

$$u(g - g^- \cdot (\varphi \circ \psi) - g^+ \cdot ((1 - \varphi) \circ \psi)) = 0,$$

hence

$$u(g) = g^- \cdot u(\varphi \circ \psi) + g^+ \cdot u((1 - \varphi) \circ \psi) = 0,$$

since $u \in \mathcal{A}(\psi)^\circ$ and both $\varphi \circ \psi$ and $(1 - \varphi) \circ \psi$ are elements of $\mathcal{A}(\psi)$.
This shows that every $f \in \mathcal{D}(I_x)$ is contained in the subspace $D\left(\overline{\mathcal{A}(\psi)}\right) = \left\{f' : f \in \overline{\mathcal{A}(\psi)}\right\}$ of derivatives of $\overline{\mathcal{A}(\psi)}$. Since $D\left(\overline{\mathcal{A}(\psi)}\right)$ is also a subspace we can deduce

$$\mathcal{D}(E(\psi)^c) = \operatorname{span}\{\mathcal{D}(I_x) : x \in E(\psi)^c\} \subseteq D\left(\overline{\mathcal{A}(\psi)}\right)$$

and the continuity of the differentiation D on $\mathcal{E}(\mathbb{R}, \mathbb{R})$ implies the inclusion

$$\mathcal{D}(E(\psi)^c) \subseteq D\left(\overline{\mathcal{A}(\psi)}\right) \subseteq \overline{D\mathcal{A}(\psi)}.$$

The set on the right hand side being closed we end up with $\mathcal{I}(E(\psi)) = \overline{\mathcal{D}(E(\psi)^c)} \subseteq \overline{D\mathcal{A}(\psi)}$. Now consider some $g \in \mathcal{E}(\mathbb{R}, \mathbb{R})$ with $g' \in \mathcal{I}(E(\psi))$. We obtain $g' \in \mathcal{I}(E(\psi)) \subseteq \overline{D\mathcal{A}(\psi)}$, hence we can find some $g_n \in \mathcal{A}(\psi)$ such that $g_n' \to g'$. The modifications defined by $f_n(t) = g_n(t) - g_n(0) + g(0)$ are also contained in $\overline{\mathcal{A}(\psi)}$, since every constant function is an element of the algebra. Finally, it remains to note that f_n' converges to f' and $f_n(0) = g(0)$, which implies $f_n \to g$ and therefore $g \in \overline{\mathcal{A}(\psi)}$.

$\square$

Let us now reformulate theorem 1.1 in order to give a complete proof.

Theorem 1.12 (Allan, Kakiko, O'Farrell, and Watson)**.**
For an injective map $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ and $f \in \mathcal{E}(\mathbb{R}, \mathbb{R})$ the following are equivalent.

(i) The map f is contained in the closure of the algebra $\mathcal{A}(\psi) = \{F \circ \psi : F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})\}$.

*(ii) For all $x \in \mathbb{R}$ we have $T_x^\infty f \in T_x^\infty \mathcal{A}(\psi)$
(i.e. $\forall x \, \exists \, F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ such that $f^{(k)}(x) = (F \circ \psi)^{(k)}(x)$ for all $k \in \mathbb{N}_0$) .*

(iii) For all $e \in E(\psi)$ we have $T_e^\infty f \in T_e^\infty \mathcal{A}(\psi)$.

The proof of the implication (iii) $\Rightarrow$ (i) in [AKOW98] is rather technical and involves explicit estimates to deduce the local proximity to $\mathcal{A}(\psi)$ for a function satisfying the pointwise Taylor condition (iii). Moreover there is little hope whatsoever to generalize this proof for a generator ψ that does not only depend on one variable. Let us therefore give a slightly modified proof that relies a little less on computation.

Proof. The implication (i)$\Rightarrow$(ii) has been proven in proposition 1.7 and (ii)$\Rightarrow$(iii) is trivial as $E(\psi) \subseteq \mathbb{R}$. Let us therefore concentrate on the proof of the implication (iii)$\Rightarrow$(i).
Consider the subspace

$$\mathcal{T}(\psi) = \bigcap_{e \in E(\psi)} (T_e^\infty)^{-1}(T_e^\infty \mathcal{A}(\psi)).$$

of functions with the "right" Taylor series on $E(\psi)$.
Let us explain the basic idea behind this proof. We will use the theorem of bipolars which states that the closure of the subspace L consists of exactly those $f \in \mathcal{E}(\mathbb{R},\mathbb{R})$ such that $u(f) = 0$ for all $u \in L^\perp$. It will turn out that those distributions have support in the critical set $E(\psi)$, which can have a quite complicated shape. However we can restrict ourselves to some weak-$*$ dense subset of $\mathcal{A}(\psi)^\perp$ using the theorem of Krein-Milman. We will see that elements of this dense subset, the linear hull of some specific extreme points having a much simpler support, allow us to estimate $u(f)$ more explicitly. Compared to the proof in [AKOW98] this reduces the required amount of calculus.
Since we have already seen in proposition 1.7 that $T_e^\infty \mathcal{A}(\psi)$ is always closed and the Taylor maps $T_e^\infty : \mathcal{E}(\mathbb{R},\mathbb{R}) \to X_1 = \mathbb{R}^{\mathbb{N}_0}$ are continuous, the subspace $\mathcal{T}(\psi)$ is closed as intersection of closed sets. Obviously we aim at proving that $\overline{\mathcal{A}(\psi)} = \mathcal{T}(\psi)$ but it will be easier to show that the respective sets of derivatives coincide. Let us mention why this is sufficient. Consider again the derivation map $D : \mathcal{E}(\mathbb{R},\mathbb{R}) \to \mathcal{E}(\mathbb{R},\mathbb{R})$, $f \mapsto f'$ which is a continuous linear surjection. The kernel of D is the set $\mathcal{C}$ of constant functions which is contained in both $\overline{\mathcal{A}(\psi)}$ and $\mathcal{T}(\psi)$. Once we have shown $D\left(\overline{\mathcal{A}(\psi)}\right) = D\mathcal{T}(\psi)$ we will obtain

$$\overline{\mathcal{A}(\psi)} = \overline{\mathcal{A}(\psi)} + \mathcal{C} = D^{-1}\left(D\left(\overline{\mathcal{A}(\psi)}\right)\right) = D^{-1}\left(D\mathcal{T}(\psi)\right) = \mathcal{T}(\psi) + \mathcal{C} = \mathcal{T}(\psi).$$

Let us now continue and prove that $D\left(\overline{\mathcal{A}(\psi)}\right)$ and $D\mathcal{T}(\psi)$ actually are the same.
It turns out that the set $AP(\psi)$ of accumulation points of $E(\psi)$ plays a central role in that matter. The key argument will be that the subspace

$$L = \{f' \in D\mathcal{T}(\psi) : \operatorname{supp}(f') \subseteq AP(\psi)^c\}$$

of $D\left(\overline{\mathcal{A}(\psi)}\right)$ is dense in $D\mathcal{T}(\psi)$. Let us first show that L is indeed contained in $D\left(\overline{\mathcal{A}(\psi)}\right)$.
Take some primitive $f \in \mathcal{T}(\psi)$ with $f' \in \mathcal{D}(AP(\psi)^c)$. The set $M = E(\psi) \cap \operatorname{supp}(f')$ is finite and by assumption we have $T_e^\infty f \in T_e^\infty \mathcal{A}(\psi)$ for every $e \in M$. We can therefore find

$F_e \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ such that $T_e^\infty f = T_e^\infty (F_e \circ \psi)$. Every $e \in M$ is an isolated point of $E(\psi)$ hence we can find $\varphi_e \in \mathcal{E}(\mathbb{R}, \mathbb{R})$ such that $\varphi_e = 1$ near $\{e\}$ and $\varphi = 0$ near $E(\psi) \setminus \{e\}$. By corollary 1.11 we have $\varphi_e \in \overline{\mathcal{A}(\psi)}$ since φ_e' is flat on $E(\psi)$. The function $\sum_{e \in M} (F_e \circ \psi) \cdot \varphi_e$ is contained in the closure of $\mathcal{A}(\psi)$ and the derivative of $f - \sum_{e \in M} (F_e \circ \psi) \cdot \varphi_e$ is flat on $E(\psi)$, hence the function is also contained in the closure of $\mathcal{A}(\psi)$ by corollary 1.11. This implies

$$f = f - \sum_{e \in M} (F_e \circ \psi) \cdot \varphi_e + \sum_{e \in M} (F_e \circ \psi) \cdot \varphi_e \in \overline{\mathcal{A}(\psi)},$$

hence $f' \in D(\overline{\mathcal{A}(\psi)})$.
Next we show that the derivative of $f \in \mathcal{T}(\psi)$ is flat on the set $AP(\psi)$ of accumulation points of $E(\psi)$, which will turn out to be a a rather simple consequence of the mean value theorem. For any given accumulation point $e_\infty \in AP(\psi)$ we can find a sequence $(e_n)_{n \in \mathbb{N}} \in E(\psi) \setminus \{e_\infty\}$ converging to e_∞. Without loss of generality we can suppose this sequence to be strictly monotone. For $f \in \mathcal{T}(\psi)$ we have $f'(e_n) = 0$ and the mean value theorem gives another strictly monotone sequence $(\xi_n)_{n \in \mathbb{N}}$ such that ξ_n lies between e_n and e_{n+1} and $f''(\xi_n) = \frac{f'(e_n) - f'(e_{n+1})}{e_n - e_{n+1}} = 0$. This leads to $f''(e_\infty) = 0$ by continuity and applying the mean value theorem inductively to the strictly monotone sequence of zeros of $f^{(k)}$ converging to e_∞ implies $f^{(k+1)}(e_\infty) = 0$. As a direct consequence every $g \in D\mathcal{T}(\psi)$ is flat on the set $AP(\psi)$ of accumulation points of $E(\psi)$.
To prove $\overline{L} = D\mathcal{T}(\psi)$ we will use the theorem of bipolars which states that a function $g \in \mathcal{E}(\mathbb{R}, \mathbb{R})$ is contained in $\overline{L}$ if and only if $u(g) = 0$ for all u in the annihilator $L^\perp$ of L. Since $\overline{L} \subseteq D\mathcal{T}(\psi)$ it is therefore sufficient to prove $u(g) = 0$ whenever u vanishes on L and $g \in D\mathcal{T}(\psi)$. Our aim is to find a more convenient dense subset of $L^\perp$ that allows us to estimate $u(g)$ for $g \in D\mathcal{T}(\psi)$ explicitly. In order to do so, we require further knowledge about $u \in L^\perp$.
Let us begin by showing that every $u \in L^\perp$ has support in $E(\psi)$. Indeed, the set $\mathcal{D}(E(\psi)^c)$ is contained in $D\left(\overline{\mathcal{A}(\psi)}\right) \subseteq D\mathcal{T}(\psi)$ by corollary 1.11, which implies $\mathcal{D}(E(\psi)^c) \subseteq L$. We obtain that every $u \in L^\perp$ vanishes on $\mathcal{D}(E(\psi)^c)$, hence $\operatorname{supp}(u) \subseteq E(\psi)$ whenever $u \in L^\perp$. Unfortunately the unknown structure of $E(\psi)$, especially near accumulation points, makes it difficult to estimate $u \in L^\perp$ on $D\mathcal{T}(\psi)$ which we require to prove $u|_{D\mathcal{T}(\psi)} = 0$. In fact, as mentioned in [AKOW98], $E(\psi)$ can even be the Cantor set. To this end we can further restrict ourselves to a weak-$*$ dense subset of $L^\perp$ using the Krein-Milman theorem. Indeed as a continuous linear map on the set $\mathcal{I}(AP(\psi))$ of flat functions on $AP(\psi)$ (endowed with the relative topology of $\mathcal{E}(\mathbb{R}, \mathbb{R})$), every $u \in L^\perp$ can be absorbed into a polar set $B(K,n)^\circ$ of some unit ball $B(K,n) = \{f \in \mathcal{I}(AP(\psi)) : \|f\|_{K,n} \leq 1\}$. This polar set and its intersection with $L^\perp$ being weak-$*$ compact and convex the theorem of Krein-Milman states that it is the weak-$*$ closure of the convex hull of its extreme points and we only have to prove $u|_{D\mathcal{T}(\psi)} = 0$ for such an extreme point u.
We will now further characterize these extreme points. To this end fix some unit ball $B(K,n) = \{f \in \mathcal{I}(AP(\psi)) : \|f\|_{K,n} \leq 1\}$ in $\mathcal{I}(AP(\psi))$. We will show that every extreme

point u of $B(K,n)^\circ \cap L^\perp$ is such that $\operatorname{supp}(u) \setminus AP(\psi)$ lies in a convex subset of $AP(\psi)^c$. This means that it is contained in some subinterval I of $AP(\psi)^c$ and therefore encased between two consecutive accumulation points of $E(\psi)$.
Supposing that this not to be the case, we can find an accumulation point $e \in AP(\psi)$ as well as $x, y \in \operatorname{supp}(u) \setminus AP(\psi)$ such that $x < e < y$. Denote by $\mathbb{I}_A$ the indicator function of a subset A of $\mathbb{R}$. For an arbitrary $\varphi \in \mathcal{I}(AP(\psi))$ both functions $\varphi^- = \varphi \cdot \mathbb{I}_{]-\infty,e[}$ and $\varphi^+ = \varphi - \varphi^- = \varphi \cdot \mathbb{I}_{]e,\infty[}$ are smooth and we can define $u^-(\varphi) = u(\varphi^-)$ and $u^+(\varphi) = u(\varphi^+)$. Let us prove that $u^-, u^+ \in B(K,n)^\circ \cap L^\perp$. For $h \in L$ its support is a subset of $AP(\psi)^c$ hence $\operatorname{supp}(h^-) = \operatorname{supp}(h) \cap]-\infty, e[$ is also a subset of $AP(\psi)^c$. Every primitive h^- of h^- is contained in $\mathcal{T}(\psi)$ since $T_z^\infty h^- \in T_z^\infty \mathcal{A}(\psi)$ either directly by assumption for $z < e$ or because for $z \geq e$ it has the same Taylor series as the constant function $t \mapsto h^-(e)$ which is contained in $\mathcal{A}(\psi)$. We therefore obtain $u^- \in L^\perp$ and since $u^+ = u - u^-$ also $u^+ \in L^\perp$. Moreover we have $\|u^-\|_{K,n}^* + \|u^+\|_{K,n}^* = \|u\|_{K,n}^* = 1$. Indeed, $\|u\|_{K,n}^* = 1$, since it is an extreme point of $B(K,n)^\circ$ and the triangle inequality implies

$$\|u\|_{K,n}^* = \|u^- + u^+\|_{K,n}^* \leq \|u^-\|_{K,n}^* + \|u^+\|_{K,n}^*.$$

On the other hand, for $\varphi, \eta \in \mathcal{I}(AP(\psi))$ satisfying $\|\varphi\|_{K,n} \leq 1$ and $\|\eta\|_{K,n} \leq 1$, we obtain $\|\varphi^- + \eta^+\|_{K,n} \leq 1$. Multiplying φ, η with appropriate scalars of absolute value 1 we can suppose $u^-(\varphi^-), u^+(\eta^+) \geq 0$, hence $u^-(\varphi) + u^+(\eta) = u(\varphi^- + \eta^+) \leq 1$. Taking the supremum over all $\varphi, \eta \in B(K,n)$ we obtain $\|u^-\|_{K,n}^* + \|u^+\|_{K,n}^* \leq 1$. Note that since $x \in \operatorname{supp}(u^-)$ and $y \in \operatorname{supp}(u^+)$ neither norm can be zero. We can now write u as the convex combination

$$u = \|u^-\|_{K,n} \frac{u^-}{\|u^-\|_{K,n}} + \|u^+\|_{K,n} \frac{u^+}{\|u^+\|_{K,n}} = \|u^-\|_{K,n} \frac{u^-}{\|u^-\|_{K,n}} + (1 - \|u^-\|_{K,n}) \frac{u^+}{\|u^+\|_{K,n}}$$

of $\frac{u^-}{\|u^-\|_{K,n}}$ and $\frac{u^+}{\|u^+\|_{K,n}}$ which are both contained in $L^\perp \cap B(K,n)^\circ$. The fact that u is an extreme point implies $u = \frac{u^-}{\|u^-\|_{K,n}} = \frac{u^+}{\|u^+\|_{K,n}}$. This means that u vanishes both on $\{f \in \mathcal{I}(AP(\psi)) : \operatorname{supp}(f) \subseteq]-\infty, e]\}$ and $\{f \in \mathcal{I}(AP(\psi)) : \operatorname{supp}(f) \subseteq [e, \infty[\}$ but since their direct sum is $\mathcal{I}(AP(\psi))$ we obtain the contradiction $u = 0$.
Let us now show that such an extreme point u vanishes on $D\mathcal{T}(\psi)$. First we will decompose u to further simplify its support. As seen above the support of u is contained in some compact interval $I = [a,b]$ such that $]a,b[\cap AP(\psi) = \emptyset$. Since the support of u is contained in $E(\psi) \cap I$ which is a countable set having no more than two accumulation points (a and b), every $x \in I \setminus E(\psi)$ has a neighborhood contained in $E(\psi)^c$. Fix such a point $x \in I \setminus E(\psi)$. Once again, we can find some $\varphi \in \mathcal{E}(\mathbb{R}, \mathbb{R})$ such that $\varphi = 1$ near $E(\psi) \cap]-\infty, x[$ and $\varphi = 0$ near $]x, \infty[\cap E(\psi)$. Since the derivatives of φ and $1 - \varphi$ are flat on $E(\psi)$, both are contained in the closure of the algebra $\mathcal{A}(\psi)$. Moreover for $F \in \mathcal{E}(\mathbb{R}, \mathbb{R})$ we have $(F \cdot \varphi)' = F' \cdot \varphi$ near $E(\psi)$ which implies that the "localizations" $\varphi \cdot u$ and $(1 - \varphi) \cdot u$ are contained in $L^\perp$. The intersection of the supports of the localizations and the set of accumulation points contains at most one element.
Considering the localizations separately if necessary, we can suppose without loss of generality that u is such that $\operatorname{supp}(u) \cap AP(\psi)$ consists of no more than one point.

If the support of u is finite, we can easily show that u must vanish on $D\mathcal{T}(\psi)$. Indeed, we can refine our localization φ by some function $\tau \in \mathcal{E}(\mathbb{R}, \mathbb{R})$ that is constant one near the finite set of isolated points $\operatorname{supp}(u)$ and zero near the rest of $E(\psi)$. Since the support of u is a compact set, we can even choose τ to have compact support. For $f \in D\mathcal{T}(\psi)$ the difference $f - \tau \cdot f$ is flat near $\operatorname{supp} u$ and since $\tau \cdot f \in D\mathcal{T}(\psi)$ has compact support in $AP(\psi)^c$, we obtain $u(f) = u(\tau f) = 0$.

If the support of u is not finite, it consists of a countable union of discrete points in $\mathbb{R}$. We can order these points into a strictly monotone sequence $(e_n)_{n\in\mathbb{N}} \subseteq E(\psi)$ that converges to the accumulation point $e_\infty \in AP(\psi)$. Without loss of generality we may suppose that this sequence to be strictly increasing. The basic idea will be to localize u further using a sequence of functions φ_n that switches from 0 to 1 in the biggest gaps of $E(\psi)^c$ to estimate $u(f)$ explicitly. We will now explain how to chose this sequence.

The sequence $|e_n - e_{n+1}|$ converges to 0 hence we can find a smallest integer n_1 such that $d_1 = |e_{n_1} - e_{n_1+1}| \geq |e_j - e_{j+1}|$ for all $j \in \mathbb{N}$. Inductively we can construct a "distance maximizing" subsequence $(n_k)_{k\in\mathbb{N}}$ such that $d_k = |e_{n_k} - e_{n_k+1}| \geq |e_j - e_{j+1}|$ for all $j \geq n_{k-1}$. To use this sequence to evaluate u, consider some function η that is constant one near $]-\infty, 0]$ and zero near $[1, \infty[$. The norms $\|\eta\|_{\mathbb{R},j} = \sup\{|\eta^{(\ell)}(x)| : \ell \leq j, x \in \mathbb{R}\}$ are finite for every $j \in \mathbb{N}$. The "compressions" $\eta_k(t) = \eta\left(\frac{t-e_{n_k}}{d_k}\right)$ are locally constant near $E(\psi)$ and the derivatives satisfy $|\eta_k^{(j)}(z)| \leq \frac{\|\eta\|_{\mathbb{R},j}}{d_k^j}$. Since η_k is locally constant near $E(\psi)$, we obtain $u(\eta_k \cdot f) = 0$ whenever $f \in D\mathcal{T}(\psi)$ since the support of $\eta_k \cdot f$ is a compact subset of $AP(\psi)^c$. As the distribution u has support in the compact interval $I = [e_1, e_\infty]$ and finite order N, we can estimate $|u(f)| \leq C \cdot \|f\|_{I,N}$ using for instance theorem 2.3.10 from Hörmander's book [Hör03]. Since the function $(1 - \eta_k) \cdot f$ is constant 0 on $[e_1, e_{n_k}]$, we obtain for $f \in D\mathcal{T}(\psi)$:

$$|u(f)| = |u((1-\eta_k) \cdot f)| \leq C \cdot \|(1-\eta_k) \cdot f\|_{I,N} = C \cdot \|(1-\eta_k) \cdot f\|_{[e_{n_k}, e_\infty],N}.$$

Applying the Leibniz formula we obtain the estimate $\|g \cdot h\|_{K,\ell} \leq \sum_{|\alpha|\leq\ell} \|g\|_{K,\ell} \cdot \|h\|_{K,\ell}$, which we can dominate by $C_\ell \cdot \|g\|_{K,\ell} \cdot \|h\|_{K,\ell}$. Applying this estimate to $g = 1 - \eta_k$ and $h = f$ leads to

$$|u(f)| \leq C \cdot C_N \cdot \|(1-\eta_k)\|_{[e_{n_k}, e_\infty],N} \cdot \|f\|_{[e_{n_k}, e_\infty],N} \leq \tilde{C}_N \cdot \frac{\|\eta\|_{\mathbb{R},N}}{d_k^N} \cdot \|f\|_{[e_{n_k}, e_\infty],N}.$$

Now fix some $x \geq e_{n_k}$. By definition we have $d_k = |e_{n_k} - e_{n_k+1}| \geq |e_\ell - e_{\ell+1}|$ for all $\ell \geq n_k$, hence the interval $[x, x+(j+1)d_k]$ contains at least $j+1$ different elements of the sequence $(e_n)_{n\in\mathbb{N}}$. Since $f \in D\mathcal{T}(\psi)$, we have $f(e_n) = 0$ and the interval $[x, x+(j+1)d_k]$ must contain $j+1$ different zeros of f. The mean value theorem implies that it contains at least one zero y_j of $f^{(j)}$ for $j \leq N$. This leads to the estimate

$$|f^{(j)}(x)| = |f^{(j)}(x) - f^{(j)}(y_j)| = |f^{(j+1)}(\xi_j)| \cdot |x - y_j| \leq |f^{(j+1)}(\xi_j)| \cdot (j+1) \cdot d_k.$$

Replacing x by ξ_j and iterating the estimate, we obtain for $j \leq N$:

$$\begin{aligned}
|f^{(j)}(x)| &= \left|f^{(j+N+1)}(\xi_{j+N+1})\right| \cdot \prod_{\ell=1}^{N+1} ((j+\ell) \cdot d_k) \\
&\leq \left|f^{(j+N+1)}(\xi_{j+N+1})\right| \cdot (j+N+1)! \cdot d_k^{N+1} \\
&\leq \|f\|_{I,2N+1} \cdot (2N+1)! \cdot d_k^{N+1}.
\end{aligned}$$

This implies that

$$\begin{aligned}
|u(f)| &\leq \tilde{C}_N \cdot \frac{\|\eta\|_{\mathbb{R},N}}{d_k^N} \cdot \|f\|_{[e_{n_k}, e_\infty],N} \\
&\leq \tilde{C}_N \cdot \frac{\|\eta\|_{\mathbb{R},N}}{d_k^N} \cdot \|f\|_{I,2N+1} \cdot (2N+1)! \cdot d_k^{N+1} \\
&\leq \tilde{C}_N \cdot \|\eta\|_{\mathbb{R},N} \cdot \|f\|_{I,2N+1} \cdot (2N+1)! \cdot d_k \to 0,
\end{aligned}$$

hence the required $u(f) = 0$, since $d_k \to 0$, thus completing ending the proof.

$\square$

Chapter 2

Characterization of closed composition algebras in one dimension

The main aim of this section will be to characterize when an injective smooth symbol $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ generates a closed composition algebra $\mathcal{A}(\psi) = \{F \circ \psi : F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})\}$. This property was called "composite function property" by Bierstone and Milman (Definition 1.7 in [BM98]). It will turn out that, in the one-dimensional case, a smooth injective curve $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ has the composite function property if and only if its inverse $\psi^{-1} : \psi(\mathbb{R}) \to \mathbb{R}$ is locally Hölder continuous. We will also see that this is equivalent to the three conditions of properness, Whitney-regularity of the image, and finite order of each critical point. Most of the results have already been published in the joined work [KW11] with Wengenroth in a somewhat shorter form but we have slightly improved proposition 2.9.

2.1 Introduction and previous results

As an introduction, let us briefly cite the work already done by Bierstone, Glaeser, Milman, and Pawłucki.
In [Gla63], Glaeser deals with a smooth map $\theta : U \to V$ where $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^p$ are open sets for $p \leq n$. He gives the following sufficient conditions regarding the question whether a smooth analytic map $\theta : U \to V$ has the composite function property.

Theorem 2.1 (Glaeser, [Gla63], Théorème II)**.**
If θ satisfies the conditions $\Theta_1, \Theta_2, \Theta_3, \Theta_4$ below, the subalgebra $\mathcal{A}(\theta)$ is closed in $\mathcal{E}(U, \mathbb{R})$.

Θ_1. *θ is a real-analytic map from U to V*

Θ_2. *The rank of θ is equal to p on an everywhere-dense open subset of U.*

Θ_3. *The image $\theta(U)$ is closed in V.*

Θ_4. *For every compact set $K \subseteq \theta(U)$ there exists a compact subset $L \subseteq U$ such that $K = \theta(L)$.*

We remind the reader that every map satisfying Θ_4 is called semiproper. In [BMP96] the authors gave the following equivalent definition. They called a map $f : A \to B$ between topological spaces semiproper if $f(A)$ is closed in B and the quotient topology on $f(A)$ coincides with the one induced by B. They also mentioned that this property is equivalent to the following definition, which will better suit our needs.

Definition 2.2.
A continuous map $\varphi : X \to Y$ is called semiproper if, for every compact set $K \subseteq Y$, there is a compact set $L \subseteq X$ such that $\varphi(L) = K \cap \varphi(X)$.

Since we mainly deal with injective functions, let us mention that injective semiproper maps are actually proper, i.e. preimages of compact sets are compact themselves. Obviously, since φ is injective, the set L in the definition above is unique and must coincide with the preimage of K, hence must be compact.
The results of Bierstone and Milman are not restricted by the dimension condition seen in Glaeser's work. They also pointed out that the composition property, the fact that the composition algebra $\mathcal{A}(\varphi)$ is closed, was more a geometrical characteristic of the image of φ. They considered the more general case of a real-analytic mapping $\varphi : M \to N$ between real-analytic manifolds where they used the symbol $\varphi^*(C^\infty(N))$ for the set $\{F \circ \varphi : F \in C^\infty(N)\}$ and $\varphi^*(C^\infty(N))\widehat{\ }$ for the set of functions with the right Taylor series (i.e. the Taylor series of a composition $F \circ \varphi$) at every point. For $M = \mathbb{R}$ this set is the closure of the algebra $\mathcal{A}(\varphi) = \varphi^*(C^\infty(\mathbb{R}^d))$ as we have seen in proposition 1.12. Let us first clarify some notions necessary to formulate their result.

Definition 2.3.

(i) ([BM88], definition 2.1) A subset $X \subseteq M$ is called semianalytic if every point $x \in X$ has a neighborhood U such that $X \cap U \subseteq \mathcal{S}(\mathcal{O}(U))$. Here $\mathcal{O}(U)$ denotes the set of real-analytic maps on U and $S(\mathcal{O}(U))$ stands for the smallest family of subsets of U that is stable under finite union, finite intersection, and complement, and also contains all preimages $\{f(x) > 0\}$ for every $f \in \mathcal{O}(U)$.

(ii) ([BM88], definition 3.1) A subset $X \subseteq M$ is called subanalytic if every point $x \in X$ has a neighborhood U such that $X \cap U$ is the image under a projection of some relatively compact semianalytic set.

(iii) We write $\mathcal{E}^m(\mathbb{R}^d, \mathbb{R})$ for the space of m-times continuously differentiable, real-valued functions on $\mathbb{R}^d$. For a given subset $M \subseteq \mathbb{R}$ we denote by $\mathcal{I}^m(M)$ the subspace $\{F \in \mathcal{E}^m(\mathbb{R}^d, \mathbb{R}) : f^{(k)}(x) = 0 \text{ for all } k \leq m, x \in M\}$ of m-flat functions on M. We define the space of m-times continuously differentiable germs on M as the quotient $\mathcal{E}^m(M, \mathbb{R}) = \mathcal{E}^m(\mathbb{R}^d, \mathbb{R})/\mathcal{I}^m(M)$. For $m = \infty$, we will abbreviate $\mathcal{I}(M) = \mathcal{I}^\infty(M)$ and $\mathcal{E}(M, \mathbb{R}) = \mathcal{E}^\infty(M, \mathbb{R})$.

In this setting and in the case of the real-analytic manifold $M = \mathbb{R}$, the first part of their result reads as follows.

Theorem 2.4 (Bierstone, Milman [BM98] Theorem 1.13)**.**
Let $\varphi : \mathbb{R} \to \mathbb{R}^n$ be a real-analytic and proper (or even semiproper) map such that the image $\varphi(\mathbb{R})$ is a closed subanalytic subset of $\mathbb{R}^n$. The algebra $\mathcal{A}(\psi)$ is closed if and only if the space $\mathcal{E}(\varphi(\mathbb{R}), \mathbb{R})$ of smooth germs on $\varphi(\mathbb{R})$ is the same as the intersection of all $\mathcal{E}^m(\varphi(\mathbb{R}), \mathbb{R})$.

Bierstone and Milman also give additional characterizations of closed composition algebras. For instance they show that it is also equivalent to the semicoherence of $\varphi(\mathbb{R})$ or the fact that φ satisfies certain types of *uniform Chevalley estimate*. They also note in [BM98], theorem 1.23 that the composite function property implies the existence of an extension operator $E : \mathcal{E}(\varphi(\mathbb{R}), \mathbb{R}) \to \mathcal{E}(\mathbb{R}^d, \mathbb{R})$.

For us, the interesting fact about the results of both Glaeser and Bierstone and Milman is that they deal with maps that need not be injective and therefore require the semiproperness of the generator. This is not a coincidence but actually a necessary condition no matter the injectivity of the generator or the dimensions of its domain and codomain. This might hint to a way of getting rid of the fact that ψ has to be injective, thereby allowing us to generalize our results to the case of a smooth curve. Unfortunately, this is out of our range for now.

Let us proceed with results of our own. We will mostly seize a suggestion by the referee of [AKOW98] mentioned in the previous section. He conjectured that the composition algebra $\mathcal{A}(\psi) = \{F \circ \psi : F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})\}$ should already be closed whenever the composition symbol $\psi \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ is a (smooth) injective and proper curve which has only critical points of finite order.

As we have mentioned above, this question has already been intensely discussed by Bierstone, Glaeser, and Milman in the case of a real-analytic map, but to our knowledge thus far smooth injective maps $\psi : \mathbb{R}^n \to \mathbb{R}^{n+k}$ have not been dealt with.

To gain a better understanding of the conjecture mentioned above, let us give a formal definition of the crucial notion of finite order.

Definition 2.5.
A smooth function $f \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ is said to be of finite order at x if there exists some differentiation order $n \in \mathbb{N}$ such that $D^n f(x) \neq 0$. The smallest such integer $n = n(x)$ is called the order of f at x.

Let us now show that these two conditions of properness and finite order of the critical points are not sufficient to obtain a closed composition algebra by constructing a rather simple counterexample.

Example 2.6.
Consider the map $\varphi : \mathbb{R} \to \mathbb{R}$ defined by $\varphi(t) = \exp\left(-\frac{1}{t}\right)$ for $t > 0$ and $\varphi(t) = 0$ for $t \leq 0$. This map is smooth and flat at $t = 0$. In order to obtain both finite order and injectivity, consider $\psi : \mathbb{R} \to \mathbb{R}^2$ defined by $\psi(t) = (t^2, \varphi(t))$. Its image around zero is given by the following curve.

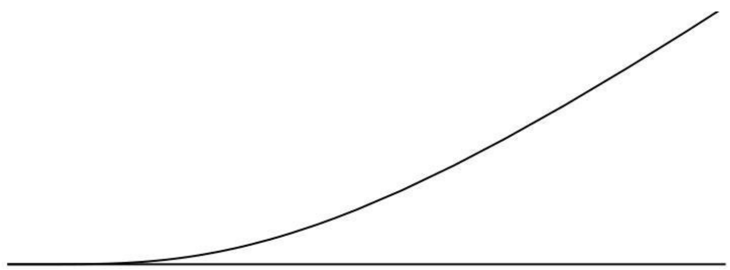


Figure 2.1: The sharp cusp of the image $\psi(\mathbb{R})$ near 0.

Obviously, the only critical point of ψ is $t_0 = 0$ which has order 2, since its first coordinate has non-vanishing second derivative. To see the injectivity, consider $x, y \in \mathbb{R}$ satisfying $\psi(x) = \psi(y)$. Suppose $x \neq y$. The equality of the first coordinate implies $x^2 = y^2$ and therefore $x = -y$. Without loss of generality we can assume $x < y$ and thus $x < 0 < y$ which results in the contradiction $\varphi(x) = 0 < \varphi(y)$. To prove the properness of ψ, fix some compact set $K \subseteq \mathbb{R}^2$. The first coordinate ψ_1 of ψ is the monomial $t \mapsto t^2$, which is proper. Considering the projection $K_1 = \pi_1(K)$ of K onto the first coordinate we obtain $\psi^{-1}(K) \subseteq \psi_1^{-1}(K_1)$, which must be compact as a closed subset of a compact set.
The function $\sqrt{\varphi}$ is flat in the only critical point $z = 0$ of ψ. It therefore has the same Taylor series in z as the function that is constant zero and theorem 1.12 implies that $\sqrt{\varphi}$ must be contained in the closure of $\mathcal{A}(\psi)$. Supposing that the composition algebra is closed, we can find some $F \in \mathcal{E}(\mathbb{R}^2, \mathbb{R})$ such that $\sqrt{\varphi} = F \circ \psi$. For $t > 0$ we obtain

$$\exp\left(-\frac{1}{2t}\right) = \sqrt{\varphi(t)} = \sqrt{\varphi(t)} - \sqrt{\varphi(-t)} = F \circ \psi(t) - F \circ \psi(-t) = F(t^2, \varphi(t)) - F(t^2, 0).$$

But as continuously differentiable function, F is Lipschitz-continuous near $(0,0)$ and we can find some constant $C \geq 0$ such that

$$\exp\left(-\frac{1}{2t}\right) = \left|F(t^2, \varphi(t)) - F(t^2, 0)\right| \leq C \cdot \|(t^2, \varphi(t)) - (t^2, 0)\| = C \cdot \|\varphi(t)\| = C \exp\left(-\frac{1}{t}\right)$$

for all $t > 0$ sufficiently small. Multiplying both sides with $\exp\left(\frac{1}{2t}\right)$, we obtain

$$1 \leq C \cdot \exp\left(-\frac{1}{2t}\right),$$

which leads to a contradiction since the right hand side converges to 0 for $t \to 0$.

The reason why we can not consider the branches $\{\psi(t) : t \leq 0\}$ and $\{\psi(t) : t \geq 0\}$ separately is our use of the Lipschitz continuity of F. Indeed we can consider two distances between points on $\psi(\mathbb{R})$. The first would be the usual Euclidean distance in $\mathbb{R}^2$, in this example $|\psi(t) - \psi(s)|$. The second is the geodesic distance of $\psi(s)$ and $\psi(t)$ on $\psi(\mathbb{R})$. Supposing $s \leq t$, this is precisely the length of the curve $\psi|_{[s,t]}$, namely $\int_s^t \|\psi'(x)\| dx$. The composition property links both distances prohibiting the case encountered above and requires that the euclidean distance can be dominated locally by (some potential power of) the geodesic distance. We have given a sketch of the image of ψ around zero in figure 2.1 to better demonstrate this fact. As we can see, the ratio between the geodesic and Euclidean distances of two points on the same vertical line but different branches of $\psi(\mathbb{R})$ explodes as the vertical line approaches zero.
Having this in mind, the following definition of Schwartz ([Sch66], 8, III.9) will prove useful.

Definition 2.7.
A closed arcwise connected subset $M \subseteq \mathbb{R}^d$ is called Whitney-regular if for every point $z \in M$ one can find a relative ball $B(z,\varepsilon)$ and constants $C, \gamma > 0$ such that the geodesic distance (the smallest length of a rectifiable curve in $M \cap B(z,\varepsilon)$) between two points $x, y \in M \cap B(z,\varepsilon)$ can be dominated by $C \cdot \|x - y\|^\gamma$.

The most interesting property of a connected Whitney-regular set M is that distributions with support in M can be estimated by seminorms involving only evaluations of the functions and its derivatives in M ([Hör03], Theorem 2.3.11). This is gravely false if the set M is not Whitney-regular as mentioned in [Hör03], example 2.3.2, where Hörmander constructed a distribution u with compact support and of order one that can not be estimated by a seminorm $\|\cdot\|_{\mathrm{supp}(u),k}$ for any $k \in \mathbb{N}$.
Let us first prove that in the injective case, the two previous conditions of properness and finite order together with the Whitney-regularity of the image $\psi(\mathbb{R}^d)$ are necessary for the composition algebra to be closed.

2.2 Necessity of the conditions

Throughout this section we will suppose that the algebra $\mathcal{A}(\psi) = \{F \circ \psi : F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})\}$ is a closed subspace of $\mathcal{E}(\mathbb{R}, \mathbb{R})$, hence a Fréchet space itself when endowed with the relative topology of $\mathcal{E}(\mathbb{R}, \mathbb{R})$. As a continuous surjective linear map between Fréchet spaces the composition operator $C_\psi : \mathcal{E}(\mathbb{R}^d, \mathbb{R}) \to \mathcal{A}(\psi)$ is open and we can use this fact to obtain our main tool.

Proposition 2.8.
If $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ generates a closed composition algebra $\mathcal{A}(\psi)$, then the following condition holds: For every compact set $K \subseteq \mathbb{R}^d$ and $n \in \mathbb{N}$ we can find $L \subseteq \mathbb{R}^q$ compact and $k \in \mathbb{N}$ as well as some constant $c > 0$ such that any $f \in \mathcal{A}(\psi)$ can be factorized as $f = F \circ \psi$ with

F satisfying

$$\|F\|_{K,n} \leq \begin{cases} c \cdot \|f\|_{L,k}, & \textit{if } \|f\|_{L,k} \neq 0 \\ 1/2, & \textit{if } \|f\|_{L,k} = 0 \end{cases}.$$

Proof. As mentioned above, $C_\psi : \mathcal{E}(\mathbb{R}^d, \mathbb{R}) \to \mathcal{A}(\psi)$ is a continuous linear surjection between Fréchet spaces, hence open. This means that the image $C_\psi(U)$ of a zero neighborhood U in $\mathcal{E}(\mathbb{R}^d, \mathbb{R})$ is again a zero neighborhood in $\mathcal{A}(\psi)$ with respect to the relative topology induced by $\mathcal{E}(\mathbb{R}, \mathbb{R})$. Applying this to the zero neighborhood U given by the unit ball $B_{K,n} = \{F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R}) : \|F\|_{K,n} \leq 1\}$, we can find an open subset V of $\mathcal{E}(\mathbb{R}, \mathbb{R})$ such that $V \cap \mathcal{A}(\psi) \subseteq C_\psi(B_{K,n})$. The multiples $\varepsilon B_{L,k} = \{f \in \mathcal{E}(\mathbb{R}, \mathbb{R}) : \|f\|_{L,k} \leq \varepsilon\}$ of unit balls in $\mathcal{E}(\mathbb{R}, \mathbb{R})$ form a basis of the zero-neighborhoods in $\mathcal{E}(\mathbb{R}, \mathbb{R})$, where ε ranges over all positive numbers, k over all integers, and L over all compact subsets of $\mathbb{R}$. We can therefore find a compact set $L \subseteq \mathbb{R}$, $k \in \mathbb{N}$, and $\varepsilon > 0$ such that $\varepsilon B_{L,k} \subseteq V$. We obtain $\varepsilon B_{L,k} \cap \mathcal{A}(\psi) \subseteq C_\psi(B_{K,n})$ and the linearity of C_ψ implies that every $f \in \mathcal{A}(\psi)$ with $\|f\|_{L,k} \leq 1$ can be represented as $f = F \circ \psi$ with $\|F\|_{K,n} \leq c = \frac{1}{\varepsilon}$.
For $\|f\|_{L,k} \neq 0$ we have $\frac{f}{\|f\|_{L,k}} \in B_{L,k}$, hence we can represent it as some composition $G \circ \psi = \frac{f}{\|f\|_{L,k}} \in B_{L,k}$ with $\|G\|_{K,n} \leq c$. For $F = \|f\|_{L,k} \cdot G$ we obtain both $F \circ \psi = f$ and $\|F\|_{K,n} \leq c\|f\|_{L,k}$.
For $\|f\|_{L,k} = 0$ the unit ball $B_{L,k}$ contains the entire line $\{\lambda f : \lambda \in \mathbb{R}\}$. Therefore every multiple λf can be represented as $\lambda f = G_\lambda \circ \psi$ with $\|G_\lambda\|_{K,n} \leq c$. For $\lambda = 2c$ we obtain $F = \frac{G_\lambda}{\lambda}$ such that both $F \circ \psi = f$ and $\|F\|_{K,n} \leq \frac{1}{2}$.

□

The deduced condition of proposition 2.8 is too technical to easily decide if, for a given $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R})$, the composition algebra is closed or not. However it can be used to determine simple criterias for closed algebras. The next necessary condition is a slight improvement upon the statement we made in [KW11] (3, step 1). While it uses the same proof, it is neither restricted to injective maps nor to curves (i.e. maps of one variable). It was already mentioned in 1.4 of [BS83] that this condition is necessary to obtain a closed composition algebra though the proof was omitted.

Proposition 2.9.
If a function $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ generates a closed algebra $\mathcal{A}(\psi) = \{F \circ \psi : F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})\}$ in $\mathcal{E}(\mathbb{R}^q, \mathbb{R})$, it must be a semiproper map.

Proof. We recall that by definition 2.2 we need to find for every compact subset $K \subseteq \mathbb{R}^d$ another compact subset $L \subseteq \mathbb{R}^q$ such that $\psi(L) = K \cap \psi(\mathbb{R}^q)$.
Fix a compact subset K of $\mathbb{R}^d$. As mentioned above, we can choose $L \subseteq \mathbb{R}$ compact, $k \in \mathbb{N}$, and $c > 0$ such that $B_{L,k} \cap \mathcal{A}(\psi) \subseteq C_\psi(cB_{K,0})$. It will turn out that $K \cap \psi(\mathbb{R}^q)$ is contained in the image $\psi(L)$ of the compact set $L \subseteq \mathbb{R}$, which implies semiproperness since then $\psi(L \cap \psi^{-1}(K)) = \psi(\mathbb{R}^d) \cap K$.

Now suppose there is some $\psi(x) \in K \setminus \psi(L)$. The two sets $\{\psi(x)\}$ and $\psi(L)$ are compact and disjoint. Using again corollary 1.4.11 from Hörmander's book [Hör03], we can find $G \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ such that $G(\psi(x)) = 1$ and $G = 0$ on an open neighborhood V of $\psi(L)$. The composition $g = G \circ \psi$ is constant zero on the open neighborhood $\psi^{-1}(V)$ of L hence $\|g\|_{L,k} = 0$. Applying proposition 2.8, we can write $g = \Gamma \circ \psi$ with some $\Gamma \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ satisfying $\|\Gamma\|_{K,0} \leq \frac{1}{2}$. Since $\psi(x) \in K$ we obtain the contradiction

$$1 = |g(x)| = |\Gamma(\psi(x))| \leq \|\Gamma\|_{K,0} \leq \frac{1}{2}.$$

□

As a corollary, we obtain the first necessary condition in the injective case.

Corollary 2.10.
If an injective function $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ generates a closed composition algebra $\mathcal{A}(\psi)$, then ψ must be a proper map.

Proof. As seen in proposition 2.9, for every compact set $K \subseteq \mathbb{R}^d$ we can find a compact set $L \subseteq \mathbb{R}^q$ such that $\psi(L) = K \cap \psi(\mathbb{R})$. The injectivity of ψ therefore implies that $L = \psi^{-1}(K)$ is compact. □

Note that, since $\|F\|_{K,n} \leq c$ implies $\|F\|_{K,0} \leq c$, we even obtain a slightly stronger result, namely that $\psi^{-1}(K) \subseteq L$ whenever $B_{L,k} \cap \mathcal{A}(\psi) \subseteq C_\psi(cB_{K,n})$.
Alternatively one could prove this fact by applying the surjectivity criterion already used in proposition 1.7 (cf. 26.1 in [MV97]). It states that a continuous linear map $T : X \to Y$ between Fréchet spaces is surjective if and only if preimages under T^t of bounded sets in X' are again bounded in Y'. For a compact set $K \subseteq \mathbb{R}^d$, the set of evaluations $\{\delta_x : x \in K\}$ is bounded, since it is contained in the polar set of the unit ball $B_{K,0}$. For $s \in \psi^{-1}(K)$, we have $\delta_{\psi(s)}(F) = F(\psi(s)) = \delta_s(F \circ \psi) = C_\psi^t(\delta_s)(F)$, hence the set $\{\delta_s : s \in \psi^{-1}(K)\} \subseteq C_\psi^{-t}(B_{K,0}^\circ)$ is bounded in $\mathcal{A}(\psi)'$. We can therefore find some unit ball $B_{L,n}$ and a $\lambda > 0$ such that $\delta_s \in \lambda B_{L,n}^\circ$ for all $s \in \psi^{-1}(K)$, which also implies $\psi^{-1}(K) \subseteq L$.
An easy consequence of the properness of ψ in the injective case is the following statement about the continuity of the inverse map.

Remark 2.11.
If $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ is proper and injective, and if we endow $\psi(\mathbb{R}^q)$ with the relative topology of $\mathbb{R}^d$, the inverse map $\psi^{-1} : \psi(\mathbb{R}^q) \to \mathbb{R}^q$ must be continuous. In particular, this is always the case when an injective ψ generates a closed composition algebra.

Proof. We will show that images of closed sets under ψ are closed (remark 1.4.1 in [BS83]). Consider a closed set $C \subseteq \mathbb{R}^q$ as well as a sequence $(\psi(c_n))_{n \in \mathbb{N}}$ in $\psi(C)$ converging to some $x \in \mathbb{R}^d$. The set $K = \{\psi(c_n) : n \in \mathbb{N}\} \cup \{x\}$ is compact, hence $\psi^{-1}(K)$ is also compact.

We can therefore find a subsequence φ such that $c_{\varphi(n)}$ converges in $\mathbb{R}^q$ to some c_∞. The set C being closed, we have $c_\infty \in C$ and the continuity of ψ implies

$$x = \lim_{n\to\infty} \psi(c_{\varphi(n)}) = \psi(c_\infty) \in \psi(C).$$

The preimage of C under $\gamma = \psi^{-1}$ is given by $\gamma^{-1}(C) = \psi(C)$ which is closed in $\psi(\mathbb{R}^q)$ hence $\gamma = \psi^{-1}$ is continuous. It follows that for any open set $O \subseteq \mathbb{R}$ the image $\psi(O) = \gamma^{-1}(O)$ is open in $\psi(\mathbb{R}^q)$ and can therefore be represented as the intersection $\psi(O) = U \cap \psi(\mathbb{R}^q)$ where $U \subseteq \mathbb{R}^d$ is open.
If $\mathcal{A}(\psi)$ is closed, we have seen in proposition 2.9 that ψ must be a semiproper map. Its injectivity together with definition 2.2 implies that it is proper hence ψ^{-1} is continuous. $\square$

We will now use the fact that ψ must be a proper map to extract geometrical conditions on ψ' required for the proof of both the finite order condition and the Whitney-regularity of $\psi(\mathbb{R})$.

Lemma 2.12.
If $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ is injective and the composition algebra $\mathcal{A}(\psi)$ is closed, then ψ' must satisfy the following estimate:
For all compact sets $J \subseteq \mathbb{R}$, there are constants $c > 0$ and $\ell \in \mathbb{N}$ such that

$$\operatorname{dist}(x, E(\psi))^\ell \le C \cdot \|\psi'(x)\|$$

for all $x \in J$.

Proof. Fix a compact set $J \subseteq \mathbb{R}$ and consider $K = \psi(J)$ and $n \ge 1$. We can choose L, k and c as in proposition 2.8. We obtain that every $f \in \mathcal{A}(\psi)$ has a representation $f = F \circ \psi$ such that $\|F\|_{K,1} \le c \cdot \|f\|_{L,k}$. For $L \subseteq M$ and $k \le r$ we have $\|f\|_{L,k} \le \|f\|_{M,r}$. Therefore we can enlarge L to be a compact interval containing the set $J = \psi^{-1}(K)$, which is compact by the properness of ψ. Furthermore we can also suppose $k \ge 2$. For $f = F \circ \psi$ and $x \in L$ we then have $\psi(x) \in K$ and applying the Cauchy-Schwarz inequality to the formula given by the chain rule implies that

$$|f'(x)| = |(F\circ\psi)'(x)| \le \sum_{j=1}^d |\partial_j F(\psi(x))\psi_j'(x)| \le \|\nabla F(\psi(x))\| \cdot \|\psi'(x)\| \le \|F\|_{K,n} \cdot \|\psi'(x)\|.$$

For the specific choice of F given by proposition 2.8, namely $\|F\|_{K,n} \le c \cdot \|f\|_{L,k}$, we obtain $|f'(x)| \le c \cdot \|f\|_{L,k} \cdot \|\psi'(x)\|$ whenever $\|f\|_{L,k} \ne 0$. Define $d(x) = \operatorname{dist}(x, E(\psi))$ and consider $\varphi \in \mathcal{D}((-1,1))$ satisfying $\varphi'(0) = 1$. For any $x \notin E(\psi)$ we can construct the bump function $f_x \in \mathcal{E}(\mathbb{R}, \mathbb{R})$ by $f_x(t) = \varphi\left(\frac{t-x}{d(x)}\right)$. The support of f_x does not intersect $E(\psi)$ hence proposition 1.9 implies $f_x \in \overline{\mathcal{A}(\psi)} = \mathcal{A}(\psi)$. If we further restrict ourselves to $x \in L$, the norm $\|f_x\|_{L,k}$ cannot vanish since this would mean that $f_x' = 0$ on L which contradicts $f_x'(x) = \frac{1}{d(x)}$. For $x \in L \setminus E(\psi)$ we obtain a function $F_x \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ satisfying

$$\frac{1}{d(x)} = |f_x'(x)| \leq \|F_x\|_{K,n} \cdot \|\psi'(x)\| \leq c \cdot \|f_x\|_{L,k} \cdot \|\psi'(x)\| \leq c \cdot \frac{\|\varphi\|_{\mathbb{R},k}}{d(x)^k} \cdot \|\psi'(x)\|,$$

and multiplying both sides by $d(x)^k$ leads to the required estimate

$$\operatorname{dist}(x, E(\psi))^{k-1} \leq c \cdot \|\varphi\|_{\mathbb{R},k} \cdot \|\psi'(x)\| \leq C \cdot \|\psi'(x)\|.$$

Since $\ell = k - 1 \geq 1$, the inequality is trivially true for $x \in E(\psi)$ hence is satisfied on L, which contains J.

$\square$

As a direct consequence, we can deduce the following geometrical property of the critical set $E(\psi)$.

Proposition 2.13.
If an injective map $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ generates a closed composition algebra $\mathcal{A}(\psi)$, every critical point must be of finite order.

Proof. Suppose there is a critical point $e \in E(\psi)$ that does not have finite order.
Consider any compact neighborhood I of e and choose $\ell = \ell(I)$ and $C = C(I)$ as in lemma 2.12 such that

$$\operatorname{dist}(x, E(\psi))^{\ell} \leq C \cdot \|\psi'(x)\|$$

holds for all $x \in I$.
Choose $\delta > 0$ such that $\|\psi^{(\ell+1)}(x)\| \leq \varepsilon_\ell = (2^{\ell+1} C (\ell+1)^\ell)^{-1}$ for all x in the compact subinterval $J = [e - \delta, e + \delta]$ of I. The distance to $E(\psi)$ being a continuous function, we can find some $y \in J$ such that $d(x) \leq d(y)$ for all $x \in J$. In order to shorten the following estimates a bit, we will write $d = d(y)$. By definition of d any closed subinterval of I of length $2d$ contains at least one critical point. Therefore every closed subinterval J of length $2(\ell+1)d$ contains at least $\ell + 1$ different zeros of ψ_j' for any coordinate $j \leq d$. Applying Rolle's theorem repeatedly we see that J contains at least one (actually $\ell - k + 2$) zero, z_k, of $\psi_j^{(k)}$ for all $1 \leq k \leq \ell + 1$ and each $j \leq d$.
If $2(\ell+1) \cdot d \leq \delta$, fix some coordinate $j \leq d$ of ψ. We can repeatedly apply the mean value theorem to the different zeros $z_k \in J$ of $\psi_j^{(k)}$ to obtain a sequence $\xi_1, ..., \xi_\ell \in J$ such that

$$\begin{aligned}
|\psi_j'(y)| &= |\psi_j'(y) - \psi_j'(z_1)| = |\psi_j''(\xi_1)| \cdot |y - z_1| \\
&\leq |\psi_j''(\xi_1)| \cdot 2(\ell+1) \cdot d = |\psi_j''(\xi_1) - \psi''(z_2)| \cdot 2(\ell+1) \cdot d \\
&\leq |\psi_j^{(3)}(\xi_2)| \cdot |\xi_1 - z_2| \cdot 2(\ell+1) \cdot d \leq |\psi_j^{(3)}(\xi_2)| \cdot (2(\ell+1) \cdot d)^2 \\
&\dots \\
&\leq |\psi_j^{(\ell+1)}(\xi_\ell)| \cdot (2(\ell+1) \cdot d)^\ell.
\end{aligned}$$

If otherwise, we have $2(\ell+1)\cdot d > \delta$. We can then successively apply Rolle's theorem to the zeros $z_k = e$ of $\psi_j^{(k)}$ to obtain $|\psi_j'(y)| \leq |\psi_j^{(\ell+1)}(\xi_\ell)|\cdot(2(\ell+1)d)^\ell$ with some $\xi_\ell \in I$.
By assumption we can dominate $|\psi_j^{(\ell+1)}(z)|$ by $\frac{1}{2^{\ell+1}C(\ell+1)^\ell}$ on J and lemma 2.12 implies the following inequality

$$\operatorname{dist}(y, E(\psi))^\ell \leq C\cdot\|\psi'(y)\| \leq \varepsilon_\ell\cdot C\cdot(2(\ell+1)\cdot d(y))^\ell = \frac{\operatorname{dist}(y, E(\psi))^\ell}{2}.$$

for all $y \in J$. This of course can only be true if $\operatorname{dist}(y, E(\psi)) = 0$ on J. Therefore ψ must be constant on J in contradiction to the injectivity. □

As a simple consequence, let us show that the set $E(\psi)$ of critical points must be discrete whenever $\mathcal{A}(\psi)$ is closed and injective.
Indeed, suppose $E(\psi)$ has an accumulation point e_∞. We can find a strictly monotone sequence $(e_n)_{n\in\mathbb{N}}$ in $E(\psi)\setminus\{e_\infty\}$ converging to e_∞. As seen in the proof of theorem 1.12 we can iteratively apply Rolle's theorem to a strictly monotone sequence of zeros of $\psi_j^{(k)}$ to obtain another strictly monotone sequence of zeros of $\psi_j^{(k+1)}$ converging to e_∞. This implies $\psi^{(k)}(e_\infty) = 0$ by continuity for all $k \in \mathbb{N}$. The point e_∞ does therefore not have finite order in contradiction to the previous proposition 2.13.
Let us now proceed with the fact that ψ must have a Whitney-regular image. According to definition 2.7 we need to prove that for any point $x \in \psi(\mathbb{R})$ we can find $\delta, \gamma, C > 0$ such that every two points $y, z \in \psi(\mathbb{R})$ with $|z-x|, |y-x| < \delta$ can be connected by a curve Γ in $\psi(\mathbb{R})$ of length inferior to $C\cdot|y-z|^\gamma$. Since ψ is injective the only possible curve connecting $y = \psi(s)$ and $z = \psi(t)$ is $\Gamma = \psi|_{[s,t]}$, which has the length $\int_s^t \|\psi'(w)\|dw$. Dominating $\|\psi'\|$ by some constant C_I on a compact interval I containing $\psi^{-1}(B(x,\delta))$, we obtain the estimate

$$\int_s^t \|\psi'(w)\|dw \leq C_I\cdot|t-s|.$$

It is therefore sufficient to prove that $|t-s| \leq C\cdot\|\psi(t)-\psi(s)\|^\gamma$ for all $\|\psi(s)-x\| < \delta$ and $\|\psi(t)-x\| \leq \delta$.

Proposition 2.14.
If an injective map $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ generates a closed composition algebra $\mathcal{A}(\psi)$, then the inverse map $\psi^{-1} : \psi(\mathbb{R}) \to \mathbb{R}$ must be locally Hölder continuous. In particular $\psi(\mathbb{R})$ is a Whitney-regular subset of $\mathbb{R}^d$.

Proof. Let us begin by proving the local Hölder continuity of the inverse. This is always true near some point x if it is the image $x = \psi(a)$ of some regular point $a \in E(\psi)^c$, and will be proved using the fact that ψ is a diffeomorphism between the one-dimensional manifolds $\mathbb{R}\setminus E(\psi)$ and its image under ψ. Indeed, by assumption we can choose a coordinate j such that $\psi_j'(a) \neq 0$. Restricting ourselves to some interval $I = [a-\varepsilon, a+\varepsilon]$ such that $|\psi_j'| > c > 0$

on I, we can suppose that the sign of ψ_j' is constant hence we obtain

$$c \cdot |t-s| < \int_s^t |\psi_j'(w)| dw = \left| \int_s^t \psi_j'(w) dw \right| = |\psi_j(t) - \psi_j(s)| \leq \|\psi(t) - \psi(s)\|$$

for all $s, t \in I$. Choosing $\delta > 0$ such that $B(x,\delta) \cap \psi(\mathbb{R}) \subseteq \psi(I)$,we obtain the Hölder continuity around $x = \psi(a)$ with Hölder exponent $\gamma = 1$ and $C = 1/c$.
Now consider some critical point $e \in E(\psi)$. Since the translation $\tau_e(f)(t) = f(t-e)$ defines an isomorphism on $\mathcal{E}(\mathbb{R},\mathbb{R})$, the algebra $\tau_e(\mathcal{A}(\psi))) = \mathcal{A}(\tau_e(\psi))$ is also closed and we can suppose $e = 0$. By proposition 2.13 the order n of 0 must be finite and we can find an interval $I = [-\varepsilon, \varepsilon]$ and a coordinate j such that $\psi_j^{(n)}(x) \neq 0$ on I. We recall that the order n of x was defined as the smallest integer such that $\psi^{(n)}(x) \neq 0$. The derivative ψ_j' has no zero other than 0 on I since otherwise applying Rolle's theorem $n-1$ times would yield a zero $\xi \in I$ of $\psi_j^{(n)}$. Using l'Hospital's rule $n-1$ times for the quotients $\psi_\ell'(t)/\psi_j'(t)$, we see that each of them has a (finite) limit at 0, namely $\psi_\ell^{(n)}(0)/\psi_j^{(n)}(0)$. By reducing ε we can dominate every coordinate $|\psi_\ell'|$ by $|\psi_j'|$ and suppose that

$$\|\psi'(t)\| \leq c \cdot |\psi_j'(t)|$$

with some constant $c \geq 1$.
The set $E(\psi)$ being discrete, we can further reduce ε to suppose $\operatorname{dist}(t, E(\psi)) = |t-e| = |t|$ on I. Applying lemma 2.12 to I, we obtain $\ell \in \mathbb{N}$ and $C > 0$ such that $|t|^\ell \leq C \cdot \|\psi'(t)\|$ on I.
If t and s are on the same side of 0, we can suppose w.l.o.g. that $0 < |s| < |t|$. Since ψ_j' has no zero on the interval $\{\lambda t + (1-\lambda)s : \lambda \in [0,1]\}$, the sign of ψ_j' is constant there. For $x = t/s \geq 1$ and $k \in \mathbb{N}$, we can use the estimate

$$(x-1)^k \leq (x-1) \cdot x^{k-1} \leq (x-1) \sum_{j=0}^{k-1} x^j = x^k - 1$$

to obtain $|t-s|^k \leq |t^k - s^k|$ by multiplying both sides with $|s|^k = |s^k|$. Note that this estimate is also trivially satisfied for $|s| = 0$ and we can use it for $k = \ell + 1$ to obtain

$$\frac{1}{\ell+1}|t-s|^{\ell+1} \leq \frac{1}{\ell+1}|t^{\ell+1} - s^{\ell+1}| = \left| \int_s^t w^\ell dw \right|.$$

Note that the sign of w is constant. We can further estimate using $|w|^\ell \leq C \cdot \|\psi'(w)\|$ and get

$$\begin{aligned}
\frac{1}{\ell+1}|t-s|^{\ell+1} &\leq \int_s^t C \cdot \|\psi'(w)\| dw \leq \int_s^t Cc \cdot |\psi_j'(w)| dw \\
&= Cc \cdot \left| \int_s^t \psi_j'(w) dw \right| = Cc \cdot |\psi_j(t) - \psi_j(s)| \\
&\leq Cc \cdot \|\psi(t) - \psi(s)\|,
\end{aligned}$$

where we have used that the sign of ψ_j' is constant to pull the absolute value out of the integral.
Now suppose $s < 0 < t$ and consider the interval $I = [-\varepsilon, \varepsilon]$ for $\varepsilon > 0$. We need to refine the estimate given in corollary 2.12. To do so, fix a function $\varphi \in \mathcal{D}(-1,1)$ satisfying $\varphi(0) = 1$ and define the compressions φ_x by $\varphi_x(w) = \varphi\left((w-x)/|x|\right)$ for $x \neq 0$. If $|x|$ is sufficiently small, we have $\operatorname{supp}(\varphi_x) \subseteq B(x, |x|)$, which is contained in $E(\psi)^c$. Proposition 1.9 thus implies that $\varphi_x \in \mathcal{A}(\psi)$. Applying proposition 2.8 to a compact set K containing $\psi(I)$ and $n = 1$, we obtain another compact set $L \supseteq \psi^{-1}(K)$ and $c, k \in \mathbb{N}$ such that any $f \in \mathcal{A}(\psi)$ can be factorized as $F \circ \psi$ with $\|F\|_{K,1} \leq c \cdot \|f\|_{L,k}$ whenever $\|f\|_{L,k} \neq 0$. Since $\|\varphi_x\|_{L,k} \neq 0$ we get a factorization $F_x \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ such that $\|F_x\|_{K,1} \leq c \cdot \|\varphi_x\|_{L,k} \leq c \cdot |x|^{-k} \cdot \|\varphi\|_{\mathbb{R},k}$. We can now use the mean value inequality for F_t to estimate $1 = \varphi_t(t) = \varphi_t(t) - \varphi_t(s)$ and obtain

$$1 = \varphi_t(t) - \varphi_t(s) = |F_t(\psi(t)) - F_t(\psi(s))| \leq \|F_t\|_{K,1} \cdot \|\psi(t) - \psi(s)\| \leq \frac{C_k}{|t|^k} \cdot \|\psi(t) - \psi(s)\|,$$

where $C_k = c \cdot \|\varphi\|_{L,k}$. Multiplying both sides with $|t|^k$ the estimate reads

$$|t|^k \leq C_k \cdot \|\psi(t) - \psi(s)\|,$$

and the same procedure applied to φ_s gives us $|s|^k \leq C_k \cdot \|\psi(t) - \psi(s)\|$. Using the trivial inequality $|t - s| \leq 2 \cdot \max\{|t|, |s|\}$ leads to

$$|t - s|^k \leq 2^k \cdot \max\{|t|^k, |s|^k\} \leq 2^k \cdot C_k \cdot \|\psi(t) - \psi(s)\|$$

and taking the k-th root on both sides implies the local Hölder continuity of the inverse.
Finally, let us prove that $\psi(\mathbb{R})$ is Whitney-regular. The set $\psi(\mathbb{R})$ is arcwise connected since the curve $\psi|_{[s,t]}$ connects $\psi(s)$ and $\psi(t)$ in $\psi(\mathbb{R})$. Corollary 2.11 states that it is also closed. For $x \in \psi(\mathbb{R})$ we can use the local Hölder continuity of ψ^{-1} to find $\delta, C, \gamma > 0$ such that $|t - s| < C \cdot \|\psi(t) - \psi(s)\|^\gamma$ for all $\psi(t), \psi(s) \in B(x, \delta)$. Reducing δ we can suppose that $B(x, \delta) \cap \psi(\mathbb{R}) \subseteq \psi(I)$ for some compact interval I. Dominating the length of the curve $\psi|_{[s,t]}$, we obtain the estimate required for the Whitney-regularity:

$$L(\psi|_{[s,t]}) = \int_s^t \|\psi'(w)\| dw \leq C_I \cdot |t - s| \leq C_I C \cdot \|\psi(t) - \psi(s)\|^\gamma.$$

□

Let us finish this subsection with the remark that in the injective case the local Hölder continuity of the inverse does not only imply the Whitney-regularity of the image but even unifies it with the two previous necessary conditions of properness and finite order.

Remarks 2.15.
Given an injective $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R})$, the local Hölder continuity of the inverse $\psi^{-1} : \psi(\mathbb{R}) \to \mathbb{R}$

implies that ψ is a proper map, possesses only critical points of finite order, and has a Whitney-regular image. Indeed, the local Hölder continuity implies continuity. Therefore preimages under ψ of compact sets in $\mathbb{R}^d$ are images under ψ^{-1} of compact sets in $\psi(\mathbb{R})$, hence compact. The finite order condition is a consequence of Taylor's formula. For some critical point $e \in E(\psi)$ we can find a neighborhood I of e such that $|t-e|^k \leq c \cdot \|\psi(t)-\psi(e)\|$ for all $t \in I$. Suppose the order of e is superior to k. Then the Taylor polynomial $T_e^k\psi$ is constant which leads to the contradiction

$$|t-e|^k \leq \|\psi(t)-\psi(e)\| = \|\psi(t) - T_e^k\psi(t)\| \leq C \cdot |t-e|^{k+1}.$$

Finally, the Whitney-regularity of $\psi(\mathbb{R})$ has been proven in 2.14.

2.3 Sufficiency of the conditions

We will now proceed to show that the local Hölder continuity of the inverse is also sufficient for the composition algebra to be closed. Throughout this section we will only consider injective curves $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ that satisfy the following three conditions:

(i) ψ is a proper map.

(ii) $\psi(\mathbb{R})$ is a Whitney-regular set.

(iii) Every critical point $e \in E(\psi)$ is of finite order.

Note that even though the local Hölder continuity of the inverse map is necessary, we do not demand ψ to satisfy it explicitly. In light of remark 2.15, it appears to be a stronger requirement than the three conditions above. It will turn out that those are actually sufficient and hence equivalent to the local Hölder continuity of the inverse map.
As we have seen before in proposition 1.9 every function with compact support outside $E(\psi)$ is contained in the closure of $\mathcal{A}(\psi)$ but is not necessarily a composition itself. We will begin by showing that this is the case if we demand that ψ is a proper map. The constructive argument given below will also be of importance to the proof of the general case. Let us emphasize that the proof neither relies on the finite order, nor on the Whitney-regularity of the image.

Proposition 2.16.
Given a proper injective map $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$, every smooth function f with support in $\mathbb{R} \setminus E(\psi)$ is contained in $\mathcal{A}(\psi)$.

Proof. Consider some $f \in \mathcal{E}(\mathbb{R}, \mathbb{R})$ with support in $\mathbb{R} \setminus E(\psi)$. For $x \in E(\psi)^c$ there is a coordinate $j = j(x)$ such that $\psi_j'(x) \neq 0$. We can therefore find an open interval I_x containing x such that $\psi_j'(t) \neq 0$ for all $t \in I_x$. The coordinate ψ_j being monotone on I_x, its image $J_x = \psi_j(I_x)$ is an interval and $\psi_j : I_x \to J_x$ is a diffeomorphism. Since ψ is an injective proper map, corollary 2.11 states that $\psi^{-1} : \psi(\mathbb{R}) \to \mathbb{R}$ is continuous. This implies that the image $\psi(I_x)$ of the open set I_x in $\mathbb{R}$ is open in $\psi(\mathbb{R})$ as the preimage under ψ^{-1} of I_x. We can therefore find some $U_x \subseteq \mathbb{R}^d$ such that $\psi(I_x) = U_x \cap \psi(\mathbb{R})$. Let us again write π_j for the continuous projection onto the j-th coordinate. Considering the intersection of U_x with $\pi_{j(x)}^{-1}(J_x)$, which is open and also contains $\psi(I_x)$, we can suppose that $\pi_{j(x)}(U_x) \subseteq J_x$. The function defined on U_x by

$$F_x = f \circ \psi_{j(x)}^{-1} \circ \pi_{j(x)}$$

is smooth on U_x and satisfies $F_x \circ \psi = f$ on I_x. The family $(U_x)_{x\in\text{supp}(f)}$ forms a cover of the closed set $\psi(\text{supp}(f))$. We can therefore find a locally finite partition of unity $(\Phi_x)_{x\in\text{supp}(f)}$ subordinated to that cover. The function

$$F = \sum_{x\in\text{supp}(f)} F_x \cdot \Phi_x$$

is well-defined and smooth as a locally finite sum of smooth maps. Since $\Phi_x \in \mathcal{D}(U_x)$, the composition $\Phi_x \circ \psi$ has support in $\psi^{-1}(U_x) = I_x$. On I_x we have $F_x \circ \psi = f$ hence $(F_x \cdot \Phi_x) \circ \psi = f \cdot (\Phi_x \circ \psi)$. Outside I_x we also have $(F_x \cdot \Phi_x) \circ \psi = 0 = f \cdot (\Phi_x \circ \psi)$. This implies the required identity $F \circ \psi = f$. $\square$

Basically, proposition 2.16 tells us that we only need to “fill in the blanks” around the critical set. Localizing, we can write f as the sum of functions f_e that have support in arbitrary neighborhoods of $e \in E(\psi)$, and a remainder g that has support outside $E(\psi)$. We know that g is contained in $\mathcal{A}(\psi)$ by the previous proposition and it remains to prove that each f_e is contained in $\mathcal{A}(\psi)$. Using the characterization of the closure of $\mathcal{A}(\psi)$ given in 1.12, we will see that considering a function f_e that is flat in e is sufficient. The proof will rely on the following special case.

Proposition 2.17.
Consider a smooth bijection $\varphi : [a,b] \to [\alpha,\beta]$ which has only one critical point $z \in \{a,b\}$. Suppose furthermore that this critical point has finite order. Then the composition operator $C_\theta : \mathcal{D}([a,b]) \to \mathcal{D}([\alpha,\beta])$ generated by the inverse map $\theta = \varphi^{-1} : [\alpha,\beta] \to [a,b]$ is continuous .

Proof. Note that the inverse map is not smooth itself. Let us first explain why we can suppose that a is the critical point with $\varphi(a) = \alpha$. Indeed, if b is the only critical point we can consider the coordinate transformation $\tau(t) = b + a - t$ which is a diffeomorphism on $\mathbb{R}$ with $\tau^{-1} = \tau$. The map $C_\tau : \mathcal{D}([a,b]) \to \mathcal{D}([a,b])$ is a continuous bijection hence an isomorphism. Since the only critical point of $\varphi\circ\tau$ is $\tau(b) = a$, the map $(\varphi\circ\tau)^{-1}$ generates a continuous composition operator by assumption. Using $\varphi^{-1} = (\varphi\circ\tau\circ\tau)^{-1} = \tau\circ(\varphi\circ\tau)^{-1}$, we obtain that $C_{\varphi^{-1}} = C_{(\varphi\circ\tau)^{-1}} \circ C_\tau$ is also continuous. The same argument with the coordinate transform $\varrho(s) = \alpha + \beta - s$ and the identity $\varrho\circ\varrho\circ\varphi = \varphi$ allows us to suppose $\varphi(a) = \alpha$.

Once we have proven that C_θ is a well-defined map between $\mathcal{D}([a,b])$ and $\mathcal{D}([\alpha,\beta])$, we can deduce the continuity from the closed graph theorem since both spaces are Fréchet and C_θ is continuous with respect to the coarser topologies of uniform convergence.
Since a is the only critical point of φ, the map $\theta = \varphi^{-1}$, and hence every composition $f \circ \theta$, is smooth on $(\alpha, \beta] = (\varphi(a), \beta]$. To prove $f \circ \theta \in \mathcal{D}([\alpha,\beta])$ we only have to prove smoothness in α which we will achieve by induction.
The case $n = 0$ is a consequence of the continuity of θ and we obtain $f \circ \theta(\alpha) = f(a) = 0$. For the induction step, suppose that $\lim_{x \to \alpha}(g \circ \theta)^{(n)}(x) = 0$ for all $g \in \mathcal{D}[a,b]$. The differentiability of $(f \circ \theta)^{(n)}$ at α, together with $(f \circ \theta)^{(n+1)}(\alpha) = 0$, will follow from the chain rule

$$(f \circ \theta)^{(n+1)} = ((f' \circ \theta) \cdot \theta')^{(n)} = \left(\frac{f' \circ \theta}{\varphi' \circ \theta}\right)^{(n)}.$$

Indeed, applying the mean value theorem to compute the difference quotient, one obtains that

$$\frac{(f \circ \theta)^{(n)}(x) - (f \circ \theta)^{(n)}(\alpha)}{x - \alpha} = (f \circ \theta)^{(n+1)}(\xi) = \left(\frac{f' \circ \theta}{\varphi' \circ \theta}\right)^{(n)}(\xi) = \left(\frac{f'}{\varphi'} \circ \theta\right)^{(n)}(\xi)$$

for some $\xi \in (a,b)$. Since $a < \xi < x \to a$, it follows that the right hand side converges to zero by the induction assumption once we have shown that $f'/\varphi' \in \mathcal{D}([\alpha,\beta])$.
Let us now prove that the quotient in question is smooth, which is also best done by induction. The quotient $g = \frac{f'}{\varphi'}$ can extended to a continuous function on $[a,b]$. Indeed, applying l'Hospital's rule sufficiently many times to $g \in \mathcal{D}((a,b])$, one sees that $\lim_{x \to a} g(x) = 0$ as a consequence of the finite order condition. For the induction step, suppose the n-th derivative of g has the form of a quotient $g^{(n)} = \frac{p}{h}$ where $h, p \in \mathcal{D}([a,b])$ and h has only one zero at a which is of finite order. Its derivative on $(a,b]$ is given by $g^{(n+1)} = \frac{P}{H}$ where $P = p' \cdot h - p \cdot h' \in \mathcal{D}([a,b])$ and $H = h^2$ has only one zero, a, which is of finite order (twice the order of h). We obtain $g^{(n+1)}(x) \to 0$ as $x \to a$, and the mean value theorem implies that the difference quotient

$$\frac{g^{(n)}(x) - g^{(n)}(\alpha)}{x - \alpha} = g^{(n+1)}(\xi) \to 0$$

since $a < \xi < x \to a$. $\square$

Let us now continue and prove that we can always write functions which are flat on $E(\psi)$ as compositions. The critical set of ψ being discrete, and since functions with support outside of it are already compositions by proposition 2.16, we can restrict ourselves to functions having support in a sufficiently small neighborhood of only one critical point. If we recall example 2.6, the problem under consideration resulted from the fact that we were not allowed to consider the two branches $\psi((-\infty,0))$ and $\psi((0,\infty))$ of the image $\psi(\mathbb{R})$ separately. This is where the Whitney-regularity comes into play.

Proposition 2.18.
For $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ that is injective, proper, has a Whitney-regular image, and only critical points of finite order, every critical point $e \in E(\psi)$ possesses a neighborhood I_e such that every function $f \in \mathcal{D}(I_e)$ which is flat in e is a contained in $\mathcal{A}(\psi)$.

Proof. Fix $e \in E(\psi)$. Using the finite order condition, we may choose $\varepsilon > 0$ sufficiently small such that some higher derivative of one coordinate ψ_j is zero-free on $I_e = [e-\varepsilon, e+\varepsilon]$. Applying the mean value theorem sufficiently many times, one sees that ψ_j' is also zero-free on $I_e \setminus \{e\}$. The coordinate ψ_j, once restricted to $I^- = [e-\varepsilon, e]$ and $I^+ = [e, e+\varepsilon]$, satisfies the condition of the previous proposition 2.17. Having this in mind, we can decompose any $f \in \mathcal{D}(I_e)$ which is flat in e into the sum $f = f^- + f^+$ of $f^- \in \mathcal{D}(I^-)$ and $f^+ \in \mathcal{D}(I^+)$. We now prove that both f^+ and f^- are contained in $\mathcal{A}(\psi)$. We will do so by showing that they are contained in the respective ranges of the related restrictions $C_\psi^- : X^- \to D(I^-)$ and $C_\psi^+ : X^+ \to \mathcal{D}(I^+)$ of the composition operator C_ψ on the sets

$$X^- = \{F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R}) : F \text{ flat on } \psi(\mathbb{R} \setminus I^-)\}, \text{ and}$$
$$X^+ = \{F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R}) : F \text{ flat on } \psi(\mathbb{R} \setminus I^+)\}.$$

Note that both X^- and X^+ are closed subspaces of $\mathcal{E}(\mathbb{R}^d, \mathbb{R})$, hence Fréchet spaces.
Let $I = [a, b]$ be either the interval I^- or I^+ and denote by $X \in \{X^-, X^+\}$ the respective definition area of the respective restriction $T : X \to \mathcal{D}(I)$ of the composition operator C_ψ. By proposition 2.16 we know that every $f \in \mathcal{D}((a, b))$ is a composition $f = F \circ \psi$, though it has not been stated explicitly that we can choose the support of F to satisfy $\text{supp}(F) \cap \psi(\mathbb{R}) \subseteq \psi((a, b))$. Let us therefore recall our constructive argument to obtain this improvement. By assumption, ψ_j' is zero-free on (a, b) hence $\psi_j : (a, b) \to \psi_j((a, b)) = (\alpha, \beta)$ is a diffeomorphism. The set $\psi(\text{supp}(f))$ is compact and has empty intersection with both closed sets $\pi_j^{-1}(\mathbb{R} \setminus (\alpha, \beta))$ and $\psi(\mathbb{R} \setminus (a, b))$. Considering some $\Phi \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ such that $\Phi = 1$ near $\psi(\text{supp}(f))$ and $\Phi = 0$ near the union L of the sets $\pi_j^{-1}(\mathbb{R} \setminus (\alpha, \beta))$ and $\psi(\mathbb{R} \setminus (a, b))$, we have $f = F \circ \psi$, where $F = \Phi \cdot (f \circ \psi_j^{-1} \circ \pi_j) \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$. The function F has support in L^c, which does not intersect $\psi(\mathbb{R} \setminus (a, b))$, hence $\text{supp}(F) \cap \psi(\mathbb{R}) \subseteq \psi([a, b])$.
This means that $\mathcal{D}((a, b))$ is contained in the range of T. This set is dense in $\mathcal{D}([a, b])$ as a consequence of the theorem of bipolars, since some $u \in \mathcal{E}(\mathbb{R}, \mathbb{R})'$ that vanishes on $\mathcal{D}((a, b))$ has support in $\mathbb{R} \setminus (a, b)$ and therefore vanishes on $\mathcal{D}([a, b])$. The range of the map $T : X \to \mathcal{D}([a, b])$ is therefore dense and it is sufficient to prove that it is also closed to obtain the surjectivity of T. Using the closed range theorem (see [MV97], theorem 26.3), we need to show

$$\text{Ker}(T)^\perp \subseteq \text{Range}(T^t),$$

where $T^t : \mathcal{D}([a, b])' \to X'$ is the transposed operator of T between the continuous duals, and $\text{Ker}(T)^\perp = \{u \in X' : u(F) = 0 \text{ for all } F \circ \psi = 0\}$ is the annihilator of the kernel of T.
Fix $u \in \text{Ker}(T)^\perp$, which we can extend to some $\tilde{u} \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})'$ via the Hahn-Banach theorem. The support of $\tilde{u}$ is contained in $\psi(\mathbb{R})$, since for $F \in \mathcal{D}(\psi(\mathbb{R})^c)$ we have $F \in X$ as well as $F \in \text{Ker}(T)$ hence $\tilde{u}(F) = u(F) = 0$. To obtain the required $v \in X'$ with

$v(F\circ\psi) = u(F)$, let us define v on the dense subset $\mathcal{D}((a,b))$ of $\mathcal{D}([a,b])$. As previously seen, we can write an arbitrary $f \in \mathcal{D}((a,b))$ as $f = F \circ \psi$ with some $F = \Phi \cdot (f \circ \psi_j^{-1} \circ \pi_j) \in X$. By assigning

$$v(f) = u(\Phi \cdot f \circ \psi_j^{-1} \circ \pi_j) = u(F)$$

whenever $F \in X$ satisfies $F \circ \psi = f$, we get a linear map on $\mathcal{D}((a,b))$ that is well-defined since $u \in \mathrm{Ker}(T)^{\perp}$. We aim at proving the existence of a continuous linear extension of v by using the Hahn-Banach theorem. To do so we need to show that v is continuous on $\mathcal{D}((a,b))$ with respect to the relative topology of $\mathcal{D}([a,b])$. This is the argument that depends on the Whitney regularity. This geometric property of $\psi(\mathbb{R})$ allows us to estimate the extension $\tilde{u}$ of u, which turns out to have support in $\psi(\mathbb{R})$, by seminorms involving only evaluations in $\psi(\mathbb{R})$. To this end fix some compact interval J such that $\mathrm{supp}(\tilde{u}) \subseteq \psi(J)$. The set $\psi(J)$ is connected and Whitney-regular and we can use theorem 2.3.11 in [Hör03] to find some $C > 0$ and $n \in \mathbb{N}$ such that such that

$$|\tilde{u}(F)| \leq C \cdot \|F\|_{\psi(J),n} = C \cdot \sup\{|D^{\alpha}F(\psi(t))| : |\alpha| \leq n, t \in J\}.$$

For our factorization $F = \Phi \cdot (f \circ \psi_j^{-1} \circ \pi_j)$ of f we can use that $\mathrm{supp}(\Phi) \cap \psi(\mathbb{R}) \subseteq \psi([a,b])$ by construction to obtain

$$|\tilde{u}(F)| \leq C \cdot \|\Phi \cdot (f \circ \psi_j^{-1} \circ \pi_j)\|_{\psi(J),n} = C \cdot \|\Phi \cdot (f \circ \psi_j^{-1} \circ \pi_j)\|_{\psi([a,b]),n}.$$

Moreover, since $\Phi = 1$ near $\psi(\mathrm{supp}(f))$, all derivatives of $\Phi \cdot (f \circ \psi_j^{-1} \circ \pi_j)$ and $f \circ \psi_j^{-1} \circ \pi_j$ coincide on a neighborhood of $\psi(\mathrm{supp}(f))$. Also both functions vanish on $\psi([a,b] \setminus \mathrm{supp}(f))$ since the derivatives of $f \circ \psi_j^{-1} \circ \pi_j$ are zero there. Combining this with the previous estimate, we obtain

$$|\tilde{u}(\Phi \cdot (f \circ \psi_j^{-1} \circ \pi_j))| \leq C \cdot \|\Phi \cdot (f \circ \psi_j^{-1} \circ \pi_j)\|_{\psi([a,b]),n} = C \cdot \|f \circ \psi_j^{-1} \circ \pi_j\|_{\psi([a,b]),n}.$$

The composition with ψ_j^{-1} is a continuous map by 2.17 and so is the one induced by the projection π_j. This implies that we can further estimate the norm on the right-hand side by some seminorm $\|f\|_{K,m}$ with some fixed compact set $K \subseteq \mathbb{R}$ and $m \in \mathbb{N}$, leading to the required continuity estimate.

It remains to show that the continuous linear extension $\tilde{v}$ satisfies $T^t(\tilde{v}) = u$. For any function $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ with $\mathrm{supp}(F \circ \psi) \subseteq (a,b)$ this is clear by construction as we already have $u(F) = v(F \circ \psi)$. To generalize this, we will use the continuity of the composition operator and show that the subspace $L = \{F \in X : F \circ \psi \in \mathcal{D}((a,b))\}$ is dense in X. By the theorem of bipolars, it is sufficient to prove $w = 0$ on X whenever $w \in \mathcal{E}'(\mathbb{R}^d, \mathbb{R})$ vanishes on L. Now consider an arbitrary $w \in L^{\perp}$. We will show that $\mathrm{supp}(w) \subseteq \psi(\mathbb{R} \setminus (a,b))$. The inclusion $\mathrm{supp}(w) \subseteq \psi(\mathbb{R})$ is clear since $\psi(\mathbb{R})$ is closed and hence we can find for each $x \in \psi(\mathbb{R})^c$ some open neighborhood $U_x \subseteq \psi(\mathbb{R})^c$. Any $G \in \mathcal{D}(U_x)$ satisfies both $G \in X$ and $\mathrm{supp}(G \circ \psi) = \emptyset \subseteq (a,b)$ hence $w(F) = 0$ and $x \notin \mathrm{supp}(w)$. For $x = \psi(t) \in \psi((a,b))$, the sets $\{x\}$ and $\psi(\mathbb{R} \setminus (a,b))$ are closed and disjoint. Separating them by open neighborhoods, we obtain again an open neighborhood U_x of x such that $U_x \cap \psi(\mathbb{R} \setminus (a,b)) = \emptyset$. Every $G \in \mathcal{D}(U_x)$ satisfies both $G \in X$ and $G \circ \psi \in \mathcal{D}(a,b)$ hence $G \in L$ and $w(G) = 0$. It

follows from theorem 2.3.3. in Hörmander's book that $w(H) = 0$ whenever H is flat on $\operatorname{supp}(w)$ hence $w|_X = 0$.
The verification of the identity $T^t(\tilde{v}) = u$ on X then follows from the continuity of the composition operator. For $F \in X$ we can find a sequence $(F_n)_{n\in\mathbb{N}}$ in L converging to F which implies

$$u(F) = \lim_{n\to\infty} u(F_n) = \lim_{n\to\infty} T^t(\tilde{v})(F_n) = T^t(\tilde{v})(F).$$

□

All we have so far are local solutions around and outside the critical set. Let us now elaborate on how to reduce the general case to the one of a flat function in $E(\psi)$ and merge the obtained local solutions.

Theorem 2.19.
For an injective $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ such that ψ

(i) is a proper map,

(ii) has a Whitney-regular image,

(iii) has only critical points of finite order,

the composition algebra $\mathcal{A}(\psi) = \{F \circ \psi : F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})\}$ is closed in $\mathcal{E}(\mathbb{R}, \mathbb{R})$.

Proof. Fix $f \in \overline{\mathcal{A}(\psi)}$. Theorem 1.12 states that, for each $e \in E(\psi)$, we can find some $F_e \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ such that $f - F_e \circ \psi$ is flat in e. According to the previous proposition 2.18 we can find, for each $e \in E(\psi)$, some open interval $I_e = (e-\varepsilon, e+\varepsilon)$ such that every function $g \in \mathcal{D}(I_e)$ that is flat in e can be written as a composition $G_e \circ \psi$ with some G_e being flat in $\psi(\mathbb{R} \setminus I_e)$. The set $E(\psi)$ being discrete, we can reduce $\varepsilon = \varepsilon(e)$ so that the intervals I_e are pairwise disjoint. The properness implies that images of open sets in $\mathbb{R}$ are open in $\psi(\mathbb{R})$ hence we can find $U_e \subseteq \mathbb{R}^d$, also pairwise disjoint, such that $U_e \cap \psi(\mathbb{R}) = \psi(I_e)$. Consider $\Phi_e \in \mathcal{D}(U_e)$ such that $\Phi_e = 1$ on a neighborhood of $\psi(e)$. The function $(f - F_e \circ \psi) \cdot (\Phi_e \circ \psi)$ is flat in e and its support is contained in $\operatorname{supp}(\Phi_e \circ \psi) \subseteq I_e$. Proposition 2.18 implies that we can find some $G_e \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ such that $(f - F_e \circ \psi) \cdot (\Phi_e \circ \psi) = G_e \circ \psi$. Multiplying G_e with a cutoff function $\Gamma_e \in \mathcal{D}(U_e)$ that is constant 1 near $\operatorname{supp}(\Phi_e)$ we can suppose $G_e \in \mathcal{D}(U_e)$. Both functions

$$F_{E(\psi)} = \sum_{e\in E(\psi)} F_e \cdot \Phi_e \quad \text{and} \quad G_{E(\psi)} = \sum_{e\in E(\psi)} G_e$$

are well defined and smooth since the supports of the respective summands are contained in U_e which are pairwise disjoint. The resulting sums are therefore locally finite.
The support of $f - F_{E(\psi)} \circ \psi - G_{E(\psi)} \circ \psi$ is contained in $\mathbb{R} \setminus E(\psi)$. Therefore, the function can be written as $H \circ \psi$ for some $H \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ by proposition 2.16. Hence we obtain

$$f = (H + F_{E(\psi)} + G_{E(\psi)}) \circ \psi \in \mathcal{A}(\psi),$$

proving the proposition.

□

Taking a look back at the necessary conditions, we can even simplify the characterization to obtain the following result.

Proposition 2.20.
An injective smooth curve $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ generates a closed composition algebra if and only if ψ^{-1} is a locally Hölder continuous map.

Proof. The necessity of the local Hölder continuity has already been proven in proposition 2.14. Now suppose the inverse map to be locally Hölder continuous. By remark 2.15, ψ must be proper, have a Whitney-regular image, and have only critical points of finite order hence the algebra is closed by proposition 2.19. □

As a byproduct we obtain a characterization of smooth injective curves with Hölder continuous inverses.

Corollary 2.21.
An injective curve $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ has a locally Hölder continuous inverse if and only if ψ is proper, has only critical points of finite order, and $\psi(\mathbb{R})$ is a Whitney-regular set.

Our next goal is to give necessary and sufficient conditions for some smooth and injective map $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ to generate a closed composition algebra. The proof of proposition 2.19 relied heavily on estimating the norm of a distribution u with support in the image $\psi(\mathbb{R})$ and this is our best shot at attempting to achieve a similar result in the case of several variables. Unfortunately the usual topology on the space of smooth functions (and of distributions), namely the one generated by the seminorms

$$\|f\|_{K,n} = \sup\{|D^\alpha f(x)| : x \in K, |\alpha| \leq n\}$$

is too rigid to be easily adapted to the rather complicated geometrical structure of the manifold $\psi(\mathbb{R}^q \setminus E(\psi))$. We can however circumvent this by considering directional derivatives rather then partial ones. Of course this is just an equivalent view on differential calculus, as partial derivatives can be easily computed from directional ones and vice versa. The following section will deal with the notions of higher derivatives as multilinear maps as well as give some explicit formula for higher derivatives of compositions. Though the computation will be lengthy and tedious, the notation introduced in definition 3.1 of the following section is crucial for understanding the techniques and results of section 4 and 5.

Chapter 3

A multidimensional version of Faà di Bruno's formula and estimates for the higher derivatives of an inverse map

In this section we give a complete proof of a version of Faà di Bruno's formula for higher derivatives of compositions of maps of several variables. We require this explicit formula to estimate the seminorms $\|\psi^{-1}\|_{K,n}$ of the inverse of a diffeomorphism ψ. This estimate will replace the argument given by proposition 2.17 in the case of several variables.
As usual, we view the total derivative of a map $f : X \to Y$ between two Banach spaces X and Y as the map $f' : X \to L(X, Y)$ associating to $x \in X$ the linear map $r \mapsto f'(x)[r]$ where

$$f'(x)[r] = \lim_{0 \neq t \to 0} \frac{f(x + tr) - f(x)}{t}$$

denotes the directional derivative at x along the direction $r \in X$.
Technically, the second order derivative would be a function $f'' : X \to L(X, L(X, Y))$ and we can identify $f''(x)$ with the 2-linear map $[r, s] \mapsto f''(x)[r][s]$ for which we simply write $f''(x)[r, s]$. Analogously, we view higher order derivatives as maps $f^{(n)} : X \to \mathcal{M}_n(X, Y)$ where $\mathcal{M}_n(X, Y)$ denotes the set of all n-linear maps from X to Y. In this context a (n-dimensional) multidirection is just an element $\mathfrak{r} = (r_1, ..., r_n) \in X^n$. We require some notations in order to simplify the rather technical computations that lie ahead of us.

Definition 3.1.
Let A be a set.

(i) We recall that a collection $\mathcal{P}$ of nonempty subsets of A is called a partition of A if $\bigcup \mathcal{P} = A$ and for $P_1, P_2 \in \mathcal{P}$ we have either $P_1 = P_2$ or $P_1 \cap P_2 = \emptyset$. We write $\mathfrak{P}(A)$ for the set of all partitions of A and abbreviate $\mathfrak{P}(n) = \mathfrak{P}(\{1, ..., n\})$.

(ii) For two multidirections $\mathfrak{r} = (r_1, ..., r_k) \in A^k$ and $\mathfrak{s} = (s_1, ..., s_\ell) \in A^\ell$ we define the (juxtaposition) operations

$$(r_1, ..., r_k) \boxtimes (s_1, ..., s_\ell) = (r_1, ..., r_k, s_1, .., s_\ell) \qquad \text{and} \qquad \mathop{\boxtimes}_{j=1}^{k} r_j = (r_1, ..., r_k).$$

We will always use small standard letters for directions in $r = \mathbb{R}^d$ and small fraktur letters for multidirections. In this setting we will sometimes use the juxtaposition operation indexed over the empty set. In order to avoid any misinterpretation, we define for a multidirection $\mathfrak{r} = (r_1, ..., r_n)$

$$\emptyset \boxtimes \mathfrak{r} = \mathfrak{r} \boxtimes \emptyset = \mathfrak{r} \qquad \text{and} \qquad \mathop{\boxtimes}_{j \in \emptyset} s_j = \emptyset.$$

For directions $s_1, .., s_k$, we will also abbreviate

$$\mathop{\boxtimes}_{j=1}^{k} s_j \boxtimes \mathfrak{r} = \left(\mathop{\boxtimes}_{j=1}^{k} s_j \right) \boxtimes \mathfrak{r} \quad \text{and} \quad \mathfrak{r} \mathop{\boxtimes}_{j=1}^{k} s_j = \mathfrak{r} \boxtimes \left(\mathop{\boxtimes}_{j=1}^{k} s_j \right).$$

(iii) Consider some multidirection $\mathfrak{r} = (r_1, ..., r_n) \in A^n$ and some subset $P \subseteq \{1, .., n\}$. We can suppose P to be ordered, meaning $P = \{p_1, ..., p_k\}$ where $k = |P|$ and $p_j < p_{j+1}$. To further simplify the notation we will write

$$\mathfrak{r}_P = \mathop{\boxtimes}_{P} \mathfrak{r} = (r_{p_1}, ..., r_{p_k})$$

for the P-selection of directions in $\mathfrak{r}$.

We only require P to be ordered to have a solid definition at hand. However, since we will use $\mathfrak{r}_P$ as the directional argument in the total derivative of some C^n-function, the theorem of Schwarz implies that the actual order of $r_{p_1}, ..., r_{p_k}$ in $\mathfrak{r}_P$ does not matter. Furthermore we will always use parentheses around the argument of the higher derivative $f^{(n)}$ of some function $f \in \mathcal{E}(\mathbb{R}^d, \mathbb{R}^q)$ and square brackets around the multidirection argument of its evaluation $f^{(n)}(x)$ to visually distinguish between nonlinear and multilinear arguments and hopefully gain some clarity.
Let us give the main result of this chapter, a higher dimension version of Faà di Bruno's formula for higher order derivatives of a composition.

Proposition 3.2.
Consider two smooth functions $\psi \in \mathcal{E}(\mathbb{R}^d, \mathbb{R}^m)$ and $\theta \in \mathcal{E}(\mathbb{R}^m, \mathbb{R}^q)$. The composition $\theta \circ \psi$ is contained in $\mathcal{E}(\mathbb{R}^d, \mathbb{R}^q)$ and for $\mathfrak{r} = (r_1, ..., r_n)$ the n-th derivative is given by the formula

$$(\theta \circ \psi)^{(n)}(x)[\mathfrak{r}] = \sum_{\mathcal{P} \in \mathfrak{P}(n)} \theta^{(|\mathcal{P}|)}(\psi(x)) \left[\mathop{\boxtimes}_{P \in \mathcal{P}} \psi^{(|P|)}(x)[\mathfrak{r}_P] \right].$$

A related expression can already be found in [Gla58], formula (9.5) in the following alternative form

$$D_x^n(\theta \circ \psi)^{(n)} = \sum \binom{\mu_1, ..., \mu_k}{n} D^\ell_{\psi(x)}\theta[T_{\mu_1,...,\mu_q}],$$

where the sum ranges over all muti-indices $\mu_1, ..., \mu_q \in \mathbb{N}_0$ satisfying $\mu_1+2\mu_2+\cdots+q\mu_q = n$ and $\ell = \mu_1 + \cdots + \mu_q$. The symbol $T_{\mu_1,...,\mu_q}$ stands for a symmetric tensor product. Glaeser states that the proof is done by copying the reasoning required to prove exercise 7 of §3 in [Bou04] which states a similar formula when ψ is a function of a real variable and θ was vector-valued. The difference to our approach is that we do not use the symmetry which, in our opinion, makes the formula easier to read and to apply for concrete computations. For reasons of completeness we shall nonetheless give an independent proof.
Since we view higher order derivatives as maps $F^{(n)}$ with values in the space $\mathcal{M}_n(X,Y)$ of n-linear maps, we shall begin with useful and well-known tools concerning multilinear maps in general and derivatives in particular.

Lemma 3.3.
Let $T \in \mathcal{M}_n(X,Y)$ be a multilinear map between the normed spaces X and Y and fix $x_j, z_j \in X$ and $j \leq n$.

(i) The difference $T[x_1, ..., x_n] - T[z_1, ..., z_n]$ is given by

$$T\left[\mathop{\boxtimes}_{j=1}^{n} x_j\right] - T\left[\mathop{\boxtimes}_{j=1}^{n} z_j\right] = \sum_{\ell=1}^{n} T\left[\mathop{\boxtimes}_{j=1}^{\ell-1} x_j \boxtimes (x_\ell - z_\ell) \mathop{\boxtimes}_{j=\ell+1}^{n} z_j\right].$$

(ii) If $|||T||| = \sup\{\|T[r_1, ..., r_n]\|_Y : r_1, .., r_n \in X, \|r_j\|_X \leq 1\}$ denotes the operator norm of T, one has the estimate

$$\|T[x_1, .., x_n]\|_Y \leq |||T||| \cdot \prod_{j=1}^{n} \|x_j\|_X.$$

(iii) For $F \in \mathcal{E}(\mathbb{R}^m, \mathbb{R}^k)$ the map $(y, r_1, ..., r_n) \mapsto F^{(n)}(y)[r_1, ..., r_n]$ is continuous. More specifically we have

$$F^{(n)}(y)[\varrho_1, ..., \varrho_n] \to F^{(n)}(x)[r_1, ..., r_n]$$

whenever $y \to x$ and $(\varrho_1, ..., \varrho_n) \to (r_1, ..., r_n)$.

Proof. The first part is best proved by induction where the case $n = 1$ follows directly from the linearity of T. The induction step $n-1 \to n$ is given by decomposing the difference

$$\begin{aligned} T\left[\mathop{\boxtimes}_{j=1}^{n} x_j\right] - T\left[\mathop{\boxtimes}_{j=1}^{n} z_j\right] =& T\left[\mathop{\boxtimes}_{j=1}^{n} x_j\right] - T\left[\mathop{\boxtimes}_{j=1}^{n-1} x_j \boxtimes z_n\right] \\ &+ T\left[\mathop{\boxtimes}_{j=1}^{n-1} x_j \boxtimes z_n\right] - T\left[\mathop{\boxtimes}_{j=1}^{n-1} z_j \boxtimes z_n\right] \end{aligned}$$

and using the induction assumption on the $n-1$-linear map $(r_1, ..., r_{n-1}) \mapsto T[r_1, ..., r_{n-1}, z_n]$ to obtain

$$\begin{aligned} T\left[\mathop{\boxtimes}_{j=1}^{n} x_j\right] - T\left[\mathop{\boxtimes}_{j=1}^{n} z_j\right] =& T\left[\mathop{\boxtimes}_{j=1}^{n-1} x_j \boxtimes (x_n - z_n)\right] \\ &+ \sum_{\ell=1}^{n-1} T\left[\mathop{\boxtimes}_{j=1}^{\ell-1} x_j \boxtimes (x_\ell - z_\ell) \mathop{\boxtimes}_{j=\ell+1}^{n-1} z_j \boxtimes z_n\right] \\ =& \sum_{\ell=1}^{n} T\left[\mathop{\boxtimes}_{j=1}^{\ell-1} x_j \boxtimes (x_\ell - z_\ell) \mathop{\boxtimes}_{j=\ell+1}^{n} z_j\right]. \end{aligned}$$

The proof of the second statement is trivial: If any argument x_j is zero, $T[x_1, ..., x_n]$ vanishes and the inequality is true. If not, the multilinearity implies

$$T[x_1, ..., x_n] = T\left[\|x_1\| \frac{x_1}{\|x_1\|}, ..., \|x_n\| \frac{x_n}{\|x_n\|}\right] = T\left[\frac{x_1}{\|x_1\|}, ..., \frac{x_n}{\|x_n\|}\right] \cdot \prod_{j=1}^{n} \|x_j\|,$$

hence the required inequality since the X-norm of the arguments of T on the right hand side are inferior or equal to 1. Note that we did not require $|||T|||$ to be finite. This is however always the case if we consider X and Y to be finite-dimensional spaces, as a multilinear map is continuous and the product of unit balls is compact.
For the third part, we first consider the case $n = 1$ and fix $x, r \in \mathbb{R}^m$. For $y, \varrho \in \mathbb{R}^m$ we have

$$\begin{aligned} \|F'(x)[r] - F'(y)[\varrho]\| &\leq \|F'(x)[r - \varrho]\| + \|F'(x)[\varrho] - F'(y)[\varrho]\| \\ &\leq |||F'(x)||| \cdot \|r - \varrho\| + |||F'(x) - F'(y)||| \cdot \|\varrho\|. \end{aligned}$$

The first summand on the right hand side converges to zero since $|||F'(x)||| < \infty$ and $\varrho \to r$. To obtain that the second half also converges to zero, it is sufficient to show that $|||F'(x) - F'(y)||| \to 0$, since ϱ converges to r and is therefore bounded. To prove this we consider the other seminorm $\|T\| = \sup\{\|T[z]\| : z \in \{e_1, ..., e_m\}\}$ on the space $L(\mathbb{R}^m, \mathbb{R}^n)$. This seminorm obviously separates points. It is therefore a norm and since the dimension of $L(\mathbb{R}^m, \mathbb{R}^n)$ is finite, we obtain some constant $C > 0$ such that $|||T||| \leq C \cdot \|T\|$ for all $T \in L(\mathbb{R}^m, \mathbb{R}^n)$. This implies $|||F'(x) - F'(y)||| \leq C \cdot \|F'(x) - F'(y)\| \to 0$ since F' is continuous. For the induction step consider

$$F^{(n+1)}(y)[\varrho_1, ..., \varrho_{n+1}] = F^{(n+1)}(y)[\varrho_1, ..., \varrho_n, \varrho_{n+1} - r_{n+1}] + F^{(n+1)}(y)[\varrho_1, ..., \varrho_n, r_{n+1}].$$

The first summand of the right hand side can be dominated by

$$\|\varrho_{n+1} - r_{n+1}\| \cdot |||F^{(n+1)}(y)||| \cdot \prod_{j=1}^{n} \|\varrho_j\|,$$

which converges to zero since the norms $|||F^{(n+1)}(y)|||$ and $\|\varrho_j\|$ are bounded when $y \to x$ and $\varrho_j \to r_j$. The second summand of the right hand side is the multidirectional derivative

$g^{(n)}(y)[\varrho_1, ..., \varrho_n]$ of the map defined by $g(z) = F'(z)[r_{n+1}]$. Applying the induction assumption, we obtain that $g^{(n)}(y)[\varrho_1, ..., \varrho_n]$ converges to $g^{(n)}(x)[r_1, ..., r_n] = F^{(n+1)}(x)[r_1, ..., r_{n+1}]$ for $y \to x$ and $\varrho_j \to \varrho$. $\square$

In order to simplify the computation and use induction, we will require the following result for the higher derivatives:

Lemma 3.4.
Consider continuously differentiable functions $F \in C^k(\mathbb{R}^m, \mathbb{R}^q)$ and $\varrho_j \in C^1(\mathbb{R}^d, \mathbb{R}^m)$ for $1 \le j \le k$. The map defined by

$$f(x) = F^{(k-1)}(\varrho_k(x)) \left[\mathop{\boxtimes}_{j=1}^{k-1} \varrho_j(x) \right]$$

is also continuously differentiable and its directional derivative $f'(x)[r]$ along $r \in \mathbb{R}^d$ is given by

$$\begin{aligned} f'(x)[r] = &\sum_{\ell=1}^{k} F^{(k-1)}(\varrho_k(x)) \left[\mathop{\boxtimes}_{\substack{j=1 \\ j\neq\ell}}^{k-1} \varrho_j(x) \boxtimes \varrho'_\ell(x)[r] \right] \\ &+ \quad F^{(k)}(\varrho_k(x)) \left[\mathop{\boxtimes}_{j=1}^{k-1} \varrho_j(x) \boxtimes \varrho'_k(x)[r] \right]. \end{aligned}$$

Proof. First, let us decompose the difference quotient in order to compute the directional derivative more easily. Take $t \in \mathbb{R} \setminus \{0\}$ then

$$\begin{aligned} \frac{f(x+tr) - f(x)}{t} =& \frac{1}{t} \left(F^{(k-1)}(\varrho_k(x+tr)) \left[\mathop{\boxtimes}_{j=1}^{k-1} \varrho_j(x+tr) \right] - F^{(k-1)}(\varrho_k(x)) \left[\mathop{\boxtimes}_{j=1}^{k-1} \varrho_j(x) \right] \right) \\ =& \frac{1}{t} \left(F^{(k-1)}(\varrho_k(x+tr)) \left[\mathop{\boxtimes}_{j=1}^{k-1} \varrho_j(x+tr) \right] - F^{(k-1)}(\varrho_k(x+tr)) \left[\mathop{\boxtimes}_{j=1}^{k-1} \varrho_j(x) \right] \right) \\ &+ \frac{1}{t} \left(F^{(k-1)}(\varrho_k(x+tr)) \left[\mathop{\boxtimes}_{j=1}^{k} \varrho_j(x) \right] - F^{(k-1)}(\varrho_k(x)) \left[\mathop{\boxtimes}_{j=1}^{k-1} \varrho_j(x) \right] \right). \end{aligned}$$

The first term on the right hand side, namely

$$\frac{1}{t} \left(F^{(k-1)}(\varrho_k(x+tr)) \left[\mathop{\boxtimes}_{j=1}^{k-1} \varrho_j(x+tr) \right] - F^{(k-1)}(\varrho_k(x+tr)) \left[\mathop{\boxtimes}_{j=1}^{k-1} \varrho_j(x) \right] \right)$$

can be rewritten by applying lemma 3.3 (i) to $T = F^{(k-1)}(\varrho_k(x+tr))$ with $z_j = \varrho_j(x+tr)$ and $x_j = \varrho_j(x)$ to obtain

$$\sum_{\ell=1}^{k-1} F^{(k-1)}(\varrho_k(x+tr)) \left[\mathop{\boxtimes}_{j=1}^{\ell-1} \varrho_j(x+tr) \boxtimes \frac{\varrho_\ell(x+tr) - \varrho_\ell(x)}{t} \mathop{\boxtimes}_{j=\ell+1}^{k-1} \varrho_j(x) \right].$$

It remains to study its convergence as t tends to zero. Since the expression in the brackets converges to $\mathop{\boxtimes}\limits_{j=1}^{\ell-1} \varrho_j(x) \boxtimes \varrho'_\ell(x)[r] \mathop{\boxtimes}\limits_{j=\ell+1}^{k-1} \varrho_j(x)$ and $\varrho_k(x+tr) \to \varrho_k(x)$ for $t \to 0$, the previous lemma 3.3 (iii) states that the sum above must converge to

$$\sum_{\ell=1}^{k-1} F^{(k-1)}(\varrho_k(x)) \left[\mathop{\boxtimes}_{j=1}^{\ell-1} \varrho_j(x) \boxtimes \varrho'_\ell(x)[r] \mathop{\boxtimes}_{j=\ell+1}^{k-1} \varrho_j(x) \right].$$

We can rewrite this limit as

$$\sum_{\ell=1}^{k-1} F^{(k-1)}(\varrho_k(x)) \left[\mathop{\boxtimes}_{\substack{j=1 \\ j\neq \ell}}^{k-1} \varrho_j(x) \boxtimes \varrho'_\ell(x)[r] \right]$$

since $F^{(k-1)}(\varrho_k(x))$ is a symmetric $(k-1)$-linear form. This coincides with the first half of the right hand side in our formula.

The second term namely

$$\frac{1}{t}\left(F^{(k-1)}(\varrho_k(x+tr)) - F^{(k-1)}(\varrho_k(x))\right) \left[\mathop{\boxtimes}_{j=1}^{k-1} \varrho_j(x) \right]$$

obviously converges to $F^{(k)}(\varrho_k(x)) \left[\mathop{\boxtimes}\limits_{j=1}^{k-1} \varrho_j(x) \boxtimes \varrho'_k(x)[r] \right]$, which is the second half of the right hand side in the formula above and thus ends the computation.

□

Let us now prove the Faà di Bruno formula for functions of several variables.

Proof of proposition 3.2. Fix $\psi \in \mathcal{E}(\mathbb{R}^d, \mathbb{R}^m)$ and $\theta \in \mathcal{E}(\mathbb{R}^m, \mathbb{R}^q)$.
The proof will be done by induction and the case $n = 1$ is nothing more than the chain-rule. For the induction step $(n-1) \to n$ fix a multidirection $\mathfrak{r} = (r_1, ..., r_n)$, where $r_1, ..., r_n \in \mathbb{R}^d$. We consider the directional derivation operator $D_{r_n} : \mathcal{E}(\mathbb{R}^d, \mathbb{R}^q) \to \mathcal{E}(\mathbb{R}^d, \mathbb{R}^q)$ defined by $D_{r_n} F(x) = F'(x)[r_n]$. Using the induction assumption on $(\theta \circ \psi)^{(n-1)}$, we get

$$\begin{aligned}
(\theta \circ \psi)^{(n)}(x)[\mathfrak{r}] &= (\theta \circ \psi)^{(n)}(x)[r_1, .., r_n] \\
&= D_{r_n}\left((\theta \circ \psi)^{(n-1)}(x)[r_1, ..., r_{n-1}] \right) \\
&= D_{r_n}\left(\sum_{\mathcal{P} \in \mathfrak{P}(n-1)} \theta^{(|\mathcal{P}|)}(\psi(x)) \left[\mathop{\boxtimes}_{P \in \mathcal{P}} \psi^{(|P|)}(x)[\mathfrak{r}_P] \right] \right) \\
&= \sum_{\mathcal{P} \in \mathfrak{P}(n-1)} D_{r_n}\left(\theta^{(|\mathcal{P}|)}(\psi(x)) \left[\mathop{\boxtimes}_{P \in \mathcal{P}} \psi^{(|P|)}(x)[\mathfrak{r}_P] \right] \right).
\end{aligned}$$

Applying lemma 3.4 to the elements of the sum we obtain

$$(\theta \circ \psi)^{(n)}(x)[\mathfrak{r}] = \sum_{\mathcal{P}\in\mathfrak{P}(n-1)} \sum_{Q\in\mathcal{P}} \theta^{(|\mathcal{P}|)}(\psi(x)) \left[\boxtimes_{\substack{P\in\mathcal{P}\\ P\neq Q}} \psi^{(|P|)}(x)[\mathfrak{r}_P] \boxtimes \psi^{(|Q|+1)}(x)\left[\mathfrak{r}_Q \boxtimes r_n\right] \right]$$
$$+ \sum_{\mathcal{P}\in\mathfrak{P}(n-1)} \theta^{(|\mathcal{P}|+1)}(\psi(x)) \left[\boxtimes_{P\in\mathcal{P}} \psi^{(|P|)}(x)[\mathfrak{r}_P] \boxtimes \psi'(x)[r_n] \right].$$

Identifying $\mathcal{P} \in \mathfrak{P}(n-1)$ with the partition $\mathcal{R} = \mathcal{P} \cup \left\{\{n\}\right\}$ in $\mathfrak{P}(n)$, one obtains a bijection between $\mathfrak{P}(n-1)$ and $\{\mathcal{R} \in \mathfrak{P}(n) : \{n\} \in \mathcal{R}\}$ where $|\mathcal{R}| = |\mathcal{P}| + 1$. Using the notation introduced in 3.1 we have

$$\boxtimes_{P\in\mathcal{P}} \psi^{(|P|)}(x)[\mathfrak{r}_P] \boxtimes \psi'(x)[r_n] = \boxtimes_{P\in\mathcal{P}\cup\{n\}} \psi^{(|P|)}(x)\left[\mathfrak{r}_P\right]$$

and the second sum can be written as

$$\sum_{\mathcal{P}\in\mathfrak{P}(n-1)} \theta^{(|\mathcal{P}|+1)}(\psi(x)) \left[\boxtimes_{P\in\mathcal{P}\cup\{n\}} \psi^{(|P|)}(x)[\mathfrak{r}_P] \right] = \underbrace{\sum_{\substack{\mathcal{R}\in\mathfrak{P}(n)\\ \{n\}\in\mathcal{R}}} \theta^{(|\mathcal{R}|)}(\psi(x)) \left[\boxtimes_{P\in\mathcal{R}} \psi^{(|P|)}(x)[\mathfrak{r}_P] \right]}_{(*)}.$$

On the other hand the first (double-)sum can be rewritten by switching the order of summation to obtain

$$\sum_{\substack{Q\subseteq\{1,\dots,n-1\}\\ Q\neq\emptyset}} \sum_{\substack{\mathcal{P}\in\mathfrak{P}(n-1)\\ Q\in\mathcal{P}}} \theta^{(|\mathcal{P}|)}(\psi(x)) \left[\boxtimes_{\substack{P\in\mathcal{P}\\ P\neq Q}} \psi^{(|P|)}(x)[\mathfrak{r}_P] \boxtimes \psi^{(|Q|+1)}(x)\left[\mathfrak{r}_Q \boxtimes r_n\right] \right].$$

By replacing the set $Q \in \mathcal{P}$ with $Q \cup \{n\}$, we obtain a unique corresponding partition $\mathcal{R}$ containing $Q \cup \{n\}$. In this case we even have $|\mathcal{P}| = |\mathcal{R}|$ and we can reformulate the double sum to

$$\sum_{\substack{Q\subseteq\{1,\dots,n-1\}\\ Q\neq\emptyset}} \sum_{\substack{\mathcal{R}\in\mathfrak{P}(n)\\ Q\cup\{n\}\in\mathcal{R}}} \theta^{(|\mathcal{R}|)}(\psi(x)) \left[\boxtimes_{\substack{P\in\mathcal{R}\\ P\neq Q\cup\{n\}}} \psi^{(|P|)}(x)[\mathfrak{r}_P] \boxtimes \psi^{(|Q\cup\{n\}|)}(x)\left[\mathfrak{r}_{Q\cup\{n\}}\right] \right].$$

Shifting the index from $Q \neq \emptyset$ to $R = Q \cup \{n\}$ with both $n \in R$ and $|R| \geq 2$ we obtain

$$\sum_{\substack{R\subseteq\{1,\dots,n\}\\ n\in R, |R|\geq 2}} \sum_{\substack{\mathcal{R}\in\mathfrak{P}(n)\\ R\in\mathcal{R}}} \theta^{(|\mathcal{R}|)}(\psi(x)) \left[\boxtimes_{\substack{P\in\mathcal{R}\\ P\neq R}} \psi^{(|P|)}(x)[\mathfrak{r}_P] \boxtimes \psi^{(|R|)}(x)\left[\mathfrak{r}_R\right] \right]$$
$$= \sum_{\substack{R\subseteq\{1,\dots,n\}\\ n\in R, |R|\geq 2}} \sum_{\substack{\mathcal{R}\in\mathfrak{P}(n)\\ R\in\mathcal{R}}} \theta^{(|\mathcal{R}|)}(\psi(x)) \left[\boxtimes_{P\in\mathcal{R}} \psi^{(|P|)}(x)[\mathfrak{r}_P] \right].$$

Adding this to the previous sum in $(*)$, namely

$$\sum_{\substack{\mathcal{R}\in\mathfrak{P}(n)\\ \{n\}\in\mathcal{R}}} \theta^{(|\mathcal{R}|)}(\psi(x))\left[\boxtimes_{P\in\mathcal{R}}\psi^{(|P|)}(x)[\mathfrak{r}_P]\right] = \sum_{\substack{R\subseteq\{1,...,n\}\\ n\in R,|R|=1}}\sum_{\substack{\mathcal{R}\in\mathfrak{P}(n)\\ R\in\mathcal{R}}} \theta^{(|\mathcal{R}|)}(\psi(x))\left[\boxtimes_{P\in\mathcal{R}}\psi^{(|P|)}(x)[\mathfrak{r}_P]\right]$$

and switching the order of summation again we obtain the desired

$$\begin{aligned}(\theta\circ\psi)^{(n)}(x)[\mathfrak{r}] &= \sum_{\substack{R\subseteq\{1,...,n\}\\ n\in R}}\sum_{\substack{\mathcal{R}\in\mathfrak{P}(n)\\ R\in\mathcal{R}}} \theta^{(|\mathcal{R}|)}(\psi(x))\left[\boxtimes_{P\in\mathcal{R}}\psi^{(|P|)}(x)[\mathfrak{r}_P]\right]\\ &= \sum_{\mathcal{R}\in\mathfrak{P}(n)} \theta^{(|\mathcal{R}|)}(\psi(x))\left[\boxtimes_{P\in\mathcal{R}}\psi^{(|P|)}(x)[\mathfrak{r}_P]\right].\end{aligned}$$

Here the last equality holds because every partition $\mathcal{R}$ of $\{1,...,n\}$ contains precisely one $R\subseteq\{1,...,n\}$ with $n\in R$. □

As we have seen in the previous section, one crucial tool to prove that $\mathcal{A}(\psi)$ is closed (in the one-dimensional case) is the ability to estimate the inverse $\psi_j'(x)^{-1}$ by some power of $\operatorname{dist}(x,E(\psi))^{-1}$. We can now use the explicit formula to obtain a similar estimate in the case of several variables.

Corollary 3.5.
Let $\psi\in\mathcal{E}(\mathbb{R}^d,\mathbb{R}^d)$ be a diffeomorphism between open sets U and V in $\mathbb{R}^d$, and let $\theta=\psi|_U^{-1}$ be its smooth inverse. The derivatives of θ satisfy the estimate

$$\|\theta^{(n)}(\psi(x))\| \le C_n\cdot\left(1+\|\psi\|_{\{x\},n}\right)^{\frac{(n-1)n}{2}}\cdot\left(1+\|\psi'(x)^{-1}\|\right)^{\frac{n(n+1)}{2}}$$

with some constant C_n depending only on n and $\|\psi\|_{\{x\},n}=\sup\{\|\psi^{(k)}(x)\| : k\le n\}$.

Proof. The case $n=1$ simply follows the chain-rule since we have $\theta'(\psi(x))=\psi'(x)^{-1}$ hence $\|\theta'(\psi(x))\|\le\|\psi'(x)^{-1}\|$. Now suppose the inequality to hold for all $0\le k\le n-1$.
Using the fact that $id^{(n)}(x)=0$ for $n\ge 2$, we can apply the formula from proposition 3.2 to the identity $id|_U=\theta\circ\psi|_U$ and some multidirection $\mathfrak{r}=(r_1,...,r_n)$ to obtain for $n\ge 2$

$$0 = id^{(n)}(x)\left[\boxtimes_{j=1}^{n} r_j\right] = \sum_{\mathcal{P}\in\mathfrak{P}(n)} \theta^{(|\mathcal{P}|)}(\psi(x))\left[\boxtimes_{P\in\mathcal{P}}\psi^{(|P|)}(x)[\mathfrak{r}_P]\right].$$

Since all partitions in $\mathfrak{P}(n)$ with the exception of $\mathcal{P}_n=\{\{1\},...,\{n\}\}$ consist of at most $n-1$ subsets of $\{1,...,n\}$, splitting the corresponding term on the right side and carrying it to the left leads to

$$\theta^{(n)}(\psi(x))\left[\boxtimes_{j=1}^{n}\psi'(x)[r_j]\right] = -\sum_{\substack{\mathcal{P}\in\mathfrak{P}(n)\\ |\mathcal{P}|\le n-1}} \theta^{(|\mathcal{P}|)}(\psi(x))\left[\boxtimes_{P\in\mathcal{P}}\psi^{(|P|)}(x)[\mathfrak{r}_P]\right].$$

Using the fact that $\psi'(x)$ is invertible we can write $\varrho_j = \psi'(x)[r_j]$. This implies

$$\theta^{(n)}(\psi(x))\left[\boxtimes_{j=1}^{n}\varrho_j\right] = -\sum_{\substack{\mathcal{P}\in\mathcal{P}(n)\\|\mathcal{P}|\le n-1}} \theta^{(|\mathcal{P}|)}\left(\psi(x)\right)\left[\boxtimes_{P\in\mathcal{P}}\psi^{(|P|)}(x)\left[\boxtimes_{k\in P}\psi'(x)^{-1}\left[\varrho_k\right]\right]\right].$$

Applying estimate (ii) from lemma 3.3, namely $\|T[r_1,...,r_n]\| \le \|\!|T|\!\| \cdot \prod_{j=1}^{n}\|r_j\|$ on the summands on the right hand side we obtain

$$\left\|\theta^{(n)}(\psi(x))\left[\boxtimes_{j=1}^{n}\varrho_j\right]\right\| \le \sum_{\substack{\mathcal{P}\in\mathfrak{P}(n)\\|\mathcal{P}|\le n-1}} \|\!|\theta^{(|\mathcal{P}|)}\left(\psi(x)\right)|\!\| \cdot \prod_{P\in\mathcal{P}}\left\|\psi^{(|P|)}(x)\left[\boxtimes_{k\in P}\psi'(x)^{-1}\left[\varrho_k\right]\right]\right\|.$$

Considering $(\varrho_1,...,\varrho_n)$ to be in the product of the $\mathbb{R}^d$ unit balls and $|P|\le n$ we can further estimate

$$\begin{aligned}
\prod_{P\in\mathcal{P}}\left\|\psi^{(|P|)}(x)\left[\boxtimes_{k\in P}\psi'(x)^{-1}\left[\varrho_k\right]\right]\right\| &\le \prod_{P\in\mathcal{P}}\|\!|\psi^{(|P|)}(x)|\!\|\cdot\|\!|\psi'(x)^{-1}|\!\|^{|P|}\\
&\le \left(1+\|\!|\psi'(x)^{-1}|\!\|\right)^n \cdot \prod_{P\in\mathcal{P}}\|\!|\psi^{(|P|)}(x)|\!\|\\
&\le \left(1+\|\!|\psi'(x)^{-1}|\!\|\right)^n \cdot \|\!|\psi|\!\|_{\{x\},n}^{|\mathcal{P}|}.
\end{aligned}$$

Taking the supremum over all $\{(\varrho_1,...,\varrho_n) : \|\varrho_j\|\le 1 \text{ for all } j\}$, we obtain

$$\begin{aligned}
\|\!|\theta^{(n)}(\psi(x))|\!\| &\le \left(1+\|\!|\psi'(x)^{-1}|\!\|\right)^n\cdot\left(\sum_{\substack{\mathcal{P}\in\mathfrak{P}(n)\\|\mathcal{P}|\le n-1}}\|\!|\theta^{(|\mathcal{P}|)}\left(\psi(x)\right)|\!\|\cdot\|\!|\psi|\!\|_{\{x\},n}^{|\mathcal{P}|}\right)\\
&\le \left(1+\|\!|\psi'(x)^{-1}|\!\|\right)^n\cdot\left(\sum_{k=1}^{n-1}\sum_{\substack{\mathcal{P}\in\mathfrak{P}(n)\\|\mathcal{P}|=k}}\|\!|\theta^{(k)}\left(\psi(x)\right)|\!\|\cdot\|\!|\psi|\!\|_{\{x\},n}^{k}\right).
\end{aligned}$$

We can dominate the sum of all cardinalities of $\{\mathcal{P}\in\mathfrak{P}(n) : |\mathcal{P}| = k\}$ in the previous sum by some common c_n independent of k. The estimate then reads

$$\begin{aligned}
\|\!|\theta^{(n)}(\psi(x))|\!\| &\le c_n\cdot\left(1+\|\!|\psi'(x)^{-1}|\!\|\right)^n\cdot\sum_{k=1}^{n-1}\|\!|\theta^{(k)}\left(\psi(x)\right)|\!\|\cdot\|\!|\psi|\!\|_{\{x\},n}^{k}\\
&\le c_n\cdot\left(1+\|\!|\psi'(x)^{-1}|\!\|\right)^n\cdot\left(1+\|\!|\psi|\!\|_{\{x\},n}\right)^{n-1}\cdot\sum_{j=1}^{n-1}\|\!|\theta^{(k)}\left(\psi(x)\right)|\!\|.
\end{aligned}$$

Using the induction assumption we know that every summand on the right hand side is dominated by

$$C_k\cdot\left(1+\|\!|\psi|\!\|_{\{x\},k}\right)^{\frac{(k-1)k}{2}}\cdot\left(1+\|\!|\psi'(x)^{-1}|\!\|\right)^{\frac{k(k+1)}{2}},$$

hence also by

$$C_k \cdot \left(1 + \|\psi\|_{\{x\},n}\right)^{\frac{(n-2)(n-1)}{2}} \cdot \left(1 + \|\psi'(x)^{-1}\|\right)^{\frac{(n-1)n}{2}}$$

since these terms increase with k. Adding up the corresponding powers and dominating $c_n \cdot \sum_{j=1}^{n-1} C_k$ by some C_n leads to the estimate

$$\|\theta^{(n)}(\psi(x))\| \leq C_n \cdot \left(1 + \|\psi\|_{\{x\},n}\right)^{\frac{(n-1)n}{2}} \cdot \left(1 + \|\psi'(x)^{-1}\|\right)^{\frac{n(n+1)}{2}}$$

as claimed. □

In the previous corollary 3.5 we estimated the norm of the inverse map. There, the first half of the dominating term, namely $\left(1 + \|\psi\|_{\{x\},n}\right)^{\frac{(n-1)n}{2}}$, is rather unproblematic since ψ is smooth and therefore $\|\psi\|_{\{x\},n}$ is bounded on any compact set. We are more interested in gaining information about $\|\psi'(x)^{-1}\|$ by means of ψ itself. This knowledge is given by the following definition.

Definition 3.6.
For a linear map T between two normed spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ we define the lower bound of T by

$$\lambda(T) = \inf\{\|Tx\|_Y : \|x\|_X = 1\}.$$

This definition is so important because it turns out that $1/\lambda(T)$ is actually the norm of T^{-1}.

Proposition 3.7.
For an injective linear $T : X \to Y$, the map $T : X \to \mathrm{Range}(T)$ is bijective and possesses an inverse $T^{-1} : \mathrm{Range}(T) \to X$. In this case we have

$$\|T^{-1}\| = \frac{1}{\lambda(T)}.$$

Proof. Recall that

$$\|T^{-1}\| = \sup\{\|T^{-1}y\| : y \in \mathrm{Range}(T), \|y\| = 1\}.$$

To prove "$\leq$", consider some $y \in \mathrm{Range}(T)$ with $\|y\| = 1$. We can write $y = Tx$ for some $x \in X$ and obtain $1 = \|y\| = \|x\| \cdot \left\|T\left(\frac{x}{\|x\|}\right)\right\| \geq \|x\| \cdot \lambda(T)$ hence

$$\|T^{-1}y\| = \|T^{-1}Tx\| = \|x\| \leq \frac{1}{\lambda(T)}$$

and $\|T^{-1}\| \leq 1/\lambda(T)$.

On the other hand for $\|x\| = 1$ the injectivity of T implies that $T(x) \neq 0$. We obtain the inequality

$$\frac{1}{\|Tx\|} = \frac{\|T^{-1}Tx\|}{\|Tx\|} \leq |||T^{-1}||| \cdot \frac{\|Tx\|}{\|Tx\|} = |||T^{-1}|||$$

and the fact that $\frac{1}{\lambda(T)} = \sup\left\{\frac{1}{\|Tx\|} : \|x\| = 1\right\}$ implies $\dfrac{1}{\lambda(T)} \leq |||T^{-1}|||$. We therefore obtain the claimed equality. $\square$

Chapter 4

A necessary condition in the case of several variables

Our next step will be to find necessary conditions for an injective map $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ to generate a closed composition algebra.
As we have seen in chapter 2, more precisely in proposition 2.9, the semi-properness of ψ is always necessary, regardless of the injectivity of ψ or the dimensions of its domain and codomain. Since we only consider injective maps, this means that $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ must be proper to even have a chance to generate a closed composition algebra.
Moreover, to decide if $\mathcal{A}(\psi) = \{F \circ \psi : F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})\}$ is closed or not, it would be helpful to know when a function f is an element of the closure $\overline{\mathcal{A}(\psi)}$. In general, this is a difficult question. The best result known to us is the characterization of the closure by Tougeron [Tou71] mentioned at the beginning of chapter 1. It gives sufficient conditions for the identity

$$\overline{\mathcal{A}(\psi)} = \{f \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}) : T_x^\infty f \in T_x^\infty \mathcal{A}(\psi) \text{ for all } x \in \mathbb{R}^q\}$$

to hold and is applicable to maps with Hölder continuous inverses.
We are not able to give a full description on the closure of $\mathcal{A}(\psi)$ in the case of several variables. However we can easily generalize proposition 2.16 which states that, given a proper injective curve, every function that is flat on the critical set is already a composition. This is true even if ψ does not match the sufficient conditions mentioned in [Tou71]. To extend this result to the case of an injective map $\psi : \mathbb{R}^q \to \mathbb{R}^d$ we need to adapt the notion of critical points.
We recall that by definition 1.8 the critical set $E(\psi)$, in the case of one variable, consists of all points $x \in \mathbb{R}$ such that $\psi'(x) = 0$. The reason this set is so important is that we can apply the inverse function theorem to some coordinate of ψ around any non-critical point to obtain a local smooth inverse. Obviously, for several variables, one coordinate is not enough to construct a local smooth inverse. However, having this tool in mind, the obvious definition of a non-critical point $x \in \mathbb{R}^q$ would require a full set $J = \{j_1, ..., j_q\}$ of coordinates such that $\psi'_J(x) = (\psi_{j_1}, ..., \psi_{j_q})'(x)$ is invertible. This can of course only be true if $q \leq d$ which directly follows from the injectivity of $\psi : \mathbb{R}^q \to \mathbb{R}^d$. Note that

the linear map $\psi'_J(x)$ is invertible if and only if its matrix representation $\nabla\psi_J(x)$ has full rank. We can even get rid of the specific subset J, since $\nabla\psi_J(x)$ has full rank if and only if $\nabla\psi(x)$ does. Therefore we simply require that the rank of $\nabla\psi(x)$ is q which is equivalent to the fact that the image of the linear map $\psi'(x)$ is a q-dimensional subspace of $\mathbb{R}^d$. We obtain the following definition:

Definition 4.1.
Consider an injective $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$. In analogy to definition 1.8, we write

$$\begin{aligned} E(\psi) &= \{x \in \mathbb{R}^q : \psi'(x) \in L(\mathbb{R}^q, \mathbb{R}^d) \text{ is not injective}\} \\ &= \{x \in \mathbb{R}^q : \dim(\mathrm{Range}(\psi'(x))) < q\} \\ &= \bigcap_{\substack{J \subseteq \{1,..,d\} \\ |J|=q}} \{x \in \mathbb{R}^q : \det(\psi_J(x)') = 0\} \end{aligned}$$

for the set of critical points. Obviously all three definitions coincide and one can easily see that the third set is closed as a finite intersection of continuous preimages of the closed set $\{0\}$.

In the one-dimensional case, i.e. $q = 1$, the definitions 1.8 and 4.1 coincide since the total derivative $\psi'(x) : \mathbb{R} \to \mathbb{R}^q$, as a linear map, is injective if and only if the vector $(\psi'_1(x), ..., \psi'_d(x)) \in \mathbb{R}^d$ describing it does not vanish. Moreover for every $x \notin E(\psi)$ the matrix $\nabla\psi(x) = (\nabla\psi_1(x), ..., \nabla\psi_d(x))$ has full rank and we can choose $J \subseteq \{1, ..., d\}$ with $|J| = q$ such that $\det(\nabla\psi_J(x)) \neq 0$. With this definition at hand we can proceed to prove the multidimensional analogon to proposition 2.16

Proposition 4.2.
For an injective smooth map $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ the following holds:

(i) The set $\mathcal{I}(E(\psi)) = \{g \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}) : g$ is flat in every $x \in E(\psi)\}$ is contained in $\overline{\mathcal{A}(\psi)}$.

(ii) If ψ is proper, every $f \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ with $\mathrm{supp}(f) \cap E(\psi) = \emptyset$ is a composition $f = F \circ \psi$.

Proof. (i) Fix $x \in E(\psi)^c$. By definition 4.1 we can find a full set of coordinates J, meaning $|J| = q$, such that $\psi'_J(x)$ is a linear bijection. The inverse function theorem states the existence of open neighborhoods V_x of x and W_x of $\psi_J(x) \in \mathbb{R}^J$ such that $\psi_J : V_x \to W_x$ is a diffeomorphism. For $f \in \mathcal{D}(V_x)$ we define $F_x : \pi_J^{-1}(W_x) \to \mathbb{R}$ by

$$F_x = f \circ \psi_J^{-1} \circ \pi_J,$$

where π_J denotes the canonical projection $(x_i)_{i \leq d} \mapsto (x_j)_{j \in J}$. Since $\pi_J : \pi_J^{-1}(W_x) \to W_x$ is smooth, as a linear map, and $\psi_J^{-1} \in \mathcal{E}(W_x, V_x)$, the function F_x is smooth on the open set $\pi_J^{-1}(W_x)$. We also obtain the identity

$$F_x \circ \psi(x) = f \circ \psi_J^{-1} \circ \pi_J \circ \psi(x) = f(x)$$

for all $x \in V_x$.
To prove $f \in \overline{\mathcal{A}(\psi)}$ we will use the theorem of bipolars and show that $u(f) = 0$ whenever u vanishes on $\mathcal{A}(\psi)$.
To this end fix some $u \in \mathcal{A}(\psi)^\perp$ and consider a compact set K such that $\operatorname{supp}(u)$ is contained in the interior of K. The set $\operatorname{supp}(f)$ is compact and does not intersect $K \cap V_x^c$ hence the injectivity of ψ implies that both compact sets $\psi(\operatorname{supp}(f))$ and $\psi(K \cap V_x^c)$ are disjoint. Moreover the set $\pi_J^{-1}(W_x^c)$ is closed as the continuous preimage of a closed set and its intersection with $\psi(\operatorname{supp}(f)) \subseteq \psi(V_x) \subseteq \pi_j^{-1}(W_x)$ is empty. By using for instance corollary 1.4.11 from [Hör03], we can find $\Phi \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ such that $\Phi = 1$ near $\psi(\operatorname{supp}(f))$ and $\Phi = 0$ near $\psi(K \cap V_x^c) \cup \pi_j^{-1}(W_x)$ since both sets are closed and disjoint.
The function $F_x \cdot \Phi$ is smooth on $\mathbb{R}^d$. Let us elaborate on why $(F_x \cdot \Phi) \circ \psi = f$ on K.

- For $y \in V_x$ we have $(F_x \cdot \Phi) \circ \psi(y) = f(y) \cdot (\Phi \circ \psi(y))$. Since $\Phi \circ \psi = 1$ on a neighborhood U of $\operatorname{supp}(f)$, we obtain $(F_x \cdot \Phi) \circ \psi = f$ on on U and $(F_x \cdot \Phi) \circ \psi = 0 \cdot \Phi \circ \psi = f$ on $V_x \setminus \operatorname{supp}(f)$.

- For $y \in K \cap V_x^c$ the function $\Phi \circ \psi$ is zero, hence $(F_x \cdot \Phi) \circ \psi(y) = 0 = f(y)$.

Since $f \in \mathcal{D}(V_x)$ is arbitrary and $f = (F_x \cdot \Phi) \circ \psi$ on an open neighborhood of $\operatorname{supp}(u)$ we obtain

$$u(f) = u((F_x \cdot \Phi) \circ \psi) = 0.$$

This implies $\operatorname{supp}(u) \cap V_x = \emptyset$. As $u \in \mathcal{A}(\psi)^\perp$ is also arbitrary every distribution in the annihilator of $\mathcal{A}(\psi)$ must have support contained in $E(\psi)$. By theorem 2.3.3 from [Hör03] we obtain that $u(g) = 0$ whenever $u \in \mathcal{A}(\psi)^\perp$ and $g \in \mathcal{I}(E(\psi))$, hence $\mathcal{I}(E(\psi)) \subseteq \overline{\mathcal{A}(\psi)}$ by the theorem of bipolars.
(ii) In the case of a proper map, we can refine our construction from above to actually achieve $F \circ \psi = f$. Fix $f \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ with support outside $E(\psi)$. For $x \in E(\psi)^c$ we choose $J = J(x)$ as above. Without loss of generality we can suppose V_x to be an open ball. Since ψ is proper and injective, the inverse $\psi^{-1} : \psi(\mathbb{R}^q) \to \mathbb{R}^q$ is continuous by corollary 2.11 if we endow $\psi(\mathbb{R}^q)$ with the relative topology of $\mathbb{R}^d$. A set is thus open in $\psi(\mathbb{R}^q)$ if and only if it is the image of an open subset of $\mathbb{R}^q$. This implies that we can find $U_x \subseteq \mathbb{R}^d$ such that $\psi(V_x) = U_x \cap \psi(\mathbb{R}^q)$. By intersecting U_x with the open set $\pi_{J(x)}^{-1}(W_x)$ we can even suppose $\pi_{J(x)}(U_x) \subseteq W_x$. Therefore the function

$$F_x = f \circ \psi_{J(x)}^{-1} \circ \pi_{J(x)}$$

is smooth on U_x and satisfies $F_x \circ \psi = f$ on V_x. For a locally finite partition of unity $(\Phi_x)_{x \in \operatorname{supp}(f)}$ subordinated to the cover $(U_x)_{x \in \operatorname{supp}(f)}$ of $\operatorname{supp}(f)$ the function

$$F = \sum_{x \in \operatorname{supp}(f)} F_x \cdot \Phi_x$$

is again well-defined and smooth. The summands satisfy $(F_x \cdot \Phi_x) \circ \psi = f \cdot (\Phi_x \circ \psi)$ on $V_x = \psi^{-1}(U_x)$, since $F_x \circ \psi = f$ there. Outside V_x, we also have $(F_x \cdot \Phi_x) \circ \psi = 0 = f \cdot (\Phi_x \circ \psi)$, since $\Phi_x \circ \psi$ has support in $V_x = \psi^{-1}(U_x)$. This implies the required identity $F \circ \psi = f$. $\square$

As a corollary, we can characterize the closure of composition algebras generated by a specific kind of injective maps.

Corollary 4.3.
If the critical set $E(\psi)$ of an injective $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ is discrete, then $\overline{\mathcal{A}(\psi)}$ contains exactly those functions $f \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ satisfying the Taylor condition

$$T_e^\infty f \in T_e^\infty \mathcal{A}(\psi).$$

for all $e \in E(\psi)$

Proof. The necessity of the Taylor condition has been shown in proposition 1.7. As already noted in proposition 4.2 (i), every $f \in \mathcal{D}(E(\psi)^c)$ is contained in $\overline{\mathcal{A}(\psi)}$. Therefore, $\mathcal{A}(\psi)^\perp$ contains only those $u \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})'$ that have support in $E(\psi)$. Since this set is discrete, the support of $u \in \mathcal{A}(\psi)^\perp$ is a finite collection $\{e_1, ..., e_N\}$ of critical points. If $f \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ satisfies the pointwise Taylor condition, we can find a function $F_{e_j} \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ such that $T_{e_j}^\infty f = T_{e_j}^\infty (F_{e_j} \circ \psi)$. Considering functions $\Phi_j \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ such that $\Phi_j = 1$ near $\psi(e_j)$ and $\Phi_j = 0$ near $\psi(E(\psi) \setminus \{e_j\})$, we can localize to obtain

$$u(f) = u\left(\sum_{j=1}^N f \cdot (\Phi_j \circ \psi)\right) = \sum_{j=1}^N u((F_{e_j} \circ \psi) \cdot (\Phi_j \circ \psi)) = 0.$$

This shows $f \in \overline{\mathcal{A}}(\psi)$. □

As we have seen in section 2, the one-dimensional injective case allows for a rather elementary characterization of closed composition algebras, namely those generated by a map with a locally Hölder continuous inverse map. Moreover the proof of the necessity uses the fact that the Hölder continuity is equivalent to the following three conditions altogether:

(i) ψ is a proper map.

(ii) Every critical point has finite order.

(iii) The image $\psi(\mathbb{R})$ is a Whitney-regular set.

The Hölder continuity looks like a promising approach to deal with the case of several variables as it does not depend in any way on the dimensions of the domain of ψ (unlike condition (ii) in the equivalent decomposition).
Let us begin by giving a positive result dealing with a very special case.

Proposition 4.4.
Consider some injective $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ such that the critical set $E(\psi)$ is discrete. If the composition algebra is closed, then ψ must have a locally Hölder-continuous inverse.

Proof. By assumption the composition operator is a continuous linear surjection between Fréchet spaces and hence open. As we have seen in proposition 2.8, this implies that for every compact set $K \subseteq \mathbb{R}^d$ and $n \in \mathbb{N}$ there is a compact set $L \subseteq \mathbb{R}^q$, a differentiation order $k \in \mathbb{N}$ and a constant $c > 0$ such that for every $f \in \mathcal{A}(\psi)$ we can find $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ with $f = F \circ \psi$ and

$$\|F\|_{K,n} \leq \begin{cases} c \cdot \|f\|_{L,k}, & \text{if } \|f\|_{L,k} \neq 0 \\ 1/2, & \text{otherwise} \end{cases}.$$

Now suppose the inverse map $\psi^{-1} : \psi(\mathbb{R}^q) \to \mathbb{R}^q$ fails to be locally Hölder continuous. This means that we can find some $v = \psi(z)$ and sequences $(\zeta_n)_{n\in\mathbb{N}}$ and $(\eta_n)_{n\in\mathbb{N}}$ in $\psi(\mathbb{R}^q)$ with $\|\zeta_n - v\| < 1/n$ and $\|\eta_n - v\| < 1/n$ such that

$$|\psi^{-1}(\zeta_n) - \psi^{-1}(\eta_n)\| > n \cdot \|\zeta_n - \eta_n\|^{\frac{1}{n}}.$$

Writing $\zeta_n = \psi(x_n)$ and $\eta_n = \psi(y_n)$ we obtain the more handy condition

$$\|\psi(x_n) - \psi(y_n)\| < \frac{\|x_n - y_n\|^n}{n^n}.$$

By corollary 2.10, ψ must be proper and remark 2.11 implies that ψ^{-1} is continuous. Therefore both $x_n = \psi^{-1}(\xi_n)$ and $y_n = \psi^{-1}(\eta_n)$ must converge to $z = \psi^{-1}(v)$. Note that $\|x_n - y_n\| > 0$ and we can therefore choose a subsequence such that the directions $r_n = \frac{y_n - x_n}{\|y_n - x_n\|}$ converge to some $r \in \mathbb{R}^q$ with $\|r\| = 1$. Applying the mean value theorem to every coordinate of ψ we obtain some $\xi_n^j = x_n + t(y_n - x_n)$ such that

$$\|\psi_j'(\xi_n^j)[r_n]\| = \left\| \psi_j'(\xi_n^j) \left[\frac{y_n - x_n}{\|y_n - x_n\|} \right] \right\| = \frac{\|\psi_j(x_n) - \psi_j(y_n)\|}{\|x_n - y_n\|} < \frac{\|x_n - y_n\|^{n-1}}{n} \to 0.$$

Since $r_n \to r$ and $\xi_n^j \to x_\infty$, the left hand side must converge to $\psi_j'(x_\infty)[r] = 0$. Thus x_∞ is a critical point. Without loss of generality we can suppose $\|x_n - x_\infty\| \geq \|y_n - x_\infty\|$ and hence $\|x_n - y_n\| \leq 2 \cdot \|x_n - x_\infty\|$.

Now consider again $\varphi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ with support in the open unit ball $B(0, 1)$ in $\mathbb{R}^d$ and satisfying $\varphi(0) = 1$. The functions defined by

$$f_n(z) = \varphi\left(\frac{2}{\|x_n - y_n\|} \cdot (z - x_n) \right)$$

have support in $B\big(x_n, \frac{\|x_n - y_n\|}{2}\big)$ which contains neither x_∞ nor y_n. Since the critical set is discrete, it does not intersect the support of f_n for n sufficiently large. By proposition 4.2 (ii) we can deduce that those f_n belong to $\mathcal{A}(\psi)$. Now we can use proposition 2.8 for a compact set $K \subseteq \mathbb{R}^d$ containing every line segment $\{x_n + t(y_n - x_n) : t \in [0, 1]\}$ and $\tilde{n} = 1$. We obtain some compact $L \subseteq \mathbb{R}^q, k \in \mathbb{N}$, and $c \geq 0$ such that f_n can be factorized as $f_n = F_n \circ \psi$ with $\|F_n\|_{K,1} \leq c \cdot \|f_n\|_{L,k} \leq c \cdot \|f_n\|_{\mathbb{R}^q,k}$. Applying the mean value inequality to F_n we get

$$\begin{aligned} 1 = f_n(x_n) - f_n(y_n) = F_n(\psi(x_n)) - F_n(\psi(y_n)) &\leq \|F\|_{K,1} \cdot \|\psi(x_n) - \psi(y_n)\| \\ &\leq c \cdot \|f_n\|_{L,k} \cdot \|\psi(x_n) - \psi(y_n)\|. \end{aligned}$$

Dominating $\|f_n\|_{L,k}$ by $2^k \cdot \|x_n - y_n\|^{-k} \cdot \|\varphi\|_{\mathbb{R}^q,k}$ we obtain for $C = c \cdot 2^k \cdot \|\varphi\|_{\mathbb{R}^q,k}$ the contradiction

$$1 \leq C \cdot \|\psi(x_n) - \psi(y_n)\| \cdot \|x_n - y_n\|^{-k} < \frac{1}{n^n}\|x_n - y_n\|^{n-k} \to 0.$$

□

Unfortunately, we are not able to give a proof for the rather plausible conjecture that the local Hölder continuity of the inverse is always necessary in order to obtain a closed composition algebra. This is due to the fact that we have no understanding on the geometric properties of the critical set $E(\psi)$ and the resulting behavior of compositions $F \circ \psi$ on it. This is also the reason why we are still unable to give a characterization of the closure of the composition algebra in the setting of several variables. However we can note a rather intriguing fact about locally Hölder continuity, namely a stability of closedness, hinting towards the conjecture that the Hölder condition could be a characterization of closed composition algebras.

Proposition 4.5.
If $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ and $\gamma \in \mathcal{E}(\mathbb{R}, \mathbb{R}^q)$ are smooth and injective and ψ^{-1} is locally Hölder continuous, then $\mathcal{A}(\psi \circ \gamma)$ is closed whenever $\mathcal{A}(\gamma)$ is closed. This means that the composition operator C_ψ maps a closed composition algebra $\mathcal{A}(\gamma)$ into a closed composition algebra $C_\psi(\mathcal{A}(\gamma)) = \mathcal{A}(\psi \circ \gamma)$.

Proof. By assumption, γ is injective and it follows from proposition 2.20 that $\mathcal{A}(\gamma)$ is closed if and only if γ^{-1} is locally Hölder-continuous. Since $\psi \circ \gamma \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ is also injective, we can again apply proposition 2.20 to obtain that $\mathcal{A}(\psi \circ \gamma)$ is closed if and only if its inverse $(\psi \circ \gamma)^{-1} = \gamma^{-1} \circ \psi^{-1}$ is also locally Hölder continuous. It remains to show that the compositions of locally Hölder continuous maps is again locally Hölder continuous. For this purpose consider two locally Hölder continuous maps $F : X \to Y$ and $G : Y \to Z$. For $x \in X$ we can find $\varepsilon_F, c_F, \alpha_F > 0$ such that

$$\|F(v) - F(w)\|_Y \leq c_F \cdot \|v - w\|_X^{\alpha_F}$$

for all $v, w \in B_X(x, \varepsilon_F)$ as well as $\varepsilon_G, c_G, \alpha_G > 0$ such that

$$\|G(s) - G(t)\|_Z \leq C_G \cdot \|s - t\|_Y^{\alpha_G}$$

for all $s, t \in B_Y(F(x), \varepsilon_G)$. Choosing $\varepsilon < \varepsilon_F$ such that $B(x, \varepsilon)$ is contained in the set $F^{-1}(B_Y(F(x), \varepsilon_G))$, which is open by the continuity of F, we can use both inequalities. For $v, w \in B(x, \varepsilon)$ we have $F(v), F(w) \in B_Y(F(x), \varepsilon_G)$, hence

$$\|G \circ F(v) - G \circ F(w)\|_Z \leq c_G \cdot \|F(v) - F(w)\|_Y^{\alpha_G} \leq c_G \cdot c_F^{\alpha_G} \cdot \|v - w\|_X^{\alpha_F \cdot \alpha_G}.$$

□

As a by-product we obtain the following corollary.

Corollary 4.6.
If an injective $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ has a locally Hölder continuous inverse, then every directional map $\psi_{x,r} : \mathbb{R} \to \mathbb{R}^d, t \mapsto \psi(x + tr)$ generates a closed algebra. This implies that the directional maps have only critical points of finite order and Whitney-regular image.

Proof. If $r = 0$, this is obvious as the composition algebra $\mathcal{A}(\psi_{x,r})$ contains only constant functions. If $r \neq 0$, the inverse of the curve defined by $\gamma(t) = x + t \cdot r$ is Hölder continuous since

$$|\gamma^{-1}(x + t \cdot r) - \gamma^1(x + s \cdot r)| = |t - s| = \frac{1}{\|r\|}\|x + t \cdot r - (x + s \cdot r)\|$$

The previous proposition implies that $\mathcal{A}(\psi_{x,r}) = \mathcal{A}(\psi \circ \gamma)$ must be closed. □

Remark 4.7.
Using our notation from chapter 3, more precisely 3.1, we can compute the higher order derivatives of the directional maps and obtain

$$\psi_{x,r}^{(k)}(0) = \psi^{(k)}(x) \left[\underset{j=1}{\overset{k}{\boxtimes}} r \right].$$

If ψ^{-1} is locally Hölder continuous, the directional map $\psi_{x,r}$ has only critical points of finite order. This means that for any direction $r \in \mathbb{R}^q \setminus \{0\}$ we can find some "directional order" $k = k(r)$ such that $\psi^{(k)}(x) \left[\boxtimes_{j=1}^k r\right] \neq 0$, which seems to be a good way to generalize the notion of order to the case of several variables. Moreover the directional orders are bounded in any point $x \in \mathbb{R}^q$. Indeed the map $r \mapsto \psi^{(k)}(x) \left[\boxtimes_{j=1}^k r\right]$ is continuous. Thus we can find a neighborhood U of r such that $\psi^{(k)}(x) \left[\boxtimes_{j=1}^k \varrho\right] \neq 0$ for all $\varrho \in U$ and every directional order at x along $\varrho \in U$ is at most $k(r)$. The boundary of the unit ball being a compact set we can find a finite cover $(U_r)_{r \in F}$ of $\{r \in \mathbb{R}^q : \|r\| = 1\}$ and thus every directional order at x is dominated by $k_x = \max\{k(r) : r \in F\}$.

Proposition 4.4 gives us an easy tool to determine, in the special case of a discrete critical set, which functions ψ are unable to generate a closed composition algebra. We give two examples to illustrate this fact.

Example 4.8.
Consider the map $\varphi : \mathbb{R} \to \mathbb{R}$ defined by $\varphi(z) = \exp\left(-\frac{1}{|z|}\right)$ for $z \neq 0$ and $\varphi(0) = 0$. Let us define the generator $\psi : \mathbb{R}^2 \to \mathbb{R}^3$ by

$$\psi(z) = \begin{pmatrix} |z|^2 \\ z_1 \cdot \varphi(|z|^2) \\ z_2 \cdot \varphi(|z|^2) \end{pmatrix}.$$

It seems rather intuitive that the sharp cusp in the center, as depicted in the following figure 4.1, does not allow for a closed composition algebra. This turns out to be correct as the function ψ has only a discrete critical set, but does not possess a locally Hölder continuous inverse.

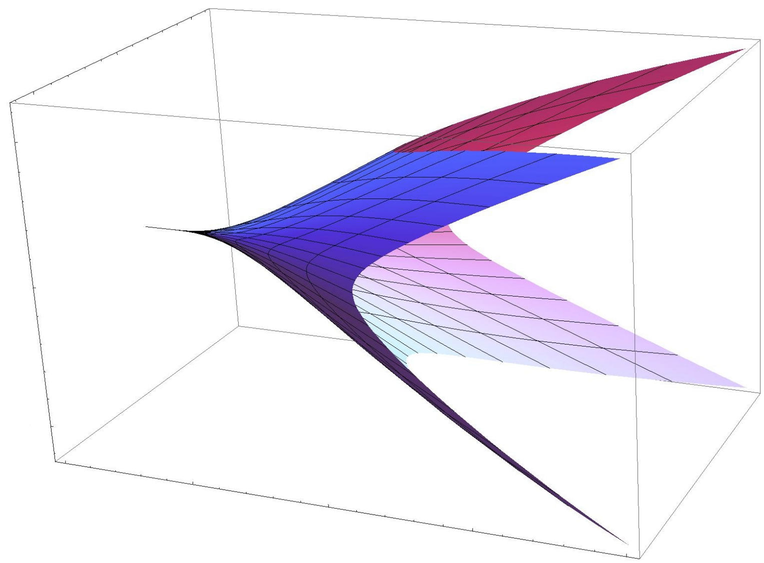


Figure 4.1: The sharp cusp of ψ around $(0,0)$.

The injectivity is obvious since $\psi(z) = \psi(w)$ implies $|z|^2 = |w|^2$. If $|z| = |w| \neq 0$, we can divide the last two coordinates of ψ by $\varphi(|z|^2) = \varphi(|w|^2) \neq 0$ to obtain $z = w$.
We will now use proposition 4.4 and corollary 4.6 to show that $\mathcal{A}(\psi)$ cannot be closed. To apply 4.4 we first need to check that the critical set $E(\psi)$ is discrete.
The matrix representation of the derivative of ψ is given by

$$\nabla\psi(z) = \begin{pmatrix} 2 \cdot z_1 & 2 \cdot z_2 \\ \varphi(|z|^2) + 2 \cdot z_1^2 \cdot \varphi'(|z|^2) & 2 \cdot z_1 \cdot z_2 \cdot \varphi'(|z|^2) \\ 2 \cdot z_1 \cdot z_2 \cdot \varphi'(|z|^2) & \varphi(|z|^2) + 2 \cdot z_2^2 \cdot \varphi'(|z|^2) \end{pmatrix}.$$

Obviously $(0,0)$ is a critical point as $\nabla\psi(0,0) = 0_{\mathbb{R}^{2\times 3}}$ and it will turn out to be the only one. Indeed for $z_1 \neq 0$, the derivative of $\psi_I = (\psi_1, \psi_3)$ is invertible since

$$\begin{aligned} \det(\nabla\psi_I(z)) &= \begin{vmatrix} 2 \cdot z_1 & 2 \cdot z_2 \\ 2 \cdot z_1 \cdot z_2 \cdot \varphi'(|z|^2) & \varphi(|z|^2) + 2 \cdot z_2^2 \cdot \varphi'(|z|^2) \end{vmatrix} \\ &= 2 \cdot z_1 \cdot \varphi(|z|^2) + 4 \cdot z_1 \cdot z_2^2 \cdot \varphi'(|z|^2) - 4 \cdot z_1 \cdot z_2^2 \cdot \varphi'(|z|^2) \\ &= 2 \cdot z_1 \cdot \varphi(|z|^2) \neq 0. \end{aligned}$$

The rank of $\nabla\psi(z)$ is therefore 2 and $\psi'(z)$ is injective.
For $z_1 = 0$ and therefore $z_2 \neq 0$, consider instead $\psi_{II} = (\psi_1, \psi_2)$. An analogue computation

gives us that

$$\begin{aligned}\det(\nabla\psi_{II}(z)) &= \begin{vmatrix} 2\cdot z_1 & 2z_2 \\ \varphi(|z|^2)+2\cdot z_1^2\cdot\varphi'(|z|^2) & 2\cdot z_1\cdot z_2\cdot\varphi'(|z|^2)\end{vmatrix} \\ &= 4\cdot z_1^2\cdot z_2\cdot\varphi'(|z|^2) - 2\cdot z_2\cdot\varphi(|z|^2) - 4\cdot z_1^2\cdot z_2\cdot\varphi'(|z|^2) \\ &= -2\cdot z_2\cdot\varphi(|z|^2)\neq 0\end{aligned}$$

and again the rank of $D\psi(z)$ is 2.
Now suppose $\mathcal{A}(\psi)$ to be closed. By proposition 4.4 the inverse ψ^{-1} is locally Hölder continuous. For $z=\psi(0,0)$ there are $\varepsilon>0$ as well as $c,\alpha>0$ such that

$$\|\psi^{-1}(x)-\psi^{-1}(y)\|\leq c\cdot\|x-y\|^\alpha$$

for all $x,y\in\psi(\mathbb{R}^2)\cap B(z,\varepsilon)$. To see that this can not be the case, we evaluate this estimate in the points $x_n=\psi(1/n,0)$ and $y_n=\psi(-1/n,0)$. We obtain the estimate

$$\frac{2}{n}=\|\psi^{-1}(x_n)-\psi^{-1}(y_n)\|\leq c\cdot\|\psi\left(1/n,0\right)-\psi\left(-1/n,0\right)\|^\alpha = c\cdot\left|\frac{2}{n}\cdot\varphi\left(\frac{1}{n^2}\right)\right|^\alpha$$

and multiplying both sides with n brings the required contradiction $2\leq\frac{c\cdot n}{n^\alpha}\varphi\left(\frac{1}{n^2}\right)^\alpha\to 0$.

Example 4.9.
Consider $\varphi:\mathbb{R}\to\mathbb{R}$ as in example 4.8. We define the map $\psi:\mathbb{R}^2\to\mathbb{R}^3$ by

$$\psi(z)=\begin{pmatrix} z_1\cdot z_2 \\ z_1\cdot\varphi(|z|^2) \\ z_2\cdot\varphi(|z|^2)\end{pmatrix}.$$

To better visualize its shape, we have given a sketch in the figure below.

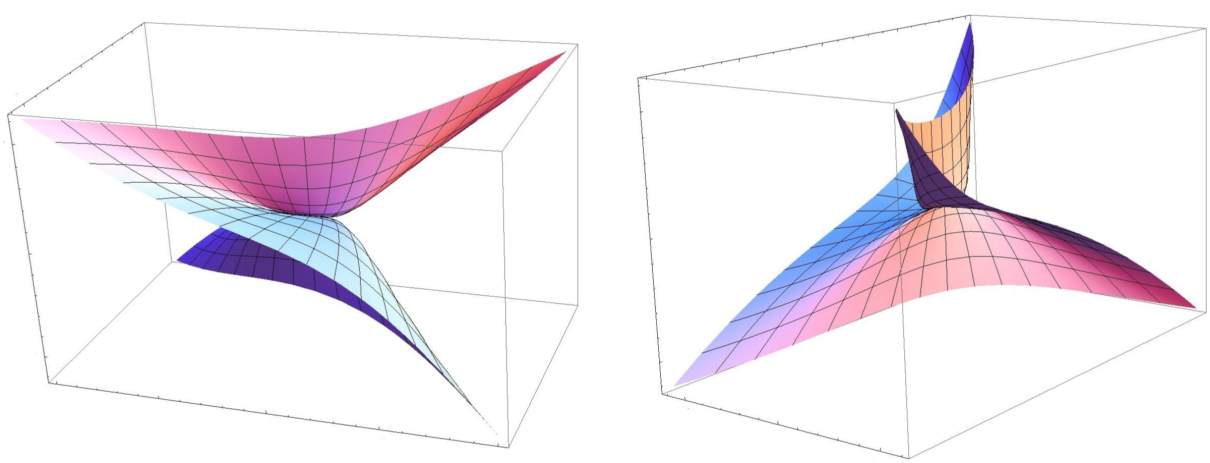


Figure 4.2: The Torsion of ψ around $(0,0)$ from two different angles.

To prove the injectivity, consider $\psi(z) = \psi(w)$ and distinguish the following three cases. If $\psi_2(z) = \psi_2(w) = 0$, we have $z_1 = 0 = w_1$. This implies $z_2^2 = |z|^2$ and $w_2^2 = |w|^2$ hence $z_2 \cdot \varphi(z_2^2) = \psi_3(z) = \psi_3(w) = w_2 \cdot \varphi(w_2^2)$ and the injectivity of $t \mapsto t \cdot \varphi(t)$ on $\mathbb{R}$ gives us the required $z_2 = w_2$.
If $\psi_3(z) = \psi_3(w) = 0$, we have $z_2 = 0 = w_2$ and the same argument applied to ψ_2 gives us the required $z_1 = w_1$.
Elsewise, both $\psi_2(w) = \psi_2(z)$ and $\psi_3(w) = \psi_3(z)$ do not vanish. Therefore we have $z_1, z_2, w_1, w_2 \neq 0$, and hence also $\psi_1(z) = \psi_1(w) \neq 0$. We can multiply the last two coordinates to obtain $z_1 \cdot z_2 \cdot \varphi(|z|^2)^2 = w_1 \cdot w_2 \cdot \varphi(|w|^2)^2$ and dividing both sides by $z_1 \cdot z_2 = \psi_1(z) = \psi_1(w) = w_1 \cdot w_2$ and taking the square root yields $\varphi(|z|^2) = \varphi(|w|^2)$. As in the previous example, we can divide ψ_2 and ψ_3 by $\varphi(|z|^2) = \varphi(|w|^2)$ to see that $z = w$.
To prove that $\mathcal{A}(\psi)$ cannot be closed we will again use proposition 4.4 in combination with corollary 4.6. To do so, we need once more to show that the critical set is discrete. It will turn out that the only critical point is $z_0 = (0, 0)$. Computing the derivative we obtain

$$\nabla\psi(z) = \begin{pmatrix} z_2 & z_1 \\ \varphi(|z|^2) + 2z_1^2 \cdot \varphi'(|z|^2) & 2z_1 \cdot z_2 \cdot \varphi'(|z|^2) \\ 2z_1 \cdot z_2 \cdot \varphi'(|z|^2) & \varphi(|z|^2) + 2z_2^2 \cdot \varphi'(|z|^2) \end{pmatrix}.$$

Evaluating this expression for $z = 0$, we see that the origin is indeed a critical point and it remains to prove that it is the only one. Consider some point $z \neq 0$ and $\lambda, \mu \in \mathbb{R}$ such that $\lambda D_1\psi(z) = \mu D_2\psi(z)$. We obtain $\lambda z_1 = \mu z_2$ from the first coordinate ψ_1 and inserting this identity into the second coordinate we get

$$\lambda\varphi(|z|^2) + 2\lambda z_1^2 \cdot \varphi'(|z|^2) = \lambda D_2\psi_3(z) = \mu D_2\psi_3(z) = \mu 2z_1 \cdot z_2 \cdot \varphi'(|z|^2) = \lambda 2z_1^2\varphi'(|z|^2),$$

hence $\lambda\varphi(|z|^2) = 0$. Since $z \neq 0$, this implies $\lambda = 0$ and the same computation for $\lambda D_1\psi_3(z) = \mu D_2\psi_3(z)$ gives us $\mu = 0$.
Suppose that $\mathcal{A}(\psi)$ is closed. Proposition 4.4 states that the inverse of ψ must be locally Hölder continuous and corollary 4.6 implies that every directional map generates a closed composition algebra. To obtain a contradiction consider the direction $r = (1, 0)$. We obtain the directional map

$$\psi_{0,r}(t) = \psi\begin{pmatrix} t \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ t \cdot \varphi(t^2) \\ 0 \end{pmatrix}$$

and $\psi_{0,r}$ is flat in $t = 0$ which is prohibited by proposition 2.13

Although we are not able to prove that the Hölder condition is necessary, studying the proof of the example 4.4 we can prove a similar necessary condition. The key argument for evaluating the condition $\|\psi(x_n) - \psi(y_n)\| \leq \frac{1}{n}\|x_n - y_n\|^n$ was the construction of the functions $f_n \in \mathcal{A}(\psi)$, for which we could give precise estimates for the norms $\|f_n\|_{L,k}$. The proof that $f_n \in \mathcal{A}(\psi)$ simply relied on proposition 4.2 and was done by showing $\operatorname{supp}(f_n) \cap E(\psi) = \emptyset$. Unfortunately we can not apply this construction and the related estimates without further knowledge of the exact shape of the critical set. However, we can give the following necessary condition, which is a generalization of proposition 2.12, for an injective smooth map $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ which does not require any knowledge about the structure of $E(\psi)$.

Definition 4.10.
The function $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ is said to satisfy the (local) lower distance estimate if for any compact set $K \subseteq \mathbb{R}^q$ there is a $c_K > 0$ and "order" $\alpha_K \in \mathbb{N}$ such that

$$\|\psi(x) - \psi(y)\| \geq c_K \cdot \|x - y\| \cdot \max\{\operatorname{dist}(x, E(\psi)), \operatorname{dist}(y, E(\psi))\}^{\alpha_K}$$

for all $x, y \in K$.

We will deduce weaker conditions from the lower distance estimate that are somewhat easier to read, but let us first prove that the lower distance estimate is indeed a necessary condition.

Proposition 4.11.
If $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ is injective and $\mathcal{A}(\psi)$ is closed, then ψ must satisfy the lower distance estimate.

Proof. We can basically proceed as in the case of a discrete critical set. Suppose that the lower distance estimate fails on a fixed compact set L. We obtain the existence of some sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ in L such that

$$\|\psi(x_n) - \psi(y_n)\| < \frac{1}{n^n} \cdot \|x_n - y_n\| \cdot \max\{\operatorname{dist}(x_n, E(\psi)), \operatorname{dist}(y_n, E(\psi))\}^n.$$

This implies that $x_n \neq y_n$. Without loss of generality we can again suppose that the sequences $(x_n)_{n\in\mathbb{N}}, (y_n)_{n\in\mathbb{N}}$ and $r_n = \frac{y_n - x_n}{\|y_n - x_n\|}$ converge to some x_∞, y_∞ and r_∞ respectively. Since $\operatorname{dist}(z, E(\psi))$ is bounded on the compact set L, the right hand side converges to zero and the injectivity of ψ implies $x_\infty = y_\infty$. Applying the mean value theorem to every coordinate ψ_j of ψ we obtain an intermediate point $\xi_n^j = x_n + \lambda(y_n - x_n)$ with

$$\|\psi_j'(\xi_n^j)[r_n]\| \leq \frac{1}{n^n} \cdot \max\{\operatorname{dist}(x_n, E(\psi)), \operatorname{dist}(y_n, E(\psi))\}^n \to 0.$$

The left hand side converges to $\|\psi'(x_\infty)[r_\infty]\|$ and the right hand side converges to zero. This implies that x_∞ must be a critical point.
In the previous example, we used the fact that the composition operator was open by constructing appropriate functions in the algebra that lead to a contradiction. Let us recall the precise argument we are going to use.
As previously seen in remark 2.8, the fact that C_ψ is an open map implies that for every compact set $K \subseteq \mathbb{R}^d$ and $m \in \mathbb{N}$ we can find some compact set $L \subseteq \mathbb{R}^q$ as well as $c > 0$

and $k \in \mathbb{N}$ such that every $f \in \mathcal{A}(\psi)$ can be factorized by some $F \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ satisfying

$$\|F\|_{K,m} \leq \begin{cases} c \cdot \|f\|_{L,k}, & \text{if } \|f\|_{L,k} \neq 0 \\ \frac{1}{2}, & \text{otherwise} \end{cases}.$$

Before evaluating this condition, let us explain how we can connect this criterion with the sequences $(x_n)_{n\in\mathbb{N}}$ and $(y_n)_{n\in\mathbb{N}}$ to obtain a contradiction. Suppose we have already constructed some function $\varphi_n \in \mathcal{A}(\psi)$ such that $|\varphi_n(x_n) - \varphi_n(y_n)|$ is "large". We can apply the condition mentioned above to some compact set K containing every line segment $\{x_n + \lambda(y_n - x_n) : \lambda \in [0,1]\}$ and $m = 1$. We obtain some $F_n \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ for which we can use the mean value inequality leading to

$$\begin{aligned} |\varphi_n(x_n) - \varphi_n(y_n)| = |F_n(\psi(x_n)) - F_n(\psi(y_n))| &\leq \|F_n\|_{K,1} \cdot \|\psi(x_n) - \psi(y_n)\| \\ &\leq c \cdot \|\varphi_n\|_{L,k} \cdot \|\psi(x_n) - \psi(y_n)\|. \end{aligned} \tag{$*$}$$

This leads to the required contradiction if $|\varphi_n(x_n) - \varphi_n(y_n)|$ is large compared to the right hand side above.
Unfortunately we need to be a little more cautious with our construction of φ_n than in the case of isolated critical points in order to get a contradiction. To simplify the computations, suppose without loss of generality that $\operatorname{dist}(x_n, E(\psi)) \geq \operatorname{dist}(y_n, E(\psi))$. We distinguish the two following cases.
If $\frac{1}{2}\operatorname{dist}(x_n, E(\psi)) \leq \|x_n - y_n\|$, consider some $\varphi \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ with support in the open unit ball and $\varphi(0) = 1$. Define

$$\varphi_n(z) = \varphi\left(\frac{z - x_n}{\frac{1}{2}\operatorname{dist}(x_n, E(\psi))}\right).$$

The support of φ_n is contained in the open unit ball around x_n with radius $\frac{1}{2}\operatorname{dist}(x_n, E(\psi))$ and therefore does not meet y_n nor $E(\psi)$. By proposition 4.2 we have $\varphi_n \in \mathcal{A}(\psi)$ with $\varphi_n(x_n) = 1$ and $\varphi_n(y_n) = 0$. With the estimate $(*)$ above, we get a parametrization F_n satisfying

$$1 = |\varphi_n(x_n) - \varphi_n(y_n)| = |F(\psi(x_n)) - F(\psi(y_n))| \leq c \cdot \|\varphi_n\|_{L,k} \cdot \|\psi(x_n) - \psi(y_n)\|.$$

Dominating the norm $\|\varphi_n\|_{L,k}$ by $2^k \cdot \operatorname{dist}(x_n, E(\psi))^{-k} \cdot \|\varphi\|_{L,k}$ and using our assumption that $\|\psi(x_n) - \psi(y_n)\| \leq \frac{1}{n^n} \cdot \|x_n - y_n\| \cdot \operatorname{dist}(x_n, E(\psi))^n$ we obtain the required contradiction

$$1 \leq c \cdot \|\varphi_n\|_{L,k} \cdot \|\psi(x_n) - \psi(y_n)\| \leq \frac{C}{n^n} \cdot \operatorname{dist}(x_n, E(\psi))^{n-k} \cdot \|x_n - y_n\| \to 0$$

Otherwise, we have $\frac{1}{2}\operatorname{dist}(x_n, E(\psi)) \geq \|x_n - y_n\|$.
For our construction we require some $\Phi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ with support in the open unit ball $B(0,1)$ and such that $\alpha = \inf\{\Phi'(x)[e_1] : \|x\| \leq 1/2\} > 0$. We define

$$\varphi_n(z) = \Phi\left(R_n \frac{z - x_n}{\operatorname{dist}(x_n, E(\psi))}\right),$$

where R_n is a normed rotation matrix mapping $\frac{y_n - x_n}{\|y_n - x_n\|}$ to the first unit vector e_1
Since R_n maps the open unit ball to itself, the support of φ_n is contained in the open ball around x_n with radius $\operatorname{dist}(x_n, E(\psi))$. It does therefore not intersect $E(\psi)$, which implies $\varphi_n \in \mathcal{A}(\psi)$ by proposition 4.2.
Define the direction $r_n = \frac{1}{\operatorname{dist}(x_n, E(\psi))} \cdot R_n(y_n - x_n)$. Applying the mean value theorem to the map $\mu : \mathbb{R} \to \mathbb{R}$ defined by $\mu(t) = \Phi(t \cdot r_n)$, we obtain some $\xi \in [0,1]$ such that

$$\begin{aligned}\varphi_n(y_n) - \varphi_n(x_n) = \mu(1) - \mu(0) = \mu'(\xi) &= \Phi'(\xi \cdot r_n)[r_n] \\ &= \Phi'(\xi \cdot r_n)\left[R_n \frac{y_n - x_n}{\|y_n - x_n\|}\right] \cdot \frac{\|y_n - x_n\|}{\operatorname{dist}(x_n, E(\psi))} \\ &= \Phi'(\xi \cdot r_n)[e_1] \cdot \frac{\|y_n - x_n\|}{\operatorname{dist}(x_n, E(\psi))}.\end{aligned}$$

Since $\|\xi \cdot r_n\| \leq \|r_n\| \leq \frac{1}{2}$, we obtain $\Phi'(\xi \cdot r_n)[e_1] \geq \alpha$ and hence

$$\varphi_n(y_n) - \varphi_n(x_n) \geq \alpha \cdot \frac{\|y_n - x_n\|}{\operatorname{dist}(x_n, E(\psi))}.$$

The linear map R_n has norm $|||R_n||| = 1$ and we can estimate $\|\varphi_n\|_{L,k}$ by the constant $\operatorname{dist}(x_n, E(\psi))^{-k} \cdot \|\Phi\|_{\mathbb{R}^q,k}$. Using $(*)$ again, we obtain

$$\alpha \cdot \frac{\|y_n - x_n\|}{\operatorname{dist}(x_n, E(\psi))} \leq |\varphi_n(y_n) - \varphi_n(x_n)| \leq \frac{\tilde{C}}{n^n} \cdot \operatorname{dist}(x_n, E(\psi))^{n-k} \cdot \|x_n - y_n\|,$$

hence the required contradiction

$$0 < \alpha \leq \frac{C_k}{n^n} \cdot \operatorname{dist}(x_n, E(\psi))^{n+1-k} \to 0$$

when dividing both sides with $\frac{\|y_n - x_n\|}{\operatorname{dist}(x_n, E(\psi))}$.

$\square$

Remark 4.12.
The lower distance estimate of definition 4.10 is quite lengthy but implies the following, formally weaker yet simpler conditions.

(i) For all compact subsets K of $\mathbb{R}^q$ there are constants $c_K > 0$ and $\alpha_K > 0$ such that

$$\|\psi'(x)[r]\| \geq c_K \cdot \operatorname{dist}(x, E(\psi))^{\alpha_K}$$

for all $x \in K$ and $\|r\| = 1$.

(ii) For all compact subsets K of $\mathbb{R}^q$ there are constants $C_K > 0$ and $\alpha_K > 0$ such that

$$|||\psi'(x)^{-1}||| \leq C_K \cdot \left(\frac{1}{\operatorname{dist}(x, E(\psi))}\right)^{\alpha_K}$$

for all $x \in K$.

Proof. To prove (i) we can evaluate the lower distance estimate on a given compact subset K of $\mathbb{R}^q$ in the points $x \in K \cap E(\psi)^c$ and $x + tr$ for $\|r\| = 1$ and t sufficiently small. The lower distance estimate states that

$$\begin{aligned}\|\psi(x+tr) - \psi(x)\| &\geq c_K \cdot \|x + tr - x\| \cdot \max\{\operatorname{dist}(x+tr, E(\psi)), \operatorname{dist}(x, E(\psi))\}^{\alpha_K} \\ &= c_K \cdot |t| \cdot \max\{\operatorname{dist}(x+tr, E(\psi)), \operatorname{dist}(x, E(\psi))\}^{\alpha_K}.\end{aligned}$$

Dividing both sides with $|t|$ yields

$$\left\|\frac{\psi(x+tr) - \psi(x)}{t}\right\| = \frac{\|\psi(x+tr) - \psi(x)\|}{|t|} \geq c_K \cdot \max\{\operatorname{dist}(x+tr, E(\psi)), \operatorname{dist}(x, E(\psi))\}^{\alpha_K}.$$

Since the left hand side converges to $\|\psi'(x)[r]\|$ and the right hand side converges to $c_K \cdot \operatorname{dist}(x, E(\psi))^{\alpha_K}$, we obtain

$$\|\psi'(x)[r] geq c_K \cdot \operatorname{dist}(x, E(\psi))^{\alpha_K}.$$

as claimed.
Part (i) states $\|\psi'(x)[r]\| \geq c_K \cdot \operatorname{dist}(x, E(\psi))^{\alpha}$ for all $x \in K$ and $\|r\| = 1$. This implies $\lambda(\psi'(x)) \geq c_K \cdot \operatorname{dist}(x, E(\psi))^{\alpha_K}$ for all x in the compact set $K \subseteq \mathbb{R}^q$, where we recall that $\lambda(T) = \inf\{\|T[r]\| : \|r\| = 1\}$ according to definition 3.6. Remark 3.7 states that $|||T^{-1}||| = 1/\lambda(T)$. Defining $C_K = 1/c_K$ we obtain

$$|||\psi'(x)^{-1}||| = \frac{1}{\lambda(\psi'(x))} \leq C_K \cdot \frac{1}{\operatorname{dist}(x, E(\psi))^{\alpha_K}}.$$

$\square$

We have already seen in the examples 4.8 and 4.9 that the local Hölder continuity of the inverse is easy to check. However, we only know this condition to be necessary if the critical set is discrete. The lower distance estimate on the other hand is not restricted by this special case. Let us illustrate this fact in the following two examples, where the critical set $E(\psi)$ will be given by a union of lines.

Example 4.13.
Consider again the map $\varphi : \mathbb{R} \to \mathbb{R}$ defined by $\varphi(z) = \exp\left(-\frac{1}{|z|}\right)$ for $z \neq 0$ and $\varphi(0) = 0$ as mentioned in example 4.8. We define $\psi : \mathbb{R}^2 \to \mathbb{R}^3$ by

$$\psi(x,y) = \begin{pmatrix} x^2 \cdot y^2 \\ x \cdot \varphi(x^2) \\ y \cdot \varphi(y^2) \end{pmatrix}.$$

We have depicted the shape of ψ in the following figure 4.3. One can see the sharp edges along the set $\{(x, y, 0) : x = 0 \text{ or } y = 0\}$, which is actually the image of the cross $C = \{(x, y) : x = 0 \text{ or } y = 0\}$.

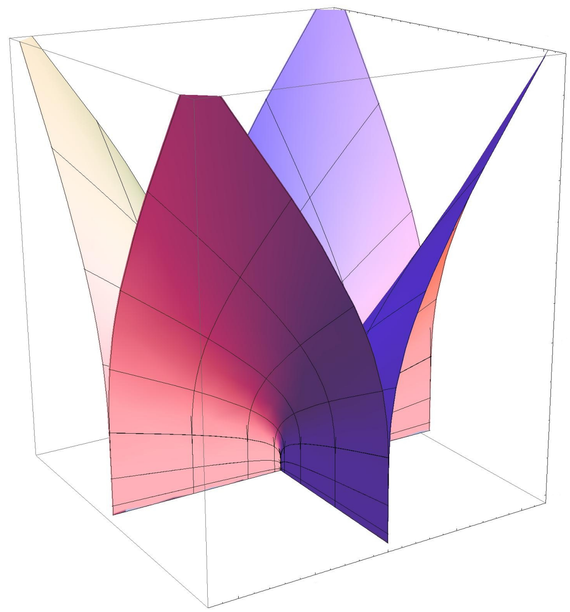


Figure 4.3: The image of $[-1,1]^2$ under ψ.

Let us first show that the injectivity of ψ is a direct consequence of the injectivity of the function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(t) = t \cdot \varphi(t^2)$. Indeed the derivative of $s \mapsto \varphi(s)$ for $s > 0$ is given by $\varphi'(s) = \frac{1}{s^2} \cdot \varphi(s)$. We can therefore compute $g'(t) = \varphi(t^2) + 2t^2 \cdot \varphi'(t^2)$, which is strictly positive for all $t \neq 0$. The function g is therefore injective and for $\psi(x,y) = \psi(z,w)$ we have $g(x) = g(z)$ and $g(y) = g(w)$, hence $(x,y) = (z,w)$. As mentioned above, the results of proposition 4.4 and corollary 4.6 cannot be applied here. Indeed the critical set is not discrete anymore and we will compute it to prove that it is instead given by the cross $C = \{(x,y) \in \mathbb{R}^2 : x = 0 \text{ or } y = 0\}$. Computing the matrix representation of the derivative we obtain

$$\nabla\psi(x,y) = \begin{pmatrix} 2 \cdot x \cdot y^2 & 2 \cdot y \cdot x^2 \\ g'(x) & 0 \\ 0 & g'(y) \end{pmatrix}.$$

Since $g'(t) \neq 0$ whenever $t \neq 0$, we see that $\psi'(x,y)$ has rank 2 whenever $(x,y) \notin C$. For $(x,y) \in C$ we have $x \cdot y = 0$ hence

$$\nabla\psi(x,y) = \begin{pmatrix} 0 & 0 \\ g'(x) & 0 \\ 0 & g'(y) \end{pmatrix},$$

which implies that the rank of $\nabla\psi'(x,y)$ is at most 1 since either $g'(x) = 0$ or $g'(y) = 0$. We can however show that ψ does not generate a closed composition algebra by proving that it fails the lower distance estimate, which is necessary by proposition 4.11. To do this, we need to find a compact set $K \subseteq \mathbb{R}^d$ such that the estimate

$$\|\psi(x,y) - \psi(z,w)\| \geq c_K \cdot \|(x,y) - (z,w)\| \cdot \max\{\operatorname{dist}((x,y),C), \operatorname{dist}(z,w),C)\}^{\alpha_K}$$

does not hold. First let us note that the distance of a point (x, y) to the critical set is given by $\operatorname{dist}((x, y), C) = \min\{|x|, |y|\}$. Now suppose the algebra $\mathcal{A}(\psi)$ to be closed and therefore ψ to satisfy the lower distance estimate by proposition 4.11. For the compact set $K = [0, 1]^2$ we would obtain appropriate constants $c_K > 0$ and $\alpha_K > 0$. Evaluating the lower distance estimate in the points (x, x) and $(-x, x)$ for $x \in [0, 1]$ leads to

$$\|\psi(x, x) - \psi(-x, x)\| \geq c_K \cdot \|(x, x) - (-x, x)\| \cdot \max\left\{\operatorname{dist}\left((x, x), C\right), \operatorname{dist}\left((-x, x), C\right)\right\}^{\alpha_K}.$$

Since $\|\psi(x, x) - \psi(-x, x)\| = |2 \cdot x \cdot \varphi(x^2)|$ and $\operatorname{dist}((x, x), C) = |x| = \operatorname{dist}((-x, x), C)$, we obtain

$$|2 \cdot x \cdot \varphi(x^2)| \geq c_K \cdot |x|^{\alpha_K + 1}.$$

This leads to the contradiction $|\varphi(x^2)| \geq \sqrt{2} \cdot c_K \cdot |x|^{\alpha_K}$ for $x > 0$ sufficiently small.

Chapter 5

Composition isomorphisms between spaces of flat functions.

In this section we will extend the results of chapter 2 to injective maps $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ of several variables. Let us recall the one-dimensional statements. Proposition 2.16 states that, if $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ is injective and proper, every function with support outside $E(\psi)$ is a composition and can even be written as $F \circ \psi$ with F having support outside $\psi(E(\psi))$. Even more, it follows from proposition 2.19 that under the stronger assumption of a locally Hölder continuous inverse the composition algebra is closed. Since every function $f \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ that is flat on $E(\psi)$ is contained in the closure of $\mathcal{A}(\psi)$, this implies that every such f can be written as a composition $f = F \circ \psi$. It also seems plausible that F can be chosen to be flat on the image $\psi(E(\psi))$ of the critical points. This is actually true in the one-dimensional case when ψ^{-1} is locally Hölder continuous, and can be deduced by the same arguments we used in the proof of theorem 2.19. We will show that this is also the case for several variables and give sufficient conditions for $\mathcal{I}(E(\psi)) \subseteq \mathcal{A}(\psi)$. By proposition 4.11, the lower distance estimate is necessary to obtain a closed composition algebra. This condition therefore seems like a plausible assumption. It will turn out that we only require the following, formally weaker notion mentioned in remark 4.12 (ii).

Definition 5.1.
An injective function $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ satisfies the "lower distance estimate for derivatives" if for any compact set $K \subseteq \mathbb{R}^q$ we can find $c_K > 0$ and $\alpha_K \in \mathbb{N}$ such that

$$\|\psi'(x)[r]\| \geq c_K \cdot \operatorname{dist}(x, E(\psi))^{\alpha_K}$$

for all $x \in K$ and $r \in \mathbb{R}^q$ with $\|r\| = 1$.

The lower distance estimate for the derivative can be viewed as a generalization of the necessary condition mentioned in proposition 2.12. In this context, our main theorem reads as follows.

Theorem 5.2.
For an injective $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ that is proper, satisfies the lower distance estimate for derivatives, and has a Whitney regular image the set $\mathcal{I}(E(\psi))$ is contained in $\mathcal{A}(\psi)$. Moreover the representation of $f \in \mathcal{I}(E(\psi))$ as $f = F \circ \psi$ with $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ can even be chosen to be "orthogonally flat with respect to ψ" meaning that F is directionally flat in $\psi(x)$ along directions that are orthogonal to the tangent space $\mathrm{Range}(\psi'(x))$ *at $\psi(x)$ as explained below.*

For the proof, we refer to 5.12

5.1 Preliminary definitions

First we need a firm definition of the subspace of "orthogonally flat" functions as well as a few topological notions. The proof of the main theorem will heavily rely on the computational tools given in section 3, more precisely propositions 3.2 and 3.7. We remind the reader that we identify the k-th derivative of a smooth map $F \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ with the function $F^{(k)}$ with values in the space $\mathcal{M}_k(\mathbb{R}^q, \mathbb{R}^d)$ of k-linear maps from $\mathbb{R}^q$ to $\mathbb{R}^d$. This differs from the viewpoint of partial derivatives, which are essential to compute the usual seminorms $\|F\|_{K,n} = \sup\{|D^\alpha F(x)| : x \in K, |\alpha| \leq n\}$ generating the Fréchet topology on $\mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$. The directional derivatives turn out to be more adapted to our computations. Let us start by defining the respective families of seminorms.

Definition 5.3.
We recall lemma 3.3 (ii), where we defined the norm $|||T|||$ of a k-linear map $T \in \mathcal{M}_k(X, Y)$ between Banach spaces by $|||T||| = \sup\{\|T[x_1, ..., x_k]\|_Y : \|x_1\|_X \leq 1, ..., \|x_K\|_X \leq 1\}$.
In analogy to the usual seminorms $\|F\|_{K,n}$ on $\mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ where $K \subseteq \mathbb{R}^q$ is compact and $n \in \mathbb{N}_0$ we define

$$\begin{aligned} |||F|||_{K,n} &= \sup\{|||F^{(k)}(x)||| : x \in K, k \leq n\} \\ &= \sup\left\{ \left\| F^{(k)}(x)[r_1, ..., r_k] \right\| : x \in K, k \leq n, \|r_1\| \leq 1, ..., \|r_k\| \leq 1 \right\}. \end{aligned}$$

As one would expect this is just a superficial change of interpretation as $\|\cdot\|_{K,n}$ and $|||\cdot|||_{K,n}$ are equivalent seminorms. Both families therefore generate the same Fréchet topology on $\mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$. For reasons of completeness we have given a separate proof of this fact in proposition 6.1 of the appendix.
We will now proceed and formally define the required spaces of flat functions mentioned in proposition 5.2.

Definition 5.4.
Consider a subset $M \subseteq \mathbb{R}^k$.

(i) We denote by

$$\mathcal{I}(M, \mathbb{R}^n) = \{F \in \mathcal{E}(\mathbb{R}^k, \mathbb{R}^n) : F^{(j)}(x) = 0 \ \forall x \in M, j \in \mathbb{N}_0\}$$

the set of flat functions on M with values in $\mathbb{R}^n$. We recall that we view the j-th derivative $F^{(j)}$ as a map from $\mathbb{R}^k$ to the space $\mathcal{M}_j(\mathbb{R}^k, \mathbb{R}^n)$ of j-linear functions from $\mathbb{R}^k$ to $\mathbb{R}^n$. We mostly deal with the case $n = 1$ and therefore simply write

$$\mathcal{I}(M) = \mathcal{I}(M, \mathbb{R}).$$

Obviously, $\mathcal{I}(M, \mathbb{R}^n)$ is a closed subspace of $\mathcal{E}(\mathbb{R}^k, \mathbb{R}^n)$.

(ii) We define the space of smooth functions on $M \subseteq \mathbb{R}^k$ as the quotient

$$\mathcal{E}(M, \mathbb{R}^n) = \mathcal{E}(\mathbb{R}^k, \mathbb{R}^n)/\mathcal{I}(M, \mathbb{R}^n)$$

which we endow with the usual quotient topology. Since $\mathcal{I}(M, \mathbb{R}^n)$ is closed, the quotient $\mathcal{E}(M, \mathbb{R}^n)$ is a Fréchet space. Again, $\mathcal{E}(M)$ will denote the special case $\mathcal{E}(M, \mathbb{R})$ of a one-dimensional range.

(iii) For $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ and its critical set $E(\psi)$, we will abbreviate

a) $\mathcal{E}(\psi) = \mathcal{E}(\psi(\mathbb{R}^q))$,

b) $\mathcal{I}(\psi) = \mathcal{I}(\psi(\mathbb{R}^q))$,

c) $\mathcal{I}(E) = \mathcal{I}(E(\psi))$ and

d) $\mathcal{I}(\psi(E)) = \mathcal{I}(\psi(E(\psi)))$.

We can now adapt these notions of quotient spaces to the generator $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$, more precisely to its image $\psi(\mathbb{R}^q)$, to obtain the definition area of the (reduced) composition operator. Even though it might not directly seem clear from the definition why we consider the upcoming set, we will elaborate an interpretation to explain why it is more natural than it seems.

Definition 5.5.
Consider a $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$.

(i) We call a function $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ orthogonally flat (with respect to ψ) if

$$F^{(k)}(\psi(x))\,[\mathfrak{r}] = 0$$

for all multidirections $\mathfrak{r} = (r_1, ..., r_k) \in \prod_{j=1}^{k} \mathrm{Range}(\psi'(x))^{\perp}$ and all $x \in \mathbb{R}^q$ and $k \in \mathbb{N}$.

(ii) If $E(\psi) \subseteq \mathbb{R}^q$ denotes the critical set of ψ, we consider the "reduced definition area"

$$\mathcal{R}(\psi) = \{F \in \mathcal{I}(\psi(E)) : F \text{ orhtogonally flat w.r.t. } \psi\}$$

and the respective quotient

$$\mathcal{E}_0(\psi) = \mathcal{R}(\psi)/\mathcal{I}(\psi).$$

Remarks 5.6.

(i) The subspace $\mathcal{R}(\psi)$ is closed since it is the intersection over all $k \in \mathbb{N}_0, x \in \mathbb{R}^q$, and $\mathfrak{r} \in \prod_{j=1}^{k} \operatorname{Range}(\psi'(x))^{\perp}$ of the sets

$$\mathcal{R}(\psi, k, x, \mathfrak{r}) = \left\{F \in \mathcal{I}(\psi(E)) : F^{(k)}(\psi(x))[\mathfrak{r}] = 0\right\}.$$

These subspaces are closed as the preimages of $\{0\}$ under the continuous evaluation maps $F \mapsto F^{(k)}(\psi(x))[\mathfrak{r}]$. Therefore $\mathcal{R}(\psi)$ is a Fréchet space and, since $\mathcal{I}(\psi) \subseteq \mathcal{R}(\psi)$, the quotient space is also Fréchet when endowed with the quotient topology.

(ii) $\mathcal{R}(\psi)$ is even an ideal. Indeed, consider $F \in \mathcal{R}(\psi)$ and $\Phi \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$. Obviously, all derivatives of the product $\Phi \cdot F$ vanish on $\psi(E)$ because those of F do. For $x \notin E$ and some normal multidirection $\mathfrak{r} = (r_1, ..., r_n) \in \prod_{j=1}^{n} \operatorname{Range}(\psi'(x))^{\perp}$ the Leibniz-formula implies

$$(\Phi \cdot F)^{(n)}(\psi(x))[\mathfrak{r}] = \sum_{A \subseteq \{1,...,n\}} \Phi^{(|A|)}(\psi(x))[\mathfrak{r}_A] \cdot F^{(|A^c|)}(\psi(x))[\mathfrak{r}_{A^c}] = 0,$$

where we recall that $\mathfrak{r}_A = \underset{j \in A}{\boxtimes} r_j$ stands for the A-selection of $\mathfrak{r} = (r_1, ..., r_n)$.

The reason to consider the subspaces $\mathcal{R}(\psi)$ and $\mathcal{E}_0(\psi)$ is that the reduced composition operator $\tilde{C}_\psi : \mathcal{E}_0(\psi) \to \mathcal{A}(\psi)$ turns out to be injective. Prior to giving a formal proof of this fact, let us explain why this space is more intuitive than the lengthy and technical definition would suggest.
Fix some $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$. For $x \in \mathbb{R}^q$, one obtains the value $F'(\psi(x))[r]$ for all directions $r = \psi'(x)[s] \in \operatorname{Range}(\psi'(x))$ directly from the behavior of $F \circ \psi$. Indeed, the values of F on $\psi(\mathbb{R}^q)$ are uniquely determined by $F \circ \psi$ and we obtain the derivative through the computation

$$F'(\psi(x))[r] = F'(\psi(x))[\psi'(x)[s]] = (F \circ \psi)'(x)[s].$$

On the other hand, other directional derivatives cannot be constructed by using only knowledge of $F \circ \psi$. This gives us too many degrees of freedom and there is little hope whatsoever to obtain an injective composition operator. We can of course compensate this problem by demanding that the function F is "as good as constant", i.e. flat in directions that are orthogonal to the tangent space, thus resulting in $F'(\psi(x))[r] = 0$ for all directions r orthogonal to $\psi'(x)[\mathbb{R}^q] = \operatorname{Range}(\psi'(x))$. We then obtain complete knowledge of $F'(\psi(x))$

and iterating the process we can gain full knowledge of $F^{(k+1)}(\psi(x))$ only by the behavior of the derivatives of F on $\psi(\mathbb{R}^q)$ together with the orthogonality condition. Let us now proceed and formalize this idea.

Proposition 5.7.
The composition operator $\tilde{C}_\psi : \mathcal{E}_0(\psi) \to \mathcal{A}(\psi),\ F+\mathcal{I}(\psi) \mapsto F\circ\psi$ is well-defined, continuous and injective.

Proof. Obviously for two representatives $F, G \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ of the same equivalence class in $\mathcal{E}_0(\psi)$ the difference $F-G$ is flat and hence zero on $\psi(\mathbb{R}^q)$. We therefore have $F\circ\psi = G\circ\psi$ and the composition with ψ is well-defined.
Let us show that the injectivity of $\tilde{C}_\psi$ results from the inclusion $\mathrm{Ker}(C_\psi)\cap\mathcal{R}(\psi) \subseteq \mathcal{I}(\psi)$. To this end, consider $\mathcal{G} \in \mathcal{E}_0(\psi)$ with $\tilde{C}_\psi(\mathcal{G}) = 0$. We can write as $\mathcal{G} = G + \mathcal{I}(\psi)$ with some $G \in \mathcal{R}(\psi)$. By the definition of $\tilde{C}_\psi$, we obtain $C_\psi(G) = \tilde{C}_\psi(\mathcal{G}) = 0$ and therefore $G \in \mathrm{Ker}(C_\psi)\cap\mathcal{R}(\psi)$. The inclusion $\mathrm{Ker}(C_\psi)\cap\mathcal{R}(\psi) \subseteq \mathcal{I}(\psi)$ then implies $G \in \mathcal{I}(\psi)$ hence $\mathcal{G} = \mathcal{I}(\psi)$ and the injectivity.
It remains to verify that $\mathrm{Ker}(C_\psi)\cap\mathcal{R}(\psi) \subseteq \mathcal{I}(\psi)$ holds. To do so, fix $F \in \mathrm{Ker}(C_\psi)\cap\mathcal{R}(\psi)$. We need to prove $F^{(k)}(\psi(x)) = 0$ for all $k \in \mathbb{N}_0$ and $x \in \mathbb{R}^q$.
For $x \in E(\psi)$ this follows from the definition since $F \in \mathcal{R}(\psi)$ is flat on $\psi(x) \in \psi(E(\psi))$.
For $x \in E(\psi)^c$ we will proceed by induction. The case $k = 0$ is obvious since we have $F \in \mathrm{Ker}(C_\psi)$ and therefore $F(\psi(x)) = 0$ for all $x \in \mathbb{R}^q$. For $k = 1$ we can write $\mathbb{R}^d$ as the direct sum $\mathrm{Range}(\psi'(x)) \oplus \mathrm{Range}(\psi'(x))^\perp$, thus obtaining for every direction $r \in \mathbb{R}^d$ a unique representation $r = \psi'(x)[s] + \eta$ with $\eta \perp \psi'(x)[\mathbb{R}^q]$. The linearity of $F'(\psi(x))$ combined with $F'(\psi(x))[\eta] = 0$ implies

$$F'(\psi(x))[r] = F'(\psi(x))[\psi'(x)[s]] + F'(\psi(x))[\eta] = (F\circ\psi)'(x)[s] = 0.$$

The key to the computation above is the fact that we can represent every direction r as a sum $\psi'(x)[s] + \eta$. For the induction step we will need a similar decomposition for multidirections $\mathfrak{r} = (r_1, ..., r_k)$. This is a more general result about symmetric multilinear maps and thus we have given a separate proof in lemma 5.8 in order to keep the computation to a minimum.
Suppose that $F^{(k-1)} = 0$ on $\psi(\mathbb{R}^q)$ and fix again some regular point $x \in \mathbb{R}^d \setminus E(\psi)$ and a multidirection $\mathfrak{r} = (r_1, ..., r_k)$. We can decompose each component r_j as the sum $\psi'(x)[s_j] + \eta_j$ with some $\eta_j \in \mathrm{Range}(\psi'(x))^\perp$. By lemma 5.8 we have

$$F^{(k)}(\psi(x))\left[\mathop{\boxtimes}_{j=1}^{k} r_j\right] = \sum_{A\subseteq\{1,...,k\}} F^{(k)}(\psi(x))\left[\mathop{\boxtimes}_{j\in A}\psi'(x)[s_j]\mathop{\boxtimes}_{\ell\in A^c}\eta_\ell\right].$$

Thus we only need to consider the “pure” cases, where the components r_j are either elements of $\mathrm{Range}(\psi'(x))$ or its orthogonal complement. Now consider the following two cases:
First case: If every r_j is situated in the orthogonal complement of $\mathrm{Range}(\psi'(x))$, the definition of $\mathcal{R}(\psi)$ implies $F^{(k)}(\psi(x))[r_1, ..., r_k] = 0$.

Second case: There exists a coordinate ℓ with $r_\ell = \psi'(x)[s_\ell]$. Using the fact that $F^{(k)}(\psi(x))$ is symmetric we can suppose $\ell = k$ simply by switching those coordinates. Defining $\Phi(z) = F^{(k-1)}(z)[r_1, ..., r_{k-1}]$ we obtain a smooth map $\Phi \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ that vanishes on $\psi(\mathbb{R}^q)$ by the induction assumption. Computing its derivative leads to the required

$$F^{(k)}(\psi(x))[r_1, ..., r_k] = \Phi'(\psi(x))[\psi'(x)[s_k]] = (\Phi \circ \psi)'(x)[s_k] = 0.$$

To prove the continuity we will show that preimages of open sets under $\tilde{C}_\psi$ are open in $\mathcal{E}_0(\psi)$. For an open subset U of $\mathcal{E}(\mathbb{R}^q, \mathbb{R})$ the preimage $C_\psi^{-1}(U)$ is open in $\mathcal{E}(\mathbb{R}^d, \mathbb{R})$. Since $\mathcal{R}(\psi)$ is a closed subspace of $\mathcal{E}(\mathbb{R}^d, \mathbb{R})$, the intersection $C_\psi^{-1}(U) \cap \mathcal{R}(\psi)$ is open in $\mathcal{R}(\psi)$ with respect to the relative topology. The injectivity of $\tilde{C}_\psi$ implies the identity $\tilde{C}_\psi^{-1}(U) = (C_\psi^{-1}(U) \cap \mathcal{R}(\psi)) + \mathcal{I}(\psi)$, which is an open set in $\mathcal{E}_0(\psi)$ with respect to the quotient topology.

$\square$

Lemma 5.8.
Consider two vector spaces X and Y and a k-linear symmetric map $F : X^k \to Y$. For all multidirections $(s_1, ..., s_k), (q_1, ..., q_k) \in X^k$ we have the identity

$$F\left[\boxtimes_{j=1}^{k} (s_j + q_j)\right] = \sum_{A \subseteq \{1,...,k\}} F\left[\boxtimes_{j\in A} s_j \boxtimes_{\ell \in A^c} q_\ell\right].$$

Proof. We recall that by definition 3.1 (ii) we have defined $\boxtimes_{j\in\emptyset} x_j \boxtimes \mathfrak{r} = \mathfrak{r} = \mathfrak{r} \boxtimes_{j\in\emptyset} x_j$. Furthermore the order of the different arguments do not matter for a symmetric k-linear map F.

Now let us proceed with the proof by induction. For $k = 1$ we have only two subsets of $\{1\}$ namely $\emptyset$ and $\{1\}$. Using the notation recalled above, the linearity of F implies

$$F[s_1 + q_1] = F[s_1] + F[q_1] = F\left[\boxtimes_{j\in\{1\}} s_j \boxtimes_{\ell\in\emptyset} q_\ell\right] + F\left[\boxtimes_{j\in\emptyset} s_j \boxtimes_{\ell\in\{1\}} q_\ell\right].$$

To prove the induction step, consider some $k+1$-linear symmetric map F. The evaluations $(r_1, .., r_k) \mapsto F[r_1, .., r_k, x]$ are k-linear and hence for $x = s_{k+1}$ and $x = q_{k+1}$ we can apply the induction assumption to obtain

$$\begin{aligned}
F\left[\mathop{\boxtimes}_{j=1}^{k+1}(s_j+q_j)\right] &= F\left[\mathop{\boxtimes}_{j=1}^{k}(s_j+q_j)\boxtimes s_{k+1}\right] + F\left[\mathop{\boxtimes}_{j=1}^{k}(s_j+q_j)\boxtimes q_{k+1}\right]\\
&= \sum_{A\subseteq\{1,\dots,k\}} F\left[\mathop{\boxtimes}_{j\in A} s_j \mathop{\boxtimes}_{\ell\in A^c} q_\ell \boxtimes s_{k+1}\right] + \sum_{A\subseteq\{1,\dots,k\}} F\left[\mathop{\boxtimes}_{j\in A} s_j \mathop{\boxtimes}_{\ell\in A^c} q_\ell \boxtimes q_{k+1}\right]\\
&= \sum_{A\subseteq\{1,\dots,k\}} F\left[\mathop{\boxtimes}_{j\in A\cup\{k+1\}} s_j \mathop{\boxtimes}_{\ell\in A^c} q_\ell\right] + \sum_{A\subseteq\{1,\dots,k\}} F\left[\mathop{\boxtimes}_{j\in A} s_j \mathop{\boxtimes}_{\ell\in A^c\cup\{k+1\}} q_\ell\right]\\
&= \sum_{\substack{A\subseteq\{1,\dots,k+1\}\\ k+1\in A}} F\left[\mathop{\boxtimes}_{j\in A} s_j \mathop{\boxtimes}_{\ell\in A^c} q_\ell\right] + \sum_{\substack{A\subseteq\{1,\dots,k+1\}\\ k+1\in A^c}} F\left[\mathop{\boxtimes}_{j\in A} s_j \mathop{\boxtimes}_{\ell\in A^c} q_\ell\right]\\
&= \sum_{A\subseteq\{1,\dots,k+1\}} F\left[\mathop{\boxtimes}_{j\in A} s_j \mathop{\boxtimes}_{\ell\in A^c} q_\ell\right].
\end{aligned}$$

□

5.2 C_ψ as an isomorphism between spaces of flat functions.

We will now prove that, under the assumptions of theorem 5.2, every $f \in \mathcal{I}(E(\psi))$ possesses a representation as $f = F \circ \psi$ with $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ satisfying the orthogonal flatness condition. We can view this statement as an improvement of proposition 4.2, since we are now able to gain control of the derivatives on $\psi(\mathbb{R}^q)$ of a certain representation F of $F \circ \psi$. For the proof of our representation we require the existence of specific diffeomorphisms between submanifolds of $\mathbb{R}^d$, which satisfy certain geometric conditions as well as norm estimates. To make this chapter more readable we have constructed these diffeomorphisms in proposition 6.5 of the appendix.
For reasons of clarity let us mention the simplified statement of proposition 6.5.

Proposition 5.9 (Diffeomorphic extension)**.**
Consider a smooth and injective map $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$. For any $x \notin E(\psi)$ we can find neighborhoods $U_x \subseteq \mathbb{R}^q$ of x, $U_0 \subseteq \mathbb{R}^{d-q}$ of 0, and $V_x \subseteq \mathbb{R}^d$ of $\psi(x)$ as well as a diffeomorphism $\Psi_x : U_x \times U_0 \to V_x$ such that:

(i) $\psi(y) = \Psi(y, 0)$ *for all* $y \in U_x$.

(ii) $\Psi'(y, 0)$ *is a bijection between* $\mathbb{R}^q \times \{0_{d-q}\}$ *and* $\psi'(y)[\mathbb{R}^d]$.

(iii) $\Psi'(y, t)$ *is a bijection between* $\{0_q\} \times \mathbb{R}^{d-q}$ *and* $\psi'(y)[\mathbb{R}^d]^\perp$ *for all* $t \in U_0$.

(iv) $\Psi''(y, t)[s_1, s_2] = 0$ *whenever* $s_1, s_2 \in \{0\} \times \mathbb{R}^{d-q}$.

(v) There are constants C_n and $\alpha(n)$ depending only on n such that

$$|||\Psi|||_{(x,0),n} \leq C_n(1 + |||\psi|||_{x,n})^{\alpha(n)} \cdot (1 + |||\psi'(x)^{-1}|||)^{\alpha(n)}.$$

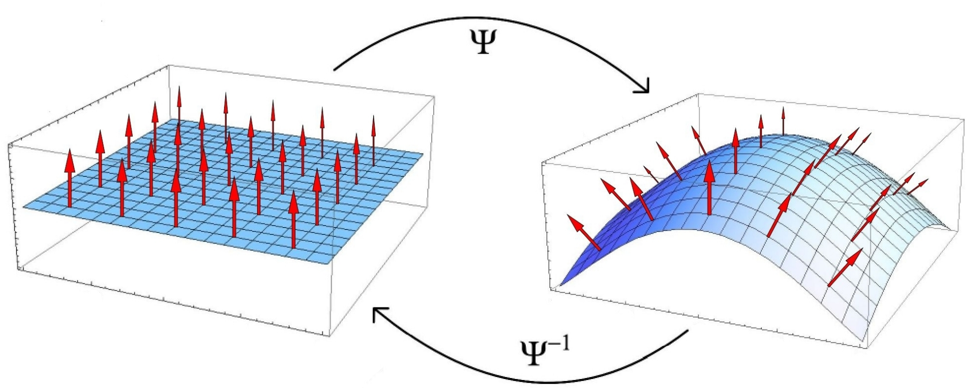
(vi) There are constants c_n and $\beta(n)$ depending only on n such that

$$|||\Psi^{-1}|||_{\psi(x),n} \leq c_n(1 + |||\psi|||_{x,n})^{\beta(n)} \cdot (1 + |||\psi'(x)^{-1}|||)^{\beta(n)}.$$

The first step is a refinement of proposition 4.2 in the sense that every $f \in \mathcal{D}(E(\psi)^c)$ does not only have a smooth parametrization $f = F \circ \psi$, but F can even be chosen to be orthogonally flat with respect to ψ. As a byproduct we also obtain that the set $\mathcal{R}(\psi)$ of orthogonally flat functions is not empty.

Proposition 5.10.
For an injective, smooth, and proper map $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ with critical set $E = E(\psi)$, every function $f \in \mathcal{D}(E^c)$ can be represented as a composition $f = F \circ \psi$ with $F \in \mathcal{R}(\psi)$.

Proof. Given a function $f \in \mathcal{D}(E(\psi)^c)$ we will explicitly construct local parametrizations F_x that are orthogonally flat and satisfy $F_x \circ \psi = f$ near x. We will then glue these local modifications together by means of an adequate partition of unity and verify that the resulting function is still orthogonally flat.
To visualize the construction we have given a simple example of a 2-dimensional submanifold of $\mathbb{R}^3$ in the following figure. The idea will be to derive the values on the left plane from the values on the image right. We then extend the functions on both sides to be constant along the red arrows.



The basic idea behind the construction is rather geometrical in nature. Indeed, around the image $\psi(x)$ of a regular point x, we can represent $\psi(\mathbb{R}^q)$ as a q-dimensional manifold in $\mathbb{R}^d$. This means that the image of ψ is locally diffeomorphic to $\mathbb{R}^q \times \{0_{d-q}\}$. For some $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ representing f, we can pull back $F|_{\psi(\mathbb{R}^q)}$ via the diffeomorphism Ψ_x^{-1}. We obtain some $\mathcal{F}$ defined on $\mathbb{R}^q \times \{0_{d-q}\}$, which we can smoothly extend by $\mathcal{F}(x, t) = \mathcal{F}(x, 0)$.

Applying the composition with the diffeomorphism Ψ_x on $\mathcal{F}$ we obtain some modification $\tilde{F}$ that coincides with F on $\psi(\mathbb{R}^q)$ near $\psi(x)$. The orthogonal flatness condition then follows from the fact that the diffeomorphism maps the last $d-q$ partial derivatives of $\mathcal{F}$, which are zero by construction, to the directional derivatives of $\tilde{F}$ along normal directions at $\psi(x)$, i.e. orthogonal directions to the tangent space $\psi'(x)[\mathbb{R}^q]$.
Let us now formalize the process stated above.
The image $\psi(E^c)$ is a smooth manifold and $\psi : E^c \to \psi(E^c)$ is a global coordinate system (cf. definition 6.3 of the appendix). We can therefore apply the previous proposition 5.9 to some $x \in E^c$ to obtain neighborhoods U_x of $x \in \mathbb{R}^q$, U_0 of $0 \in \mathbb{R}^{d-q}$ and V_x of $\psi(x) \in \mathbb{R}^d$ as well as a diffeomorphism $\Psi_x : U_x \times U_0 \to V_x$ satisfying the conditions (i)-(vi). For $f \in \mathcal{D}(U_x)$ we define $F_x : V_x \to \mathbb{R}$ by

$$F_x = f \circ \pi \circ \Psi_x^{-1},$$

where $\pi : \mathbb{R}^q \times \mathbb{R}^{d-q} \to \mathbb{R}^q$ denotes the projection $(y,t) \mapsto y$. The function Ψ_x is a diffeomorphism, hence Ψ_x^{-1} is smooth on V_x, as well as F_x. By definition we obtain the simpler identity

$$F_x(\Psi_x(y,t)) = f(y)$$

whenever $(y,t) \in U_x \times U_0 = \Psi_x^{-1}(V_x)$. This also implies $F_x \circ \psi = f$ on U_x.
Let us now prove that F_x is orthogonally flat with respect to ψ. Instead of using induction, we can prove this by direct computation using proposition 3.2. Let us nevertheless mention the case $k=1$ explicitly to better flesh out the argument.
By property (iii) of Ψ_x we can write $r \perp \psi'(y)[\mathbb{R}^q]$ as the image $r = \Psi_x'(y,0)[s]$ of some vector $s = (0,t) \in \{0_q\} \times \mathbb{R}^{d-q}$. By construction we have

$$F_x \circ \Psi_x((y,0) + \varepsilon \cdot s) = F_x \circ \Psi_x(y, \varepsilon \cdot t) = f(y) = F_x \circ \Psi_x(y,0)$$

for $|\varepsilon|$ so small that $\varepsilon \cdot t \in U_0$. Therefore the directional derivative $(F_x \circ \Psi_x)'(y,0)[s]$ along s must vanish and we obtain

$$F_x'(\Psi_x(y,0))[r] = F_x'(\Psi_x(y,0))[\Psi_x'(y,0)[s]] = (F_x \circ \Psi_x)'(y,0)[s] = 0.$$

For the general case we need to prove that

$$F^{(k)}(\psi(y))[r_1, .., r_k] = 0$$

whenever all r_j are normal directions at $\psi(y)$, i.e. $r_j \in \psi'(y)[\mathbb{R}^q]^\perp$. By (iii) from proposition 5.9 we can write each r_j as $\Psi_x'(y,0)[s_j]$ with $s_j \in \{0\} \times \mathbb{R}^{d-q}$. Since we have the identity $F \circ \Psi_x(y,t) = F \circ \Psi_x(y,0)$, proposition 3.2 applied to the multidirection $\mathfrak{s} = (s_1, ..., s_k)$ implies

$$0 = (F \circ \Psi_x)^{(k)}(y,0) \left[\boxtimes_{j=1}^{k} s_j \right] = \sum_{\mathcal{P} \in \mathfrak{P}(k)} F^{(|\mathcal{P}|)}(\Psi_x(y,0)) \left[\boxtimes_{P \in \mathcal{P}} \Psi_x^{(|P|)}(y,0)[\mathfrak{s}_P] \right].$$

By proposition 5.9, (iv) we have $\Psi_x''(y,t)[v,w] = 0$ whenever $v, w \in 0_d \times \mathbb{R}^{d-q}$. This implies that $\Psi_x^{(|P|)}(y,0)[\mathfrak{s}_P] = 0$ for all $|P| \geq 2$. Therefore, every evaluation of $F^{(|\mathcal{P}|)}(\Psi_x(y,0))$ with the exception of $\mathcal{P} = \{\{1\}, ..., \{k\}\}$ in the sum on the right hand side must vanish. This leads to the required

$$0 = (F \circ \Psi_x)^{(k)}(y,0) \left[\boxtimes_{j=1}^{k} s_j \right] = F^{(k)}(\Psi_x(y,0)) \left[\boxtimes_{j=1}^{k} \Psi_x'(y,0)[s_j] \right].$$

Note that we only have the identity $F_x \circ \Psi_x(y,0) = F_x \circ \psi(y) = f(y)$ for $y \in U_x$. To obtain the more general parametrization $F_x \circ \psi = f$ we require the properness of ψ. The continuity of ψ^{-1} given by remark 2.11 then implies that we can find an open subset $W_x \subseteq \mathbb{R}^d$ such that $\psi(U_x) = W_x \cap \psi(\mathbb{R}^q)$. By multiplying F_x with some function $\eta_x \in \mathcal{D}(W_x)$ satisfying $\eta_x = 1$ near the compact set $\psi(\mathrm{supp}(f))$ we can suppose that $F_x \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ has support in W_x. This implies $F_x \circ \psi = f$ on U_x as well as $F_x \circ \psi(z) = 0 = f(z)$ for all $z \notin U_x$. Note that the product $F_x \cdot \eta_x$, which is smooth on $\mathbb{R}^d$, is also orthogonally flat with respect to ψ and can be deduced using the Leibniz rule.
For $f \in \mathcal{D}(E^c)$ we can find a finite cover $\{U_x : x \in M\}$ of $\mathrm{supp}(f)$ by open sets U_x as above as well as a partition of unity $\{\varphi_x : x \in M\}$ subordinated to that cover. Again, we can find $F_x \in \mathcal{D}(V_x)$ that is orthogonally flat with respect to ψ and satisfies $F_x \circ \psi = f \cdot \varphi_x$. Gluing the local solutions together we obtain $F = \sum_{x \in M} F_x \in \mathcal{R}(\psi)$ that satisfies

$$F \circ \psi = \sum_{x \in M} F_x \circ \psi = \sum_{x \in M} f \cdot \varphi_x = f.$$

□

This construction is also true if we only ask for $\mathrm{supp}(f) \cap E = \emptyset$, simply by taking a locally finite partition of unity. Moreover if ψ is a proper map, it follows as a direct consequence that every $G \in \mathcal{D}(\mathbb{R}^d)$ with $\mathrm{supp}(G) \cap \psi(E) = \emptyset$ has a decomposition $G = N + K$ where $K \in \mathrm{Ker}(C_\psi)$ and $N \in \mathcal{R}(\psi)$. Indeed the function $g = G \circ \psi$ has support in the set $\psi^{-1}(\mathrm{supp}(G))$, which is compact by the assumption of properness, and contained in $E(\psi)^c$. As seen above one can find some $F \in \mathcal{R}(\psi)$ with $F \circ \psi = g = G \circ \psi$ and hence G can be written as the sum $G = F + (G - F)$ of some $F \in \mathcal{R}(\psi)$ and $G - F \in \mathrm{Ker}(C_\psi)$.
Since $\mathcal{D}(E^c)$ is dense in $\mathcal{I}(E)$, proposition 5.10 states that $\tilde{C}_\psi : \mathcal{E}_0(\psi) \to \mathcal{I}(E)$ is a continuous linear injection with dense range and we will proceed to show that it is actually surjective, hence an isomorphism between both Fréchet spaces. If we take a close look, we have actually proved a rather interesting fact, namely that $\mathcal{E}_0(\psi)$ is isomorphic to a dense subspace of the quotient $\mathcal{I}(\psi(E))/\mathrm{Ker}(C_\psi)$. This is another hint at our claim that the definition we gave is natural. In fact, if $\tilde{C}_\psi : \mathcal{E}_0(\psi) \to \mathcal{I}(E)$ is surjective, then $\mathcal{E}_0(\psi)$ and $\mathcal{I}(\psi(E))/\mathrm{Ker}(C_\psi)$ are isomorphic. It follows from proposition 5.7 that $\tilde{C} : \mathcal{E}_0(\psi) \to \mathcal{I}(E)$ is bijective, hence an isomorphism. Moreover this implies that the restricted composition operator $C_\psi : \mathcal{R}(\psi) \to \mathcal{I}(E)$, and therefore $C_\psi : \mathcal{I}(\psi(E)) \to \mathcal{I}(E)$, are both surjective. The closed range theorem states that this is exactly the case when

$\hat{C}_\psi : \mathcal{I}(\psi(E))/\operatorname{Ker}(C_\psi) \to \mathcal{I}(E)$ is an isomorphism, in which case both spaces $\mathcal{E}_0(\psi)$ and $\mathcal{I}(\psi(E))/\operatorname{Ker}(C_\psi)$ are obviously isomorphic.
It is unclear to us if the two spaces are also isomorphic in the case of a composition operator that does not have closed range.

Proposition 5.11.
If ψ is a proper and injective map, then $\mathcal{E}_0(\psi)$ is isomorphic to a dense subspace of $\mathcal{I}(\psi(E))/\operatorname{Ker}(C_\psi)$.

Proof. Let L and M be two subspaces of a topological vector space X satisfying $L \subseteq M$. We consider the map $T : X/L \to X/M$ defined by

$$T(x+L) = x+M$$

T **is well defined.** For $x, y \in X$ satisfying $x+L = y+L$ the identity $L \subseteq M$ implies $L+M = M$ and hence we obtain $x+M = x+L+M = y+L+M = y+M$.
T **is continuous.** We will show that preimages of open sets under T remain open. This will follow from the identity $T \circ q_L = q_M$, where q_M and q_L denote the respective quotient maps. For an open set U in X/M we obtain $q_M^{-1}(U) = q_L^{-1}(T^{-1}(U))$ and the surjectivity of $q_L : X \to X/L$ implies $q_L(q_M^{-1}(U)) = T^{-1}(U)$, which must be open since q_M is continuous and q_L is open.
Applying this result to $X = \mathcal{I}(\psi(E))$, $L = \mathcal{I}(\psi)$ and $M = \operatorname{Ker}(C_\psi)$ we see that the restriction of T to the subspace $\mathcal{E}_0(\psi)$ of $\mathcal{I}(\psi(E))/\mathcal{I}(\psi)$ is a continuous linear map.
$T|_{\mathcal{E}_0(\psi)}$ **is injective.** Fix $\mathcal{F} \in \mathcal{E}_0(\psi)$ with $T(\mathcal{F}) = \operatorname{Ker}(C_\psi)$. For a representative F of $\mathcal{F}$ we obtain $F \in \mathcal{R}(\psi)$ as well as $F \circ \psi = 0$ and the injectivity of the reduced composition operator $\tilde{C}_\psi$ proved in 5.7 implies $F \in \mathcal{I}(\psi)$, hence $\mathcal{F} = \mathcal{I}(\psi)$.
T **has dense range.** The set $\mathcal{D}(\psi(E)^c)$ is dense in $\mathcal{I}(\psi(E))$. Therefore, its image $\{G + \operatorname{Ker}(C_\psi) : G \in \mathcal{D}(\psi(E)^c)\}$ under the continuous quotient map is also dense in $\mathcal{I}(\psi(E))/\operatorname{Ker}(C_\psi)$. It remains to prove that every such equivalence class $G + \operatorname{Ker}(C_\psi)$ lies in the image of T. For $G \in \mathcal{D}(\psi(E)^c)$ we can find an open set U in $\mathbb{R}^d$ such that $\psi(E) \subseteq U$ and $G|_U = 0$. The composition $G \circ \psi$ is therefore constant zero on the open set $\psi^{-1}(U)$ containing E, hence $G \circ \psi \in \mathcal{D}(E^c)$. By proposition 5.10 we can find $F \in \mathcal{R}(\psi)$ such that $F \circ \psi = G \circ \psi$ and the difference $H = G - F$ is contained in $\operatorname{Ker}(C_\psi)$. This implies $F + \operatorname{Ker}(C_\psi) = G + \operatorname{Ker}(C_\psi)$ and $\mathcal{F} = F + \mathcal{I}(\psi)$ satisfies $T(\mathcal{F}) = G + \operatorname{Ker}(C_\psi)$.

$\square$

With all these preparations we are now able to give sufficient conditions for the restricted operator $C_\psi : \mathcal{I}(\psi(E)) \to \mathcal{I}(E)$ to have closed range. We recall the results from propositions 2.9 and 4.11 which state that both properness of ψ and the lower distance estimate from definition 4.10 are necessary for C_ψ to have closed range. We have also seen in remark 4.12 that the lower distance estimate implies the lower distance estimate for derivatives. It therefore seems quite natural that these conditions are part of the initial assumptions of our main theorem 5.2. The additional condition of the Whitney-regularity of $\psi(\mathbb{R}^q)$ is

plausible, given that it is already necessary in the one-dimensional case, but remains a concession to the methods used in our proof. We remind the reader that for a function ψ of one variable the main tool is given by the closed range theorem which states that $\mathrm{Range}(C_\psi)$ is closed if and only if every $u \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$ that vanishes on $\mathrm{Ker}(C_\psi)$ can be written as $u = C_\psi^t(v)$ with some $v \in \mathcal{A}(\psi)'$. To verify this we define v in a natural way (via explicit construction) on a dense subset of $\mathcal{A}(\psi)$ and need to find a continuous linear extension. This is where we heavily rely on the Whitney-regularity of the image. It allows us to dominate a distribution $u \in \mathrm{Ker}(C_\psi)^\perp$ with support in $\psi(\mathbb{R})$ only by seminorms on $\psi(\mathbb{R})$. Some further arguments then lead to the required continuity estimate $|v(F \circ \psi)| = |u(F)| \leq \|F\|_{\psi(K,n} \leq \|F \circ \psi\|_{K,\ell}$.
In the case of several variables, the Whitney regularity of $\psi(\mathbb{R}^q)$ will serve the same purpose. Unfortunately, it is still unclear to us if this condition is only a technical requirement for our approach or has deeper meaning. We believe that it is actually necessary for $\mathcal{A}(\psi)$ to be closed, or even for the weaker statement of $\mathcal{I}(E)$ to be contained in $\mathcal{A}(\psi)$.

Theorem 5.12.
If $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ is injective, proper, with Whitney-regular image, and satisfies the lower distance estimate for derivatives

$$\|\psi'(x)[r]\| \geq c_K \cdot \|r\| \cdot \mathrm{dist}(x, E(\psi))^{\alpha_K}$$

on every compact set K, then the restricted composition operator $C_\psi : \mathcal{R}(\psi) \to \mathcal{I}(E(\psi))$ has closed range. As a consequence $\mathcal{I}(E(\psi)) \subseteq \mathcal{A}(\psi)$ and the reduced composition operator $\tilde{C}_\psi : \mathcal{E}_0(\psi) \to \mathcal{I}(E(\psi))$ is an isomorphism.

Proof. The basic idea is to mimic our approach in the one-dimensional case.
We have already seen in proposition 5.10 that the range of $C_\psi : \mathcal{R}(\psi) \to \mathcal{I}(E(\psi))$ contains the $\mathcal{I}(E(\psi))$-dense subspace $\mathcal{D}(E(\psi)^c)$. It is therefore sufficient to show that $C_\psi(\mathcal{R}(\psi))$ is closed, which follows from the closed range theorem (26.3 in [MV97]) once we verify the inclusion $\mathrm{Ker}(C_\psi)^\perp \subseteq \mathrm{Range}(C_\psi^t)$.
Fix some $u \in \mathcal{R}(\psi)'$ that vanishes on $\mathrm{Ker}(C_\psi)$. Extending u via the Hahn-Banach theorem we can suppose without loss of generality that it is the restriction of a distribution on $\mathbb{R}^d$, which we will also denote by u.
Now we need to construct some $w \in \mathcal{I}(E(\psi))'$ such that $u(F) = w(F \circ \psi)$ for all $F \in \mathcal{R}(\psi)$. As an intermediary step, we define

$$v(F \circ \psi) = u(F)$$

for $F \in \mathcal{R}(\psi)$ satisfying $F \circ \psi \in \mathcal{D}(E(\psi)^c)$. Note that v is well-defined on $\mathcal{D}(E(\psi)^c)$ since by assumption $u \in \mathrm{Ker}(C_\psi)^\perp$, hence $v(F \circ \psi) = u(F) = 0$ whenever $F \circ \psi = 0$.
It will turn out that v can be extended continuously to the required w on $\mathcal{I}(E(\psi))$ satisfying $C_\psi^t(w) = u$. This extension will be obtained via the Hahn-Banach theorem once we have proved that we can dominate $v(f)$ by some seminorm $|||f|||_{K,n}$.

To achieve this, let us remark that once again the support of u is contained in $\psi(\mathbb{R}^q)$ since for $F \in \mathcal{D}(\psi(\mathbb{R}^q)^c) \subseteq \mathcal{R}(\psi)$ we have $F \circ \psi = 0$ and hence $u(F) = 0$. The Whitney-regularity of $\psi(\mathbb{R}^q)$, more precisely theorem 2.3.11 from [Hör03], implies the existence of some $L \subseteq \psi(\mathbb{R}^q)$ and $C, n \geq 0$ such that

$$|u(G)| \leq C \cdot \|G\|_{L,n}$$

for all $G \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$. We recall the definition 5.3

$$|||F|||_{L,n} = \sup\{\|F^{(j)}(x)[r_1, ..., r_j]\| : \|r_1\|, ..., \|r_j\| \leq 1, x \in K, j \leq n\}.$$

Moreover, according to proposition 6.1 of the appendix, the seminorms $\|\cdot\|_{L,n}$ and $|||\cdot|||_{L,n}$ are equivalent, resulting in the continuity estimate

$$|u(G)| \leq C \cdot |||G|||_{L,n}$$

for all $G \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})$.
By the properness of ψ every compact subset $L \subseteq \psi(\mathbb{R}^q)$ can be written as the image $L = \psi(K)$ of some compact subset $K \subseteq \mathbb{R}^q$. Now consider some arbitrary $F \in \mathcal{R}(\psi)$. By proposition 5.7 there is only one equivalence class $F + \mathcal{I}(\psi)$ representing $F \circ \psi$ and our estimation will use the fact that the derivatives of $F \in \mathcal{R}(\psi)$ in any point $\psi(x)$ depend only on $F \circ \psi$.
To formalize this argument, we remind the reader that, by remark 5.6 (ii), the subspace $\mathcal{R}(\psi)$ is an ideal. Now fix $x \in K$. We will show that if two functions $F, G \in \mathcal{R}(\psi)$ satisfy $F \circ \psi = G \circ \psi$ on an arbitrary neighborhood U of x, then all derivatives in $\psi(x)$ must coincide and therefore $F^{(k)}(\psi(x)) = G^{(k)}(\psi(x))$ for all $k \in \mathbb{N}_0$. To prove this, fix some open set V with $V \cap \psi(\mathbb{R}^q) \subseteq \psi(U)$ and $\Phi \in \mathcal{D}(V)$ that satisfies $\Phi = 1$ near $\psi(x)$. The products $F \cdot \Phi$ and $G \cdot \Phi$ are contained in $\mathcal{R}(\psi)$ and satisfy $(F \cdot \Phi) \circ \psi = (G \cdot \Phi) \circ \psi$. Proposition 5.7 implies that the difference $F \cdot \Phi - G \cdot \Phi = (F - G) \cdot \Phi$ is flat on $\psi(\mathbb{R}^q)$. Since $\Phi = 1$ near $\psi(x)$, we obtain that $F - G$ is flat in $\psi(x)$ and hence $F^{(n)}(\psi(x)) = G^{(n)}(\psi(x))$ for all $n \in \mathbb{N}_0$.
Let us now proceed with the actual estimates. Since $\mathcal{R}(\psi)$ is contained in $\mathcal{I}(\psi(E))$, we have $|||F^{(k)}(\psi(e))||| = 0$ for all $e \in E(\psi)$ and it is therefore sufficient to estimate $|||F^{(k)}(\psi(x))|||$ for $x \in K \setminus E(\psi)$ and $k \leq n$.
Now fix $x \in E(\psi)^c$ and consider the diffeomorphism $\Psi_x : U_x \times U_0 \to V_x$ as constructed in proposition 6.5. By localizing F with some function $\Phi \in \mathcal{D}(V_x)$ that is constant 1 near $\psi(x)$, we can suppose without loss of generality that $F \circ \psi$ has compact support in U_x.
We recall our local construction from 5.10: For $f = F \circ \psi \in \mathcal{D}(U_x)$ we set

$$F_x = (F \circ \psi) \circ \pi \circ \Psi_x^{-1},$$

where as usual $\pi : \mathbb{R}^q \times \mathbb{R}^{d-q} \to \mathbb{R}^q$ denotes the projection $(y, t) \mapsto y$. Since F and F_x generate the same composition and are both contained in $\mathcal{R}(\psi)$, we obtain the identity $|||F^{(n)}(\psi(x))||| = |||F_x^{(n)}(\psi(x))|||$. Either by formula 3.2 or the continuity of the composition with $\pi \circ \Psi_x^{-1}$, we can find $C_n > 0$ depending only on n such that

$$|||F^{(n)}(\psi(x))||| = |||F_x^{(n)}(\psi(x))||| \leq C_n \cdot |||F \circ \psi|||_{x,n} \cdot |||\pi \circ \Psi_x^{-1}|||_{\psi(x),n}.$$

Using the linearity of π we obtain $(\pi \circ G)^{(k)}(z)[r_1, ..., r_k] = \pi\left(G^{(k)}(z)[r_1, ...r_k]\right)$ for all $k \in \mathbb{N}_0$, hence $|||\pi \circ \Psi_x^{-1}|||_{\psi(x),n} \leq |||\Psi_x^{-1}|||_{\psi(x),n}$. This implies

$$|||F^{(n)}(\psi(x))||| \leq C_n \cdot |||F \circ \psi|||_{x,n} \cdot |||\Psi_x^{-1}|||_{\psi(x),n}.$$

Now we can apply the estimate (vi) from proposition 6.5 which states that

$$|||\Psi_x^{-1}|||_{\psi(x),n} \leq c_n \cdot (1 + |||\psi|||_{x,n})^{\gamma(n)} \cdot (1 + |||\psi'(x)^{-1}|||)^{\gamma(n)}$$

for some constants $c_n, \gamma(n)$ depending only on n. Inserting this in our estimate for $F^{(n)}$ and using the monotonicity with respect to n of the terms on the right hand side we have

$$|||F|||_{x,n} \leq C_n \cdot c_n \cdot |||F \circ \psi|||_{x,n} \cdot (1 + |||\psi|||_{x,n})^{\gamma(n)} \cdot (1 + |||\psi'(x)^{-1}|||)^{\gamma(n)}.$$

Dominating $1 + |||\psi|||_{x,n}$ for all $x \in K$ by some constant, we obtain some $C = C(K, n)$ such that

$$|||F|||_{x,n} \leq C(K, n) \cdot |||F \circ \psi|||_{x,n} \cdot (1 + |||\psi'(x)^{-1}|||)^{\gamma(n)}. \tag{$*$}$$

This is where the lower distance estimate for the derivative comes into play. It is required to estimate the norm of the inverse map $|||\psi'(x)^{-1}|||$. Indeed, we have seen in remark 4.12 (ii), that the lower distance estimate for the derivative implies

$$|||\psi'(x)^{-1}||| = \frac{1}{\lambda(\psi'(x))} \leq C_K \cdot \left(\frac{1}{\operatorname{dist}(x, E(\psi))}\right)^{\alpha_K}$$

for all $x \in K$. Using this inequality to further simplify our estimation $(*)$ of F, we get

$$\begin{aligned} |||F|||_{x,n} &\leq C(K, n) \cdot |||F \circ \psi|||_{x,n} \cdot \left(1 + \frac{C_K}{\operatorname{dist}(x, E)^{\alpha_K}}\right)^{\gamma(n)} \\ &\leq \tilde{C}(K, n) \cdot |||F \circ \psi|||_{x,n} \cdot \left(1 + \frac{1}{\operatorname{dist}(x, E)^{\alpha_K \cdot \gamma(n)}}\right). \end{aligned}$$

To deduce a global continuity estimate, it is sufficient to prove that for every $\gamma \in \mathbb{N}$ and compact set K there are some $N \in \mathbb{N}$ and $\tilde{K} \subseteq \mathbb{R}^q$ compact such that

$$\frac{|||F \circ \psi|||_{x,n}}{\operatorname{dist}(x, E)^{\gamma}} \leq |||F \circ \psi|||_{\tilde{K},N}$$

for every $x \in K$. This is best done by the mean value inequality. If the set $E(\psi)$ is empty, the support of every $f \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ does not touch $E(\psi)$ and we can write f as a composition by proposition 4.2. If not, $E = E(\psi)$ is a closed and nonempty set and we can find some $e_x \in E$ such that $\operatorname{dist}(x, E) = \|x - e_x\|$. By considering the closed convex hull $\tilde{K}$ of the set $M = K + B(0, R)$, where $R = \sup\{\operatorname{dist}(y, E) : y \in K\} < \infty$, we have $e_x \in \tilde{K}$ whenever $x \in K$. Moreover every directional derivative of $F \circ \psi$ is flat on E. For some normed multidirection $\mathfrak{r} = [r_1, ..., r_j]$ we define the smooth function $g : \mathbb{R} \to \mathbb{R}$ by

$$g(t) = (F \circ \psi)^{(j)}(e_x + t \cdot (x - e_x))\,[\mathfrak{r}]\,.$$

Since $F \circ \psi$ is flat in e_x, the k-th Taylor polynomial $T_0^{k-1} g(y) = \sum_{\ell=0}^{k-1} \frac{g^{(\ell)}(0)}{\ell!} y^\ell$ is the zero polynomial. By Taylor's theorem we can find some $0 < t < 1$ such that

$$
\begin{aligned}
|(F \circ \psi)^{(j)}(x)[r_1, ..., r_j]| = |g(1)| = \left|g(1) - T_0^{k-1} g(1)\right| &= \frac{1}{k!} \cdot |g^{(k)}(t)| \\
&= \quad \frac{1}{k!} \quad \cdot \quad \left|(F \circ \psi)^{(j+k)}(x + t(e_x - x)) \left[\mathfrak{r} \mathop{\boxtimes}_{\ell=1}^{k} x - e_x\right]\right| \\
&= \frac{|x - e_x|^k}{k!} \cdot \left|(F \circ \psi)^{(j+k)}(x + t(e_x - x)) \left[\mathfrak{r} \mathop{\boxtimes}_{\ell=1}^{k} \frac{x - e_x}{|x - e_x|}\right]\right|.
\end{aligned}
$$

The norm of every coordinate of the multidirection $\mathfrak{r} \boxtimes_{\ell=1}^{k} \frac{x - e_x}{|x - e_x|}$ is less or equal to 1 and the compact set $\tilde{K}$ contains the line segment $\{e_x + t \cdot (x - e_x) : 0 \leq t \leq 1\}$. We can therefore further dominate the right hand side leading to

$$
|(F \circ \psi)^{(j)}(x)[r_1, ..., r_j]| \leq \frac{|x - e_x|^k}{k!} \cdot \|\|F \circ \psi\|\|_{\tilde{K}, j+k} = \frac{\operatorname{dist}(x, E)^k}{k!} \cdot \|\|F \circ \psi\|\|_{\tilde{K}, j+k}.
$$

This estimate allows us to compensate the denominator in the continuity estimate. We obtain

$$
\frac{\|\|F \circ \psi\|\|_{x,n}}{\operatorname{dist}(x, E)^N} \leq \frac{1}{N!} \cdot \|\|F \circ \psi\|\|_{\tilde{K}, n+N},
$$

which implies

$$
\begin{aligned}
\|\|F\|\|_{x,n} &\leq \tilde{C}(K, n) \cdot \left(\|\|F \circ \psi\|\|_{x,n} + \frac{1}{N!} \cdot \|\|F \circ \psi\|\|_{\tilde{K}, n+N} \right) \\
&\leq 2 \cdot \tilde{C}(K, n) \cdot \|\|F \circ \psi\|\|_{\tilde{K}, n+N}.
\end{aligned}
$$

Taking the supremum over all $x \in K$ we obtain the required continuity estimate

$$
|v(F \circ \psi)| = |u(F)| \leq C \cdot \|\|F\|\|_{L,n} \leq \tilde{C} \cdot \|\|F \circ \psi\|\|_{\tilde{K}, n+k}.
$$

We can now apply the Hahn-Banach theorem and continuously extend v to some $w \in \mathcal{I}(E)'$. This functional satisfies $u(F) = w(F \circ \psi)$ whenever $\operatorname{supp}(F \circ \psi)$ is a compact subset of $E(\psi)^c$. This implies $u = C_\psi^t(w)$ on the set $\mathcal{D}(\psi(E)^c)$, which is dense in $\mathcal{I}(\psi(E))$. Both maps being continuous we finally obtain the required identity $u = C_\psi^t(w) \in \operatorname{Range}(C_\psi^t)$. $\square$

As an application of the previous theorem we can give an explicit way to construct a function f that is contained in $\mathcal{A}(\psi)$ but does not have an obvious factorization as $f = F \circ \psi$.

Example 5.13.
Suppose $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ satisfies the conditions of theorem 5.12. Fix a subset $J \subseteq \{1, ..., d\}$ with $|J| = q$, which we can suppose to be ordered, i.e. $J = \{j_1, ..., j_q\}$ with $j_\ell < j_{\ell+1}$. We define the selection $\psi_J : \mathbb{R}^q \to \mathbb{R}^q$ by $\psi_J(x) = (\psi_{j_1}(x), ..., \psi_{j_q}(x))$.

For any $\varphi \in \mathcal{E}(\mathbb{R},\mathbb{R})$ that is flat at zero, the function defined by $f(x) = \varphi(\det(\nabla\psi_J(x)))$ is contained in $\mathcal{A}(\psi)$ and can be factorized as $f = F\circ\psi$ with a smooth function $F \in \mathcal{E}(\mathbb{R}^d,\mathbb{R})$. As special case, the function $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = \exp\left(-\frac{1}{\det(\nabla\psi_J(x))}\right)$ is also a composition.

Proof. The function $x \mapsto \det(\nabla\psi_J(x))$ is smooth, which follows for instance from the Leibniz formula and the fact that $x \mapsto \nabla\psi_J(x)$ is smooth. Since J is a full set of coordinates, $\det(\nabla\psi_J(x))$ must vanish whenever x is a critical point. It follows that the composition with φ, which is flat in zero, must be flat on $E(\psi)$. Theorem 5.12 then implies that $f = \varphi \circ \det(\nabla\psi_J)$ is contained in $\mathcal{A}(\psi)$ and can even be factorized as $f = F \circ \psi$ with a function $F \in \mathcal{E}(\mathbb{R}^d,\mathbb{R})$ that is flat on $\psi(E(\psi))$. □

With theorem 5.12 at hand, we can give sufficient conditions for the composition algebra to be closed. As stated before, this requires knowledge about the behavior of compositions on the critical set $E(\psi)$. If we demand such a tame behavior, namely that the quotient $\mathcal{A}(\psi)/\mathcal{I}(E(\psi)$ is closed, we can formulate the following corollary.

Corollary 5.14.
Suppose $\psi \in \mathcal{E}(\mathbb{R}^q,\mathbb{R}^d)$ is injective and satisfies the following four condition:

(a) ψ is a proper map.

(b) The image $\psi(\mathbb{R}^q)$ is Whitney-regular.

(c) ψ satisfies the lower distance estimate for the derivative

$$\|\psi'(x)[r]\| \geq C_K \cdot \|r\| \cdot \operatorname{dist}(x,E)^{\alpha_K}$$

on every compact set K.

(d) The set $q_E(\mathcal{A}(\psi))$ is closed in $\mathcal{E}(E(\psi),\mathbb{R})$, where $q_E : \mathcal{E}(\mathbb{R}^d,\mathbb{R}) \to \mathcal{E}(E(\psi),\mathbb{R})$ denotes the quotient map $f \mapsto f + \mathcal{I}(E(\psi)$.

In this case the composition algebra $\mathcal{A}(\psi) = \{F \circ \psi : F \in \mathcal{E}(\mathbb{R}^d,\mathbb{R})\}$ is closed.

Proof. For a function $f \in \overline{\mathcal{A}(\psi)}$ the continuity of the quotient map and condition (d) imply that

$$q_E(f) \in q_E\left(\overline{\mathcal{A}(\psi)}\right) \subseteq \overline{q_E(\mathcal{A}(\psi))} = q_E(\mathcal{A}(\psi)).$$

This means that we can find some $F \in \mathcal{E}(\mathbb{R}^d,\mathbb{R})$ such that $f - F \circ \psi \in \mathcal{I}(E(\psi))$. By theorem 5.12 the first three conditions imply that $f - F \circ \psi \in \mathcal{A}(\psi)$, hence $f \in \mathcal{A}(\psi)$. □

5.3 Some special cases of closed algebras

Let us conclude this section with a few simple situations, where we can easily see that condition (d) is satisfied. The first one is the simplest extension of our one-dimensional result. As we have seen, in that case the Hölder continuity of the inverse map implies that the critical set $E(\psi)$ is discrete. This fact already facilitates proof of the theorem of Allan, Kakiko O'Farrell and Watson (cf. theorem 1.12) enormously.

Example 5.15.
If an injective $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ has a discrete critical set, the image $q_E(\mathcal{A}(\psi))$ is closed.

Proof. By Borel's theorem we know that the Taylor map $T_e^\infty : \mathcal{E}(\mathbb{R}^q, \mathbb{R}) \to X_q = \mathbb{R}^{\mathbb{N}_0^q}$ defined by $f \mapsto (f^{(\alpha)}(x))_{\alpha \in \mathbb{N}_0^q}$ is a continuous linear surjection. We will show that the continuous linear map

$$T_E^\infty : \mathcal{E}(\mathbb{R}^q, \mathbb{R}) \to \prod_{e \in E} X_q,\ \ F \mapsto (T_e^\infty F)_{e \in E}$$

is also surjective as a direct consequence of the discrete structure of E. Indeed, consider disjoint open neighborhoods U_e of e as well as a partition of unity $(\varphi_e)_{e \in E}$ subordinated to this cover of E. For a given element $x = (x_e)_{e \in E} \in X_q^E$ we can use Borel's theorem to obtain $F_e \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ satisfying $T_e^\infty F_e = x_e$. The function defined by

$$\sum_{e \in E} F_e \cdot \varphi_e$$

is smooth and satisfies $T_E^\infty F = x$. The Kernel of T_E is obviously given by the set $\mathcal{I}(E)$ of flat functions on E and the induced map $\tilde{T}_E^\infty : \mathcal{E}(\mathbb{R}^q, \mathbb{R})/\mathcal{I}(E) \to X_q^E$ is a continuous linear bijection. The open mapping theorem implies that both spaces are isomorphic.
Since by proposition 1.7 $\prod_{e \in E} T_e^\infty(\mathcal{A}(\psi)) = T_E^\infty \mathcal{A}(\psi)$ is closed in $\prod_{e \in E} X_q$, so must be $q_E(\mathcal{A}(\psi))$. □

Let us gather a few facts about this special situation that will lead to a characterization of closed composition algebras, given that the critical set of ψ is discrete.

Remarks 5.16.

(i) In the special case of a discrete critical set, the properness and lower distance estimate imply the local Hölder continuity of the inverse. Indeed suppose that the inverse $\psi^{-1} : \psi(\mathbb{R}^q) \to \mathbb{R}^q$ is not locally Hölder continuous. We can find $a = \psi(z) \in \psi(\mathbb{R}^q)$ such that for any $n \in \mathbb{N}$ there are $b_n = \psi(x_n)$ and $c_n = \psi(y_n)$ in $B(\psi(z), 1/n)$ such that

$$\|\psi^{-1}(b_n) - \psi^{-1}(c_n)\| > n \cdot \|b_n - c_n\|^{\frac{1}{n}}.$$

Obviously both sequences $(b_n)_{n\in\mathbb{N}}$ and $(c_n)_{n\in\mathbb{N}}$ converge to a and the continuity of the inverse, given by the properness of ψ, implies $x_n \to z$ as well as $y_n \to z$. Inserting $b_n = \psi(x_n)$ and $c_n = \psi(y_n)$ in the estimate above we obtain

$$\frac{1}{n^n} \cdot \|x_n - y_n\|^n > \|\psi(x_n) - \psi(y_n)\|. \tag{$*$}$$

The directions $r_n = \frac{x_n - y_n}{\|x_n - y_n\|}$ are all contained in the compact set $\{r \in \mathbb{R}^q : \|r\| = 1\}$ and we can take a subsequence such that the normed directions r_n converge to some r_∞ of norm 1. We can then apply the mean value theorem to every coordinate function $\gamma_j : \mathbb{R} \to \mathbb{R},\ t \mapsto \psi_j(y_n + t \cdot r_n)$ to obtain

$$\begin{aligned} |\psi_j(x_n) - \psi_j(y_n)| &= |\gamma_j(\|y_n - x_n\|) - \gamma_j(0)| = |\gamma_j'(\xi_n)| \cdot \|y_n - x_n\| \\ &= |\psi_j'(y_n + \xi_n \cdot r_n)[r_n]| \cdot \|y_n - x_n\|. \end{aligned}$$

Using the estimate $(*)$ above then implies

$$|\psi_j'(y_n + \xi_n \cdot r_n)[r_n]| = \frac{|\psi_j(x_n) - \psi_j(y_n)|}{\|y_n - x_n\|} < \frac{1}{n^n} \cdot \|x_n - y_n\|^{n-1} \to 0.$$

Since the left hand side converges to $\psi'(z)[r_\infty] = 0$, the point z is critical. As the set $E(\psi)$ is discrete, we can suppose that $\operatorname{dist}(x_n, E) = \|x_n - x_\infty\|$ as well as $\operatorname{dist}(y_n, E) = \|y_n - x_\infty\|$ for n sufficiently large.

Applying the lower distance estimate to a compact convex set K containing both $\{x_n : n \in \mathbb{N}\}$ and $\{y_n : n \in \mathbb{N}\}$ leads to the contradiction

$$\begin{aligned} \|x_n - y_n\|^{k+1} &\le \|x_n - y_n\| \cdot (\|x_n - x_\infty\| + \|y_n - x_\infty\|)^k \\ &\le 2^k \cdot \|x_n - y_n\| \cdot \max\{\operatorname{dist}(x_n, E), \operatorname{dist}(y_n, E)\}^k \\ &\le 2^k \cdot C_K \cdot \|\psi(x_n) - \psi(y_n)\| < \frac{1}{n^n} \cdot \|x_n - y_n\|^n \end{aligned}$$

for $n \to \infty$.

(ii) In the case of a smooth curve $\psi \in \mathcal{E}(\mathbb{R}, \mathbb{R}^d)$ the Hölder-continuity of the inverse implies the lower distance estimate. Indeed, by proposition 2.20, $\mathcal{A}(\psi)$ must be closed and we can derive the lower distance estimate from proposition 4.11.

With this fact at hand we are able to characterize closed composition algebras in this very restrictive case. The resulting theorem can be viewed as a generalization of proposition 2.20 to the case of a map of several variables.

Theorem 5.17.
An injective map $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ with a discrete critical set induces a closed composition algebra if and only if it is a proper map that satisfies the lower distance estimate.

Though this is already a characterization of closed composition algebras in the restrictive case of a discrete critical set, it is not the only implication of theorem 5.12. We can also give simple examples of closed composition algebras where the critical set of the generator is a union of lines.
Our next aim is to show that tensorizing two injective curves that generate closed composition algebras also leads to a closed composition algebra.
First let us make sure that the conditions of proposition 5.12 are satisfied.

Proposition 5.18.
Consider two smooth injective maps $\gamma \in \mathcal{E}(\mathbb{R}^d, \mathbb{R}^n)$ and $\varphi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^m)$ that generate closed composition algebras $\mathcal{A}(\gamma)$ and $\mathcal{A}(\varphi)$. The map $\psi : \mathbb{R}^{d+q} \to \mathbb{R}^{n+m}$ defined by

$$\psi(x, y) = \begin{pmatrix} \gamma(x) \\ \varphi(y) \end{pmatrix}$$

is again injective, proper, and satisfies the lower distance estimate for derivatives. If additionally both $\gamma(\mathbb{R}^d)$ and $\varphi(\mathbb{R}^q)$ are Whitney-regular, so is $\psi(\mathbb{R}^{d+q})$.

Proof. The injectivity of ψ is obvious. Since $\mathcal{A}(\gamma)$ and $\mathcal{A}(\varphi)$ are closed, proposition 2.9 implies that both γ and φ must be proper maps. Let us show that this easily extends to ψ. Consider a compact set $K \subseteq \mathbb{R}^{n+m}$. Both projections

$$\begin{aligned} K_I = \pi_I(K) &= \{x \in \mathbb{R}^n : \exists\, y \in \mathbb{R}^m \text{ with } (x, y) \in K\} \text{ and} \\ K_{II} = \pi_{II}(K) &= \{y \in \mathbb{R}^m : \exists\, x \in \mathbb{R}^n \text{ with } (x, y) \in K\} \end{aligned}$$

are compact and satisfy $K \subseteq K_I \times K_{II}$. The closed set $\psi^{-1}(K)$ is therefore contained in the compact set $\psi^{-1}(K_I \times K_{II}) = \gamma^{-1}(K_I) \times \varphi^{-1}(K_{II})$, hence compact.
Let us now prove that ψ also satisfies the lower distance estimate for derivatives. To this end we need precise knowledge regarding the shape of $E(\psi)$. For $(x, y) \in \mathbb{R}^{d+q}$ and a direction $(r, v) \in \mathbb{R}^{d+q}$ the derivative of ψ is given by $\psi'(x, y)[(r, v)] = (\gamma'(x)[r], \varphi'(y)[v])$. The dimension of $\operatorname{Range}(\psi'(x, y)) = \operatorname{Range}(\gamma'(x)) \times \operatorname{Range}(\varphi'(y))$ is therefore the sum of the dimensions of $\operatorname{Range}(\gamma'(x))$ and $\operatorname{Range}(\varphi'(y))$. It follows that (x, y) is a critical point if and only if $x \in E(\gamma)$ or $y \in E(\varphi)$ which implies

$$E(\psi) = (E(\gamma) \times \mathbb{R}^q) \cup \left(\mathbb{R}^d \times E(\varphi)\right).$$

The distance to $E(\psi)$ is therefore given by

$$\operatorname{dist}((x, y), E(\psi)) = \min \{\operatorname{dist}(x, E(\gamma)), \operatorname{dist}(y, E(\varphi))\}.$$

Now fix a compact set $K \subseteq \mathbb{R}^{d+q}$ and denote by K_I and K_{II} the projections on $\mathbb{R}^d$ and $\mathbb{R}^q$ as above. By proposition 4.11 both γ and φ must satisfy the lower distance estimate which implies its counterpart for the respective derivatives by remark 4.12. We can therefore find constants $C_\gamma, \alpha_\gamma > 0$ and $C_\varphi, \alpha_\varphi > 0$ such that

$$\begin{aligned} \|\gamma'(x)[r]\| &\geq C_\gamma \cdot \|r\| \cdot \operatorname{dist}(x, E(\gamma))^{\alpha_\gamma} \text{ and} \\ \|\varphi'(y)[v]\| &\geq C_\varphi \cdot \|v\| \cdot \operatorname{dist}(y, E(\varphi))^{\alpha_\varphi}. \end{aligned}$$

Defining $C = \min\{C_\varphi, C_\gamma\}$ and using $\psi'(x,y)[(r,v)] = (\gamma'(x)[r], \varphi'(y)[v])$ we obtain the estimate

$$\begin{aligned}\|\psi'(x,y)[(r,v)]\|^2 &= \|\gamma'(x)[r]\|^2 + \|\varphi'(y)[v]\|^2 \\ &\geq C^2 \cdot \|r\|^2 \cdot \operatorname{dist}(x, E(\gamma))^{2\alpha_\gamma} + C^2 \cdot \|v\|^2 \cdot \operatorname{dist}(y, E(\varphi))^{2\alpha_\varphi} \\ &\geq C^2 \cdot \left(\|r\|^2 + \|v\|^2\right) \cdot \min\left\{\operatorname{dist}(x, E(\gamma))^{2\alpha_\gamma}, \operatorname{dist}(y, E(\varphi))^{2\alpha_\varphi}\right\} \\ &\geq C^2 \cdot \left(\|r\|^2 + \|v\|^2\right) \cdot \min\left\{\operatorname{dist}((x,y), E(\psi))^{2\alpha_\gamma}, \operatorname{dist}((x,y), E(\psi))^{2\alpha_\varphi}\right\}\end{aligned}$$

for all $(x,y) \in K$. Taking the square root on both sides and writing $z = (x,y)$ and $w = (r,v)$ we get

$$\|\psi'(z)[w]\| \geq C \cdot \|w\| \cdot \min\left\{\operatorname{dist}(z, E(\psi))^{\alpha_\gamma}, \operatorname{dist}(z, E(\psi))^{\alpha_\varphi}\right\}$$

for all $z \in K$. It remains to prove that we can find $c, \beta > 0$ such that

$$\min\left\{\operatorname{dist}(z, E(\psi))^{\alpha_\gamma}, \operatorname{dist}(z, E(\psi))^{\alpha_\varphi}\right\} \geq c \cdot \operatorname{dist}(z, E(\psi))^\beta$$

for all $z \in K$. To this end, consider $d = 1 + \sup\{\operatorname{dist}(a, E(\psi)) : a \in K\} < \infty$. For all $z \in K$ we have $\frac{1}{d} \cdot \operatorname{dist}(z, E(\psi)) \leq 1$, hence the trivial estimate

$$\frac{\operatorname{dist}(z, E(\psi))^\beta}{d^\beta} \leq \frac{\operatorname{dist}(z, E(\psi))^\alpha}{d^\alpha}$$

for all $0 < \alpha \leq \beta$. Applying this to $\beta = \max\{\alpha_\gamma, \alpha_\varphi\}$ and $\alpha \in \{\alpha_\gamma, \alpha_\varphi\}$ leads to the required

$$c \cdot \operatorname{dist}(z, E(\psi))^\beta \leq \min\left\{\operatorname{dist}(z, E(\psi))^{\alpha_\gamma}, \operatorname{dist}(z, E(\psi))^{\alpha_\varphi}\right\},$$

where $c = \min\left\{d^{\alpha_\gamma}/d^\beta, d^{\alpha_\varphi}/d^\beta\right\}$.

To prove the second statement, let us now suppose that $\gamma(\mathbb{R}^d)$ and $\varphi(\mathbb{R}^q)$ are Whitney-regular sets. To prove that $\psi(\mathbb{R}^{d+q})$ is also Whitney-regular, we need to show the following: For every $\psi(x_0, y_0) \in \psi(\mathbb{R}^{d+q})$ we can find a neighborhood U of $\psi(x_0, y_0)$ such that any two points t, s in $U_\psi = U \cap \psi(\mathbb{R}^{d+q})$ can be joined by a curve in U_ψ of length not greater than $C_U \cdot \|t - s\|^{\alpha_U}$. By assumption we can find a neighborhood $V \subseteq \mathbb{R}^d$ of $\gamma(x_0)$ and $W \subseteq \mathbb{R}^q$ of $\varphi(y_0)$ such that this is true for V and W. Define $U = V \times W$. For $\psi(x,y), \psi(z,r) \in U$ we have $\gamma(x), \gamma(z) \in V$ as well as $\varphi(y), \varphi(r) \in W$. By assumption there is a rectifiable curve λ_I joining $\gamma(x), \gamma(z)$ in $V \cap \gamma(\mathbb{R}^d)$, and λ_{II} joining $\varphi(y)$ and $\varphi(r)$ in $W \cap \varphi(\mathbb{R}^q)$ with suitable lengths. Without loss of generality, we can suppose them to be parametrized over $[0,1]$. We define $\lambda(t) = (\lambda_I(t), \lambda_{II}(t))$. The length $L(\lambda)$ can then be estimated by the lengths $L(\lambda_I)$ and $L(\lambda_{II})$ of the the respective curves λ_I and λ_{II} in the following way:

$$\begin{aligned}L(\lambda) \leq L(\lambda_I) + L(\lambda_{II}) &\leq C_V \cdot \|\gamma(x) - \gamma(z)\|^{\alpha_V} + C_W \cdot \|\varphi(y) - \varphi(r)\|^{\alpha_W} \\ &\leq C_V \cdot \|\psi(x,y) - \psi(z,r)\|^{\alpha_V} + C_W \cdot \|\psi(x,y) - \psi(z,r)\|^{\alpha_W},\end{aligned}$$

which again implies the estimate

$$L(\lambda) \leq \tilde{C} \cdot (\|\psi(x,y) - \psi(z,r)\|^\alpha)$$

for $\alpha = \min\{\alpha_V, \alpha_W\}$ and a suitable $\tilde{C}$. □

Corollary 5.19.
If γ and φ are injective curves generating closed composition algebras and $\psi : \mathbb{R}^2 \to \mathbb{R}^{d+q}$ is defined by $\psi(t,s) = (\gamma(t), \varphi(s))$, then $\mathcal{I}(E(\psi) \subseteq \mathcal{A}(\psi)$ holds.

Corollary 5.20.
If $\gamma \in \mathcal{E}(\mathbb{R}^n, \mathbb{R}^q)$ and $\varphi \in \mathcal{E}(\mathbb{R}^m, \mathbb{R}^d)$ are injective maps with Hölder continuous inverses and both satisfy the lower distance estimate for derivatives, then $\psi(x,y) = (\gamma(x), \varphi(y))$ also satisfies $\mathcal{I}(E(\psi)) \subseteq \mathcal{A}(\psi)$.

The next step is to verify condition (iv) of proposition 5.14. To do this we need to prove that for every $f \in \overline{\mathcal{A}(\psi)}$ one can find $F \in \mathcal{E}(\mathbb{R}^{n+m}, \mathbb{R})$ such that $f - F \circ \psi$ is flat on $E(\psi) = (E(\gamma) \times \mathbb{R}) \cup (\mathbb{R} \times E(\varphi))$. Let us start with a simpler example where the second curve γ is given by the identity on $\mathbb{R}$.

Proposition 5.21.
Consider an injective curve $\gamma \in \mathcal{E}(\mathbb{R}, \mathbb{R}^n)$ generating a closed composition algebra $\mathcal{A}(\gamma)$. The injective smooth map $\psi : \mathbb{R}^2 \to \mathbb{R}^{n+1}$ defined by $\psi(t,s) = (\gamma(t), s)$ also generates a closed composition algebra.

Proof. We have already seen in the previous proposition 5.18 that ψ satisfies the conditions (i)-(iii) of proposition 5.14. It therefore remains to verify condition (iv), namely that the quotient $\mathcal{A}(\psi)/\mathcal{I}(E(\psi)$ is closed. Let us note that the critical set of ψ is given by $E(\psi) = E(\gamma) \times \mathbb{R}$ and that $E(\gamma)$ is discrete as a consequence of proposition 2.13 since $\mathcal{A}(\gamma)$ is closed by assumption.
To prove that $\mathcal{A}(\psi)$ is closed, we will construct for $f \in \overline{\mathcal{A}(\psi)}$ some $F \in \mathcal{E}(\mathbb{R}^{n+1}, \mathbb{R})$ such that $f - F \circ \psi$ is flat on $E(\psi)$. By proposition 5.12 we then obtain that $f - F \circ \psi \in \mathcal{I}(E(\psi))$ is also a composition $G \circ \psi$, hence $f = (F + G) \circ \psi$.
Let us first explain why it is sufficient to only construct $F_e \in \mathcal{E}(\mathbb{R}^{n+1}, \mathbb{R})$ such that $f - F_e \circ \psi$ is flat on $E(\gamma) \times \mathbb{R}$. The basic idea is to glue together local solutions on a line by using an appropriate partition of unity. The set $E(\gamma)$ is discrete and γ is a proper map, hence $\gamma(E(\gamma))$ is also discrete. We can therefore find a disjoint open cover $(U_e)_{e \in E(\gamma)}$ of $\gamma(E(\gamma))$ inducing a disjoint open cover $(U_e \times \mathbb{R})_{e \in E(\gamma)}$ of $\gamma(E(\gamma)) \times \mathbb{R} = \psi(E(\psi))$. Considering functions $\Phi_e \in \mathcal{D}(U_e \times \mathbb{R})$ that are constant 1 on a neighborhood of $\{\gamma(e)\} \times \mathbb{R}$ we can piece together the local solutions by defining $F = \sum_{e \in E(\gamma)} F_e \cdot \Phi_e$. Since every compact set in $\mathbb{R}^{n+1}$ intersects only finitely many open sets $U_e \times \mathbb{R}$, this function is well-defined and smooth. Moreover we have $F = F_e$ on a neighborhood V_e of $\{\gamma(e)\} \times \mathbb{R}$, hence $F \circ \psi$ and $F_e \circ \psi$ coincide on the open neighborhood $\psi^{-1}(V_e)$ of $\{e\} \times \mathbb{R}$. This obviously implies that $(F - F_e) \circ \psi$ is flat on $\{e\} \times \mathbb{R}$ for every $e \in E(\gamma)$ and therefore $f - F \circ \psi$ is flat on $E(\psi)$.
Now fix $f \in \overline{\mathcal{A}(\psi)}$ and $e \in E(\gamma)$. To construct F_e explicitly we need to refine proposition 1.7. This result states that that for every $f \in \mathcal{A}(\gamma)$ there is an $F \in \mathcal{E}(\mathbb{R}^n, \mathbb{R})$ such that $f - F \circ \gamma$ is flat in e. We will show that one can compute the Taylor coefficients of some F satisfying $T_e^\infty(F \circ \psi) = T_e^\infty f$ explicitly.

To prove this we identify again

$$\mathcal{E}(\mathbb{R}^n,\mathbb{R})/\mathcal{I}(\{\gamma(e)\}) \cong \mathbb{R}^{\mathbb{N}_0^n} \qquad \text{and} \qquad \mathcal{E}(\mathbb{R},\mathbb{R})/\mathcal{I}(\{e\}) \cong \mathbb{R}^{\mathbb{N}_0}$$

via the isomorphisms given by the Taylor maps $T^\infty_{\gamma(e)}$ and T^∞_e. We consider the continuous linear map $\tau : \mathbb{R}^{\mathbb{N}_0^n} \to \mathbb{R}^{\mathbb{N}_0}$ defined by $\tau(T^\infty_{\gamma(e)}F) = T^\infty_e(F \circ \gamma)$, which is continuous and linear. The Kernel of τ is a closed subspace of $\mathbb{R}^{\mathbb{N}_0^n}$ and we will use the fact that it is complemented, which we deduce from the book of Bonet and Pérez Carreras [PB87]. To this end we recall the definition of minimal locally convex spaces (definition 2.6.3 in [PB87]). A locally convex space is minimal if there exists no coarser locally convex separated vector space topology. This is obviously the case for $\mathbb{R}^{\mathbb{N}_0^n}$. Moreover, by theorem 2.6.4 (ii) of [PB87] every closed subspace of a minimal locally convex space is also minimal. Corollary 2.6.5 (iii) in [PB87] states that minimal subspaces are complemented. Applying this to the kernel L of τ we obtain a complementary subspace R. Obviously R and $T^\infty_e\mathcal{A}(\gamma)$ are closed subspaces of Fréchet spaces and hence Fréchet themselves. Let us prove that the restriction of $\tau : R \to T^\infty_e\mathcal{A}(\gamma)$ is a continuous linear bijection. The map τ is surjective since for $T^\infty_e(F \circ \gamma) \in T^\infty_e\mathcal{A}(\gamma)$ we can decompose $T^\infty_{\gamma(e)}F = \ell + r$, where $r \in R$ and $\ell \in L$, hence $\tau(T^\infty_{\gamma(e)}F) = \tau(r)$. The function τ is also injective since $\tau(r) = 0$ implies $r \in \mathrm{Ker}(\tau)$, which gives $r \in R \cap L = \{0\}$. As bijection between Fréchet spaces, τ is an isomorphism and hence possesses a continuous inverse θ.

Let us now extend this fact and add another variable. Our main tool will be a slight generalization of Borel's theorem (cf. proposition 6.2 of the appendix). This version states that for every collection $\{f_\alpha : \alpha \in \mathbb{N}_0^n\} \subseteq \mathcal{E}(\mathbb{R}^q,\mathbb{R})$ one can find $F \in \mathcal{E}(\mathbb{R}^{n+q},\mathbb{R})$ such that $D^{(\alpha,0_q)}F\underline{(0_n,y)} = f_\alpha(y)$ for all multi-indices $\alpha \in \mathbb{N}_0^n$.

For $f \in \underline{\mathcal{A}(\psi)}$, we consider the collection given by $f_j(y) = f^{(j,0)}(e,y)$, where $j \in \mathbb{N}_0$. By proposition 1.7, f must satisfy the pointwise Taylor condition $T^\infty_x f \in T^\infty_x\mathcal{A}(\psi)$. One can therefore find $F_{(e,y)} \in \mathcal{E}(\mathbb{R}^{n+1},\mathbb{R})$ such that $T^\infty_{(e,y)}f = T^\infty_{(e,y)}(F_{(e,y)} \circ \psi)$. This implies the identity $D^{(j,0)}f(e,y) = D^{(j,0)}(F_{(e,y)} \circ \psi)(e,y)$ for all $j \in \mathbb{N}_0$. Fix $y \in \mathbb{R}$, the map defined by $F_y(x) = F_{(e,y)}(x,y)$ is smooth. We also obtain the identity $D^{(j,0)}f(e,y) = D^j(F_y \circ \gamma)(e)$ and hence the sequence $(f_j(y))_{j\in\mathbb{N}_0}$ is contained in $T^\infty_e\mathcal{A}(\gamma)$.

For $\alpha \in \mathbb{N}_0^n$ we define

$$F_\alpha(y) = \theta_\alpha((f_j(y))_{j\in\mathbb{N}_0})$$

where $\theta_\alpha = \pi_\alpha \circ \theta$ and θ is the inverse of τ defined above by $\tau(T^\infty_{\gamma(e)}G) = T^\infty_e(G \circ \gamma)$. The map $\theta_\alpha : T^\infty_e\mathcal{A}(\gamma) \to \mathbb{R}$ is continuous and linear and can therefore be written as

$$\theta_\alpha((f_j)_{j\in\mathbb{N}_0}) = \sum_{j=0}^{N(\alpha)} c_{j,\alpha} \cdot f_j.$$

We obtain that for any $\alpha \in \mathbb{N}_0^n$

$$F_\alpha(y) = \sum_{j=0}^{N(\alpha)} c_{j,\alpha} D^{(j,y)} f(e,y),$$

which implies that F_α is a smooth map.
Applying Borel's theorem (cf. proposition 6.2) we can find $F \in \mathcal{E}(\mathbb{R}^{n+1}, \mathbb{R})$ satisfying $D^{(\alpha,0)}F(e,y) = F_\alpha(y)$ for all $\alpha \in \mathbb{N}_0^n$. It remains to show that $f - F \circ \psi$ is flat on $\{e\} \times \mathbb{R}$. By construction we already have $D^{(j,0)}(F \circ \psi)(e,y) = D^{(j,0)}f(e,y)$ for all $\alpha \in \mathbb{N}_0^n$. Computing the remaining partial derivatives we see that this also implies

$$D^{(j,k)}(F \circ \psi)(e,y) = D^{(0,k)}D^{(j,0)}(F \circ \psi)(e,y) = D^{(0,k)}D^{(j,0)}f(e,y) = D^{(j,k)}f(e,y).$$

□

Remark 5.22.
Note that if $\psi \in \mathcal{E}(\mathbb{R}^{n+m}, \mathbb{R}^{d+q})$ has the structure $\psi(x,y) = (\gamma(x), \varphi(y))$ and generates a closed composition algebra $\mathcal{A}(\psi)$, it directly follows that both $\mathcal{A}(\varphi)$ and $\mathcal{A}(\gamma)$ are closed. Indeed consider $f \in \overline{\mathcal{A}(\gamma)} \subseteq \mathcal{E}(\mathbb{R}^n, \mathbb{R}^d)$, then one can find a sequence $(F_n)_{n\in\mathbb{N}}$ in $\mathcal{E}(\mathbb{R}^n, \mathbb{R})$ such that $F_n \circ \gamma \to f$. Defining $G_n \in \mathcal{E}(\mathbb{R}^{n+m}, \mathbb{R})$ by $G_n(x,y) = F_n(x)$ we see that $G_n \circ \psi$ converges to some $g \in \mathcal{E}(\mathbb{R}^{n+m}, \mathbb{R})$. By assumption $\mathcal{A}(\psi)$ is closed and we can write $g = G \circ \psi$ for some $G \in \mathcal{E}(\mathbb{R}^{n+m}, \mathbb{R})$. Defining $F(z) = G(z, \varphi(0))$ for $z \in \mathbb{R}^n$ we obtain

$$F(\gamma(x)) = G(\gamma(x), \varphi(0)) = G \circ \psi(x,0) = \lim_{n\to\infty} G_n \circ \psi(x,0) = \lim_{n\to\infty} F_n(x) = f(x),$$

hence $f = F \circ \gamma$. As this is true for all $f \in \overline{\mathcal{A}(\gamma)}$, we see that $\mathcal{A}(\gamma)$ must be closed and for reasons of symmetry the same argument implies that $\mathcal{A}(\varphi)$ is also closed.

To generalize the idea of the previous example to some $\psi(x,y) = \gamma(x) \boxtimes \varphi(y)$ we will use the notion of regularly situated sets as found in [Mal67], definition 5.4. This is a very powerful result which we can use to decompose closed sets and related spaces of flat functions. Basically it gives precise conditions when we have $\mathcal{I}(X \cap Y) = \mathcal{I}(X) \oplus \mathcal{I}(Y)$. The main reason for the use of this technique, aside from being easy to check in the special cases we consider and thus simplifying our computations, is that this property might be useful to decompose the critical set in a less restrictive setting. This could be way to obtain a generalization of the result of Allan, Kakiko, O'Farrell and Watson ([AKOW98]) in the case of several variables and therefore a description of the closure of $\mathcal{A}(\psi)$.

Definition 5.23 ([Mal67], Def. 5.4)**.**
For $Z \subseteq \mathbb{R}^q$ we abbreviate $\mathcal{E}(Z) = \mathcal{E}(Z, \mathbb{R}) = \mathcal{E}(\mathbb{R}^q, \mathbb{R})/\mathcal{I}(Z)$ for the space of smooth germs on Z. For a smooth function $F \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ we will write $\mathcal{F}_Z = \left(D^{(\alpha)}F|_Z\right)_{\alpha\in\mathbb{N}_0^q}$ for the germ on Z generated by F.
Two closed subsets X and Y of $\mathbb{R}^q$ are called regularly situated if the sequence

$$0 \to \mathcal{E}(X \cup Y) \xrightarrow{\delta} \mathcal{E}(X) \oplus \mathcal{E}(Y) \xrightarrow{\varrho} \mathcal{E}(X \cap Y) \to 0$$

is exact, where $\delta(\mathcal{F}_{X\cup Y}) = (\mathcal{F}_X, \mathcal{F}_Y)$ and $\varrho(\mathcal{F}_X, \mathcal{G}_Y) = \mathcal{F}_{X\cap Y} - \mathcal{G}_{X\cap Y}$.

Remark 5.24.

1. We can interpret the exactness of the sequence above as a geometrical decomposition property for flat functions on $X \cap Y$, more precisely it implies
$$\mathcal{I}(X \cap Y) = \mathcal{I}(X) + \mathcal{I}(Y).$$
The implication " $\supseteq$ " is obvious and always true, no matter X and Y. To prove " $\subseteq$ " note that the exactness of the sequence in definition 5.23 is equivalent to $\mathrm{Range}(\delta) = \mathrm{Ker}(\varrho)$. For $F \in \mathcal{I}(X \cap Y)$ we see that the related germ $\mathcal{F}_X$ satisfies $\varrho(\mathcal{F}_X, 0) = 0$, hence $(\mathcal{F}_X, 0)$ is contained in $\mathrm{Ker}(\varrho) = \mathrm{Range}(\delta)$. This implies that we can find $\mathcal{G} \in \mathcal{E}(X \cup Y)$ and therefore some representative $G \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ such that $D^{(\alpha)}G(x) = D^{(\alpha)}F(x)$ for all $x \in X$ as well as $D^{(\alpha)}G(y) = 0$ for all $y \in Y$. We obtain $G \in \mathcal{I}(Y)$ and $F - G \in \mathcal{I}(X)$, hence
$$F = F - G + G \in \mathcal{I}(X) + \mathcal{I}(Y)$$

2. By theorem 5.5 in [Mal67] two closed sets $X, Y \subseteq \mathbb{R}^q$ are regularly situated if and only if they are either disjoint or satisfy the following geometric inequality.

   For any compact subsets $K \subseteq X$ and $L \subseteq Y$ there exist $C, \alpha > 0$ such that
$$\mathrm{dist}(x, L) \geq C \cdot \mathrm{dist}(x, X \cap Y)^\alpha$$
   for every $x \in K$.

   Tougeron uses this inequality as defining property of regularly situated sets (definition 4.4. in [Tou72]). He then shows the equivalence to the exactness of the sequence $0 \to \mathcal{E}(X \cup Y) \xrightarrow{\delta} \mathcal{E}(X) \oplus \mathcal{E}(Y) \xrightarrow{\varrho} \mathcal{E}(X \cap Y) \to 0$ in proposition 4.7.

Proposition 5.25.
Consider some injective $\psi : \mathbb{R}^2 \to \mathbb{R}^{d+q}$ defined by $\psi(t, s) = (\gamma(t), \varphi(s))$. If both $\mathcal{A}(\gamma)$ and $\mathcal{A}(\varphi)$ are closed, so is $\mathcal{A}(\psi)$.

Proof. As mentioned in corollary 5.19 the map ψ satisfies the conditions of proposition 5.12, hence $\mathcal{I}(E(\psi)) \subseteq \mathcal{A}(\psi)$. It remains to prove that for every $f \in \overline{\mathcal{A}(\psi)}$ one can find a composition $F \circ \psi$ such that $f - F \circ \psi$ is flat on $E(\psi)$.
Now fix some $f \in \overline{\mathcal{A}(\psi)}$ and define $X = E(\gamma) \times \mathbb{R}$ as well as $Y = \mathbb{R} \times E(\varphi)$. We will again construct some $F \in \mathcal{E}(\mathbb{R}^{n+m}, \mathbb{R})$ satisfying $f - F \circ \psi \in \mathcal{I}(E(\psi))$. To simplify this task we need to decompose f.
First we will construct a function $G \in \mathcal{E}(\mathbb{R}^{d+q}, \mathbb{R})$ such that the difference $f - G \circ \psi$ is flat on $E(\gamma) \times E(\varphi) = X \cap Y$. By assumption γ and φ are both injective and generate closed composition algebras. Proposition 2.13 states that $E(\gamma)$ and $E(\varphi)$ are discrete and hence their product $X \cap Y$ must be discrete as well. The properness of ψ implies that $\psi(X \cap Y)$ is also discrete. We can therefore find disjoint open neighborhoods $U_{e,c}$ of $(\gamma(e), \varphi(c))$ of $(e, c) \in E(\gamma) \times E(\varphi)$. Considering $\Phi_{e,c} \in \mathcal{D}(U_{e,c})$ that is constant one near

$(\gamma(e), \varphi(c)) = \psi(e, c)$ as well as $F_{e,c} \in \mathcal{E}(\mathbb{R}^{d+q}, \mathbb{R})$ such that $f - F_{e,c} \circ \psi$ is flat in (e, c) we can construct

$$G = \sum_{e \in E(\gamma),\ c \in E(\varphi)} F_{e,c} \cdot \Phi_{e,c}.$$

This map is well-defined and smooth since every compact subset of $\mathbb{R}^{d+q}$ only intersects finitely many neighborhoods $U_{e,c}$. Moreover, the resulting difference $f - G \circ \psi$ is flat on the set $E(\gamma) \times E(\varphi) = X \cap Y$.
We can now suppose without loss of generality that $f \in \mathcal{I}(X \cap Y)$ and will further decompose f into the summ $f_X + f_Y$ with $f_X \in \mathcal{I}(Y)$ and $f_Y \in \mathcal{I}(X)$ using remark 5.24 (i). To do this we will show that X and Y are regularly situated. This is best done by checking the equivalent inequality mentioned in remark 5.24 (ii), which will follow from the product structure of X and Y. To verify the inequality fix compact subsets $K \subseteq X$ and $L \subseteq Y$. For $x = (e_\gamma, s) \in K \subseteq E(\gamma) \times \mathbb{R}$ the inclusion $L \subseteq Y$ implies

$$\begin{aligned} \operatorname{dist}(x, L) \geq \operatorname{dist}(x, Y) &= \operatorname{dist}((e_\gamma, s), \mathbb{R} \times E(\varphi)) = \operatorname{dist}(s, E(\varphi)) \\ &= \operatorname{dist}((e_\gamma, s), E(\gamma) \times E(\varphi)) = \operatorname{dist}(x, X \cap Y). \end{aligned}$$

Note that both f_X and f_Y satisfy the pointwise Taylor condition and it will follow by [Tou71] Theorem 1.1 that they are therefore contained in $\overline{\mathcal{A}(\psi)}$. To apply this result are going to verify that ψ has a locally Hölder continuous inverse, which is due to the fact that $\psi = (\gamma, \varphi)$ where both γ and φ have locally Hölder continuous inverses.
Now we require to construct $F_X, F_Y \in \mathcal{E}(\mathbb{R}^{d+q}, \mathbb{R})$ such that both $f_X - F_X \circ \psi$ and $f_Y - F_Y \circ \psi$ are flat on $E(\psi) = X \cup Y$, which we will do similarly to the previous example. For reasons of symmetry it is sufficient to construct F_X, as the same techniques show how to gain F_Y. Now fix $f_X \in \overline{\mathcal{A}(\psi)}$ that is flat on $Y = \mathbb{R} \times E(\varphi)$.
For $e \in E(\gamma)$ we will construct a local solution F_e such that $f_X - F_e \circ \psi$ is flat on both $\{e\} \times \mathbb{R}$ and Y. Define the map $\tau_e : \mathbb{R}^{\mathbb{N}_0^d} \to T_e^\infty \mathcal{A}(\gamma)$ by

$$\tau_e(T^\infty_{\gamma(e)} H) = T_e^\infty (H \circ \gamma),$$

where we have again identified $\mathbb{R}^{\mathbb{N}_0^d}$ with the quotient $\mathcal{E}(\mathbb{R}^d, \mathbb{R})/\mathcal{I}(\gamma(e))$ via the isomorphism $H + \mathcal{I}(\psi(e)) \mapsto T^\infty_{\gamma(e)} H$. The kernel of τ_e is a minimal space, as closed subspace of the minimal locally convex space $\mathbb{R}^{\mathbb{N}_0^d}$, hence complemented by [PB87], corollary 2.6.5 (iii). Its complementary subspace L_e is therefore closed and the restriction $\tau_e : L_e \to T_e^\infty \mathcal{A}(\gamma)$ is a continuous linear bijection between Fréchet spaces and hence an isomorphism. We denote its inverse by $\theta_e : T_e^\infty \mathcal{A}(\gamma) \to L_e$ and write θ_e^α for the appropriate coordinate $\pi_\alpha \circ \theta_e$, where $\pi_\alpha : \mathbb{R}^{\mathbb{N}_0^d} \to \mathbb{R}$ is the usual projection and $\alpha \in \mathbb{N}_0^d$. Since $\theta_e^\alpha : T_e^\infty \mathcal{A}(\gamma) \to \mathbb{R}$ is a continuous linear map with respect to the product topology on $T_e^\infty \mathcal{A}(\gamma) \subseteq \mathbb{R}^{\mathbb{N}_0}$, we can find $N(\alpha) \in \mathbb{N}$ and coefficients $c_k \in \mathbb{R}$ such that

$$\theta_e^\alpha((\lambda_j)_{j \in \mathbb{N}_0}) = \sum_{k=0}^{N(\alpha)} c_k \cdot \lambda_k.$$

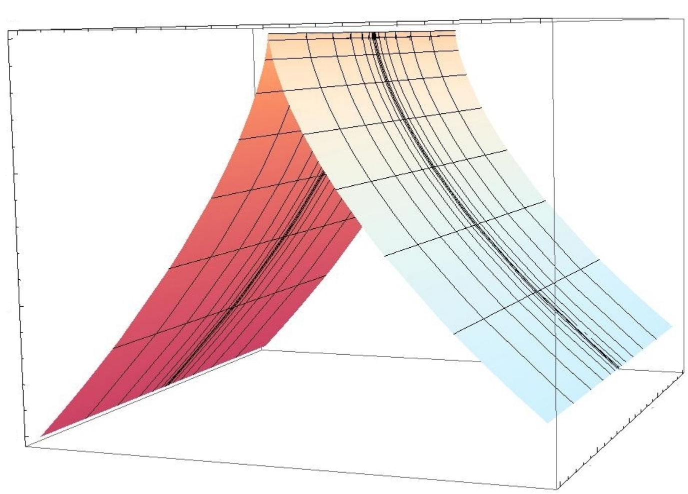
For $\alpha \in \mathbb{N}_0^d$ we define
$$h_e^\alpha(s) = \theta_e^\alpha((D^{(j,0)} f_X(e,s))_{j\in\mathbb{N}_0}),$$
which is a smooth map on $\mathbb{R}$ given the form of $\theta_e^\alpha : T_e^\infty \mathcal{A}(\gamma) \to \mathbb{R}$ as a finite sum. Moreover the representation of θ_e^α leads to the identity

$$D^k h_e^\alpha(s) = \sum_{j=0}^{N(\alpha)} c_k \cdot D^{(j,k)} f_X(e,s)$$

for the derivatives of h_e^α. Since f_X is flat on $E(\gamma) \times E(\varphi)$, this implies that h_e^α is flat on $E(\varphi)$. By assumption $\mathcal{A}(\varphi)$ is closed and we can find some function $H_e^\alpha \in \mathcal{E}(\mathbb{R}^q, \mathbb{R})$ such that $H_e^\alpha \circ \varphi = h_e^\alpha$. By proposition 5.12 we can even chose H_e^α to be flat on $\varphi(E(\varphi))$. To construct F_e we use Borel's theorem (proposition 6.2 of the appendix) to obtain F_e satisfying $D^{(\alpha,0_q)} F_e(e,y) = H_e^\alpha(y)$. This computation shows that $f_X - F_e \circ \psi$ is flat on $Y \cup (\{e\} \times \mathbb{R})$.
Now let us piece together the local solutions. Since $\gamma(E(\gamma))$ is discrete, we can find a disjoint cover U_e of $\gamma(e) \in \gamma(E(\gamma))$ generating another disjoint cover $V_e = U_e \times \mathbb{R}^q$ of $\psi(X)$. For a partition of unity $(\Phi_e)_{e\in E(\gamma)}$ subordinated to the cover $(V_e)_{e\in E(\gamma)}$ of $\gamma(E(\gamma)) \times \mathbb{R}^q$ and some F_e such that $f_X - F_e \circ \psi$ is flat on $Y \cup (\{e\} \times \mathbb{R})$, the function defined by $F_X = \sum_{e\in E(\gamma)} F_e \cdot \Phi_e$ satisfies $f_X - F_X \circ \psi \in \mathcal{I}(X)$ as well as $f_X - F_X \circ \psi \in \mathcal{I}(Y)$. □

Example 5.26.
As an example of the applicability of the previous proposition we consider $\psi : \mathbb{R}^2 \to \mathbb{R}^3$ defined by
$$\psi(t,s) = \begin{pmatrix} t^3 \\ t^2 \cdot \mathrm{Arctan}((t-3)^2) \\ s^3 \end{pmatrix}.$$

A sketch of the image of $\psi([0,1]^2)$ is depicted below.



We will now verify that $\mathcal{A}(\psi)$ is closed by using proposition 5.25 on the curves defined by $\gamma(t) = (t^3, t^2 \cdot \operatorname{Arctan}(t-3)^2)$ and $\varphi(s) = s^3$. Let us first note that φ generates a closed composition algebra since it has a locally Hölder continuous inverse. For $r < 0 < s$ we have

$$|s-r|^3 \leq (|s|+|r|)^3 \leq 8\max\{|r|,|s|\}^3 \leq 8|r|^3 + 8|s|^3 = 8|r^3 - s^3|,$$

and for $\operatorname{sgn}(r) = \operatorname{sgn}(s)$ we have

$$|r-s|^3 \leq |r-s| \cdot |r+s|^2 \leq |(r-s)(r^2+s^2+rs)| = |r^3 - s^3|.$$

This also shows that the curve γ possesses a locally Hölder continuous inverse since we have $\|\gamma(t) - \gamma(v)\| \geq |t^3 - v^3| \geq |t-v|^3$. By proposition 2.20 both curves generate closed composition algebras, hence $\mathcal{A}(\psi)$ is closed. One can even show that ψ has a local Hölder continuous inverse.

Moreover, we can use the special case of tensorized curves to verify the conditions of 5.12 in some specific cases as we demonstrate in the following example.

Example 5.27.
(i) Consider the map $\psi : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$\psi(x) = \begin{pmatrix} \exp\left(\frac{x^2+y^2}{2}\right) \\ x^3 \\ y^3 \end{pmatrix}.$$

Its critical set is given by $E(\psi) = \{(x,y) \in \mathbb{R}^2 : x \cdot y = 0\}$ and $\mathcal{I}(E(\psi)) \subseteq \mathcal{A}(\psi)$ holds. To better grasp the geometrical structure of ψ we have given a sketch below.

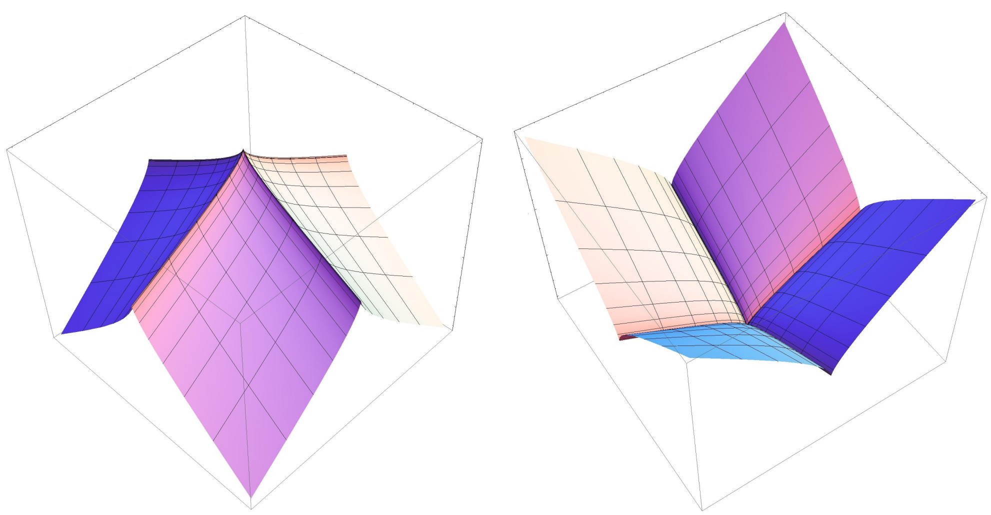


Figure 5.1: The image $\psi([-1,1]^2)$ from two different angles, above and beyond.

To verify our claim that $E(\psi)$ is the union of the sets $\mathbb{R} \times \{0\}$ and $\{0\} \times \mathbb{R}$ we need to compute the derivative of ψ. The matrix representation $\nabla\psi(x, y)$ of the linear map $\psi'(x, y)$ is given by

$$\nabla\psi(x,y) = \begin{bmatrix} x \cdot \exp\left(\frac{x^2+y^2}{2}\right) & y \cdot \left(\frac{x^2+y^2}{2}\right) \\ 3 \cdot x^2 & 0 \\ 0 & 3 \cdot y^2 \end{bmatrix}.$$

One can easily see that its rank is 0 if and only if either $x = 0$ or $y = 0$. To show that $\mathcal{I}(E(\psi)) \subseteq \mathcal{A}(\psi)$ we will verify the conditions from theorem 5.12. Given its last two coordinate functions, the injectivity of ψ is obvious. To prove the remaining conditions, namely the properness, the Whitney-regularity of the image, and the lower distance estimate for the derivative, we will show that they are directly inherited from the map $(x, y) \mapsto (x^3, y^3)$. Indeed, the curve defined by $\gamma(t) = t^3$ has a locally Hölder continuous inverse, namely $s \mapsto \sqrt[3]{s}$. Proposition 2.20 implies that γ generates a closed composition algebra. By proposition 5.18 the tensorized map defined by $\theta(x, y) = (x^3, y^3)$ is injective, proper, satisfies the lower distance estimate for derivatives, and possesses a Whitney-regular image. Theorem 4.11 then implies that $\mathcal{I}(E(\theta)) \subseteq \mathcal{A}(\theta)$ and, since $\mathcal{A}(\theta) \subseteq \mathcal{A}(\psi)$, also the required inclusion $\mathcal{I}(E(\psi)) = \mathcal{I}(E(\theta)) \subseteq \mathcal{A}(\psi)$.

(ii) This approach can also be used when dealing with maps that are not real-analytical. Indeed consider the map $\varphi : [0, \infty] \to \mathbb{R},\ t \mapsto \exp(-1/t)$ and define $\Psi : \mathbb{R}^2 \to \mathbb{R}^3$ by

$$\Psi(x) = \begin{pmatrix} \varphi(x^2 + y^2) \\ x^3 \\ y^3 \end{pmatrix}.$$

The critical set is also given by $E(\Psi) = \{(x, y) \in \mathbb{R}^2 : x \cdot y = 0\}$ and again $\mathcal{I}(E(\Psi)) \subseteq \mathcal{A}(\psi)$ holds. Surprisingly, its shape seems smoother around $(0, 0)$ then the real-analytic example (i) as the following picture shows.

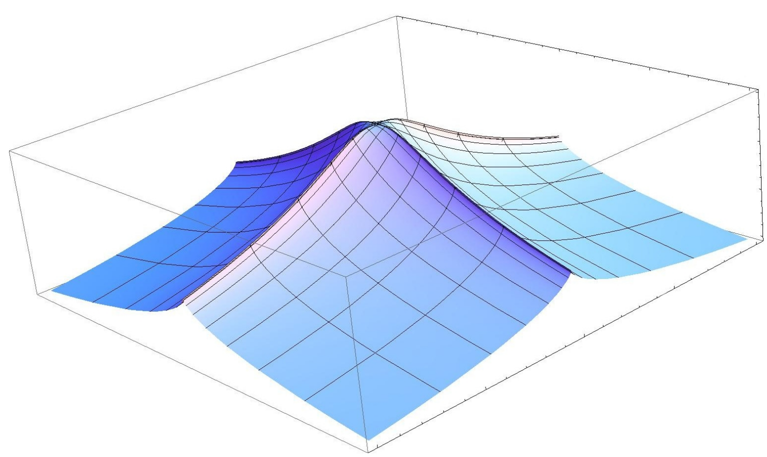


Figure 5.2: The image $\Psi([-1, 1]^2)$.

Let us first compute the derivative which is given by

$$\nabla\Psi(x,y) = \begin{bmatrix} 2 \cdot x \cdot \varphi'(x^2+y^2) & 2 \cdot y \cdot \varphi'(x^2+y^2) \\ 3 \cdot x^2 & 0 \\ 0 & 3 \cdot y^2 \end{bmatrix}.$$

We see that its rank is less than two whenever $x = 0$ or $y = 0$. If again θ is defined as in (i), the same argument as above implies $\mathcal{I}(E(\Psi)) = \mathcal{I}(E(\theta)) \subseteq \mathcal{A}(\theta) \subseteq \mathcal{A}(\Psi)$.
(iii) We can even give an example for a closed composition algebra. Let φ be defined as above and consider the map $\Phi : \mathbb{R}^2 \to \mathbb{R}^3$ defined by

$$\Phi(x) = \begin{pmatrix} \varphi(x^2) + \varphi(y^2) \\ x^3 \\ y^3 \end{pmatrix}.$$

We have again $E(\Phi) = \{(x,y) \in \mathbb{R}^2 : x \cdot y = 0\}$, and the algebra $\mathcal{A}(\Phi)$ is closed.
This time, there are no sharp edges along the critical set as the following picture shows.

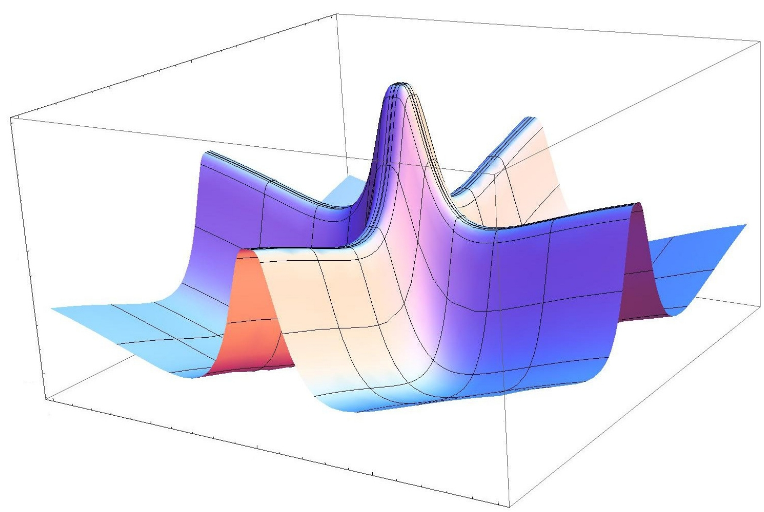


Figure 5.3: The image $\Phi([-1,1]^2)$.

Computing the derivatives we obtain

$$\nabla\Phi(x,y) = \begin{bmatrix} 2x \cdot \varphi'(x^2) & 2y \cdot \varphi'(y^2) \\ 3x^2 & 0 \\ 0 & 3y^2 \end{bmatrix}$$

and again $\Phi'(x,y)$ does not have full rank whenever $x = 0$ or $y = 0$. If $\theta : \mathbb{R}^2 \to \mathbb{R}^3$ denotes the map $(x,y) \mapsto (x^3, y^3)$ as in part (i) and (ii) above, we obtain that $\mathcal{A}(\theta)$ is closed. To prove that this is also the case for $\mathcal{A}(\Phi)$ it remains to show that the first coordinate Φ_1 is contained in $\mathcal{A}(\theta)$. Since $\mathcal{A}(\theta)$ is a subspace, it is sufficient to prove that both $(x, y) \mapsto \varphi(x^2)$ and $(x, y) \mapsto \varphi(y^2)$ are contained in $\mathcal{A}(\theta)$. This is due to the fact that both functions are flat on $E(\theta)$ and that $\mathcal{I}(E(\theta)) \subseteq \mathcal{A}(\theta)$. We obtain the inclusions

$$\mathcal{A}(\theta) \subseteq \mathcal{A}(\Phi) \subseteq \overline{\mathcal{A}(\theta)} = \mathcal{A}(\theta),$$

hence the required $\overline{\mathcal{A}(\Phi)}) = \mathcal{A}(\theta) = \mathcal{A}(\Phi)$.

Even though the first three conditions in corollary 5.14 look very similar to the ones required in the one-dimensional case of proposition 2.19, with the lower distance estimate taking the place of the finite order condition, we are not able to prove that the Hölder condition is sufficient to obtain a closed algebra. Of course, the local Hölder continuity of the inverse implies both the properness and the Whitney-regularity of the image, but we fail to see how it would imply the lower distance estimate.
Let us close this section with a few open problems.

Problem 1.
In the case of an injective $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$, is the Whitney-regularity of $\psi(\mathbb{R}^q)$ necessary in order for $\mathcal{A}(\psi)$ to be closed?

We believe that this is true. One might attack this problem by showing that every distribution u with compact support in $\psi(\mathbb{R}^q)$ can be estimated by seminorms $\|\cdot\|_{\psi(K),n}$ on $\psi(\mathbb{R}^q)$, which is the property we need in the proof of proposition 5.12. The next problem is much more difficult in nature as it requires to characterize the functions in the closure of $\mathcal{A}(\psi)$.

Problem 2.
For an injective $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$, are the three conditions of proposition 5.12 namely

(i) ψ is a proper map,

(ii) $\psi(\mathbb{R}^q)$ is a Whitney-regular set and

(iii) ψ satisfies the lower distance estimate.

sufficient to obtain that $\mathcal{A}(\psi)$ is closed?
More specifically, given proposition 5.14 it would be sufficient to show that, under these conditions, we can find for every $f \in \overline{\mathcal{A}(\psi)}$ a composition $F \circ \psi$ such that $f - F \circ \psi$ is flat on $E(\psi)$.

We recall that at least for injective generators ψ with a locally Hölder continuous inverse, the closure of $\mathcal{A}(\psi)$ is fully described by Tougeron ([Tou71], theorem (1.1)). We already pointed out at the beginning of chapter 1 that this theorem implies

$$\overline{\mathcal{A}(\psi)} = \{f \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}) : T_x^\infty f \in T_x^\infty \mathcal{A}(\psi) \ \forall x \in \mathbb{R}^q\}.$$

The heart of this problem is the behavior of $\mathcal{A}(\psi)$ on the critical set $E(\psi)$. One way to attack this might be to decompose the critical set into subsets $E_j(\psi)$ where the rank of $\psi'(x)$ is j in order to reduce the dimensions step by step. In this context, the decomposition of closed sets mentioned in definition 5.23 using regularly situated sets could be helpful.
Finally the similarities between the decomposition of the local Hölder continuity of the inverse and the three conditions of Whitney-regularity properness and finite order condition in the case of one variable suggest that this could be true even in higher dimensions.

Problem 3.
Is it true that an injective $\psi \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ generates a closed composition algebra if and only if ψ^{-1} is locally Hölder continuous?

Chapter 6

Appendix

6.1 Equivalent seminorms on $\mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$

First let us give a proof for the simple fact mentioned in chapter 5, namely that the family of seminorms $\|\!|\cdot|\!\|_{K,n}$ where $K \subseteq \mathbb{R}^q$ is a compact set and $n \in \mathbb{N}$, generates the usual Fréchet topology on $\mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$. For $F \in \mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$ we recall the definitions of both the usual seminorm

$$\|F\|_{K,n} = \sup\{\|D^\alpha F(x)\| : x \in K, |\alpha| \leq n\},$$

where $\alpha \in \mathbb{N}_0^q$ is a multi-index of length $|\alpha| = \alpha_1 + \cdots + \alpha_q$, and the multidirectional one

$$|\!|\!|F|\!|\!|_{K,n} = \sup\{|\!|\!|F^{(k)}(x)|\!|\!| : x \in K, k \leq n\}$$

from lemma 3.3, where $|\!|\!|T|\!|\!| = \sup\{\|T[r_1, ..., r_k]\|_Y : \|r_1\|_X \leq 1, ..., \|r_k\|_X \leq 1\}$ for a k-linear map $T \in \mathcal{M}_k(X, Y)$ between the normed spaces X and Y.
Let us now prove that this change of interpretation of the derivatives and their related seminorms is harmless in nature and leaves the topology of $\mathcal{E}(\mathbb{R}^d, \mathbb{R}^q)$ untouched.

Proposition 6.1.
Both families of seminorms $\|\cdot\|_{K,n}$ and $|\!|\!|\cdot|\!|\!|_{K,n}$ generate the same topology on $\mathcal{E}(\mathbb{R}^d, \mathbb{R}^q)$.

Proof. Fix a compact set $K \subseteq \mathbb{R}^d$ and differentiation order $n \in \mathbb{N}$. We will show that both seminorms $\|\cdot\|_{K,n}$ and $|\!|\!|\cdot|\!|\!|_{K,n}$ are equivalent. For any multi-index $\alpha \in \mathbb{N}_0^d$ we can use the notation from definition 3.1 to translate $\alpha \in \mathbb{N}_0^d$ into the multidirection

$$(\underbrace{e_1, ..., e_1}_{\alpha_1 \text{ times}}, ..., \underbrace{e_d, ..., e_d}_{\alpha_d \text{ times}}) = \mathop{\boxtimes}_{k=1}^{d} \left(\mathop{\boxtimes}_{j=1}^{\alpha_k} e_k \right)$$

where e_j denotes the j-th unit vector in $\mathbb{R}^d$. For $F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R}^q)$ we thereby obtain

$$D^\alpha F(x) = F^{(|\alpha|)}(x) \left[\mathop{\boxtimes}_{k=1}^{d} \left(\mathop{\boxtimes}_{j=1}^{\alpha_k} e_k \right) \right],$$

which implies $\|D^\alpha F(x)\| \leq \|\!|F^{(|\alpha|)}(x)|\!\|$ and hence $\|F\|_{K,n} \leq |\!|\!|F|\!|\!|_{K,n}$.
We will deduce the other estimate from the fact that every multilinear $T \in \mathcal{M}_n(\mathbb{R}^d, Y)$ satisfies the inequality

$$|\!|\!|T|\!|\!| \leq \sqrt{d^n} \cdot \sup\left\{\left\|T\left[\boxtimes_{j=1}^{n} \eta_j\right]\right\| : \eta_1, ..., \eta_n \in \{e_1, ..., e_d\}\right\}, \qquad (*)$$

which we will prove by induction.
For $n = 1$ and $\|x\| \leq 1$ the linearity of T implies

$$\|T[x]\| = \left\|T\left[\sum_{j=1}^{d} x_j e_j\right]\right\| \leq \sum_{j=1}^{d} |x_j| \cdot \sup\{\|T[e_\ell]\| : \ell \leq d\} \leq \sqrt{d} \cdot \sup\{\|T[e_\ell]\| : \ell \leq d\}.$$

For the induction step consider some $(n+1)$-linear map $T \in \mathcal{M}_{n+1}(\mathbb{R}^d, Y)$ and directions $x_1, ..., x_n \in \mathbb{R}^d$ satisfying $\|x_j\| \leq 1$. We can write every $\xi \in \mathbb{R}^d$ with $\|\xi\| \leq 1$ as the sum $\sum_{j=1}^d \xi_j \cdot e_j$, where $|\xi_j| \leq 1$. We obtain the estimate

$$\|T[x_1, ..., x_n, \xi]\| \leq \sum_{j=1}^{d} |\xi_j| \cdot \|T[x_1, ..., x_n, e_j]\| \leq \sum_{j=1}^{d} |\xi_j| \cdot \|T_j[x_1, ..., x_n]\|,$$

where T_j is defined by $T_j[x_1, ..., x_n] = T[x_1, ..., x_n, e_j]$. Applying the Cauchy-Schwarz inequality leads to

$$\|T[x_1, ..., x_n, \xi]\| \leq \|\xi\| \cdot \left(\sum_{j=1}^{d} \|T_j[x_1, ..., x_n]\|^2\right)^{\frac{1}{2}} \leq \left(\sum_{j=1}^{d} |\!|\!|T_j|\!|\!|^2\right)^{\frac{1}{2}},$$

hence $|\!|\!|T|\!|\!| \leq \sqrt{d} \cdot \max\{|\!|\!|T_j|\!|\!| : j \leq d\}$ once taking the supremum over all $\|x_1\|, ..., \|x_n\| \leq 1$.
We can now apply the induction assumption to every T_j to see that

$$\begin{aligned} |\!|\!|T_j|\!|\!| &\leq \sqrt{d^n} \cdot \sup\{\|T_j[\eta_1, .., \eta_n]\| : \eta_1, ..., \eta_n \in \{e_1, ..., e_d\}\} \\ &= \sqrt{d^n} \cdot \sup\{\|T[\eta_1, .., \eta_n, e_j]\| : \eta_1, ..., \eta_n \in \{e_1, ..., e_d\}\}. \end{aligned}$$

Inserting this in the estimate $|\!|\!|T|\!|\!| \leq \sqrt{d} \cdot \max\{|\!|\!|T_j|\!|\!| : j \leq d\}$ leads to the required inequality $|\!|\!|T|\!|\!| \leq \sqrt{d^{n+1}} \cdot \sup\{\|T[\eta_1, ..., \eta_{n+1}]\| : \eta_1, ..., \eta_{n+1} \in \{e_1, ..., e_d\}\}$.
Using $(*)$ for the n-linear map $F^{(n)}(x)$, we get

$$|\!|\!|F^{(n)}(x)|\!|\!| \leq \sqrt{d^n} \cdot \sup\{\|F^{(n)}[\xi_1, ..., \xi_{n+1}]\| : \xi_1, ..., \xi_{n+1} \in \{e_1, ..., e_d\}\}.$$

Finally, note that every n-tuple $(y_1, ..., y_n)$ of unit vectors $y_1, .., y_n \in \{e_1, ..., e_d\}$ adds up to some multi-index $\alpha = \sum_{j=1}^n y_j$ of length $|\alpha| = n$ that satisfies $D^\alpha F(x) = F^{(n)}(x)[y_1, ..., y_n]$.
We therefore obtain the estimate

$$\begin{aligned} |\!|\!|F^{(n)}(x)|\!|\!| &\leq \sqrt{d^n} \cdot \sup\{\|F^{(n)}[\xi_1, ..., \xi_{n+1}]\| : \xi_1, ..., \xi_{n+1} \in \{e_1, ..., e_d\}\} \\ &\leq \sqrt{d^n} \cdot \sup\{\|D^\alpha F(x)\| : \alpha \in \mathbb{N}_0^d : |\alpha| = n\} \end{aligned}$$

and hence the required $|\!|\!|F|\!|\!|_{K,n} \leq \sqrt{d^n} \cdot \|F\|_{K,n}$. $\square$

6.2 The theorem of Borel for hyperplanes

The usual version of this result, as mentioned by Trèves in Theorem 38.1 of his book [Trè67], states that every sequence $(c_\alpha)_{\alpha\in\mathbb{N}_0^d}$ can be expressed as the Taylor sequence $(D^\alpha F(0))_{\alpha\in\mathbb{N}_0^d}$ of derivatives of a smooth function $F\in\mathcal{E}(\mathbb{R}^d,\mathbb{R})$. We can extend this to allow a coefficient sequence $(f_\alpha)_{\alpha\in\mathbb{N}_0^d}$ of smooth functions on $\mathbb{R}^q$, thereby obtaining the mentioned Theorem 38.1 as the special case $q=0$ as required in the proofs of propositions 5.21 and 5.25.

Proposition 6.2.
The map $T:\mathcal{E}(\mathbb{R}^{d+q},\mathbb{R})\to\prod_{\alpha\in\mathbb{N}_0^d}\mathcal{E}(\mathbb{R}^q,\mathbb{R})$ *defined by*

$$T(F)=\left(F^{(\alpha,0_q)}(0_d,\cdot)\right)_{\alpha\in\mathbb{N}_0^d}$$

is surjective, where $G(0_d,\cdot)$ *stands for the map* $y\mapsto G(0_d,y)$. *This means that for any family* $(f_\alpha)_{\alpha\in\mathbb{N}_0^d}\subseteq\mathcal{E}(\mathbb{R}^q,\mathbb{R})$ *we can find* $F\in\mathcal{E}(\mathbb{R}^{d+q},\mathbb{R})$ *such that* $D^{(\alpha,0_q)}F(0_d,y)=f_\alpha(y)$.

Proof. First let us note that $\prod_{\alpha\in\mathbb{N}_0^d}\mathcal{E}(\mathbb{R}^q,\mathbb{R})$ together with the product topology of $\mathcal{E}(\mathbb{R}^q,\mathbb{R})$ is a Fréchet space and that T is continuous and linear. We will prove that the range of T is dense and closed.
For the proof of the density, fix a family $(f_\alpha)_{\alpha\in\mathbb{N}_0^d}$ in $\mathcal{E}(\mathbb{R}^q,\mathbb{R})$. For a given zero-neighborhood U in $\prod_{\alpha\in\mathbb{N}_0^d}\mathcal{E}(\mathbb{R}^q,\mathbb{R})$ we can find zero-neighborhoods $(V_\alpha)_{|\alpha|\le N}$ in $\mathcal{E}(\mathbb{R}^q,\mathbb{R})$ such that $\bigcap_{|\alpha|\le N}\pi_\alpha^{-1}(V_\alpha)\subseteq U$, where π_α denotes the usual projection onto $\mathcal{E}(\mathbb{R}^q,\mathbb{R})$. By defining

$$F(x,y)=\sum_{|\alpha|\le N}f_\alpha(y)\cdot x^\alpha$$

tone obtains $F^{(\alpha,0_q)}(0_d,y)=f_\alpha(y)$ and hence $\pi_\alpha(T(F)-f)=0$ for all $|\alpha|\le N$. This implies that $T(F)-f$ is contained in the intersection $\bigcap_{|\alpha|\le N}\pi_\alpha^{-1}(V_\alpha)\subseteq U$.
Let us now show that the range of T is closed, which we will prove by verifying the condition $\mathrm{Ker}(T)^\circ=\mathrm{Range}(T^t)$ of the closed range theorem (26.3 in [MV97]). The inclusion $\mathrm{Ker}(T)^\circ\supseteq\mathrm{Range}(T^t)$ is always true. To prove the other inclusion, consider some $u\in\mathrm{Ker}(T)^\circ$. For $F\in\mathcal{E}(\mathbb{R}^{d+q},\mathbb{R})$ with support outside $\{0_d\}\times\mathbb{R}^q$ all partial derivatives vanish on $\{0_d\}\times\mathbb{R}^q$, hence $T(F)=0$ and thus $u(F)=0$. This implies $\mathrm{supp}(u)\subseteq\{0_d\}\times\mathbb{R}^q$. Theorem 2.3.5 in Hörmander's book [Hör03] states that we can write

$$u(F)=\sum_{|\alpha|\le k}u_\alpha(F_\alpha)$$

where u_α is a distribution on $\mathcal{E}(\mathbb{R}^q,\mathbb{R})$ and $F_\alpha(y)=F^{(\alpha,0_q)}(0_d,y)$. Defining

$$v((f_\alpha)_{\alpha\in\mathbb{N}_0^d})=\sum_{|\alpha|\le k}u_\alpha(f_\alpha))$$

we obtain a continuous linear map u on $\prod_{\alpha\in\mathbb{N}_0^d}\mathcal{E}(\mathbb{R}^d,\mathbb{R})$ that satisfies $u(F)=v(T(F))$ and hence the required $u\in\mathrm{Range}(T^t)$. $\square$

6.3 Stable diffeomorphisms and norm estimates

Now we want to construct the specific diffeomorphisms required in the proof of propositions 5.10 and 5.12. There, we need to extend the global coordinate system $\psi : \mathbb{R}^q \setminus E(\psi) \to \mathbb{R}^d$ locally to a diffeomorphism $\Psi_x : U_x \times U_0 \to \mathbb{R}^d$ satisfying further stability conditions and estimates for the derivatives. Let us start with some basic notions of differential geometry.

Definition 6.3.
A subset M of $\mathbb{R}^d$ is called a q-dimensional manifold if for every point $m \in M$ there is an open subset U of $\mathbb{R}^d$ containing m, an open set $V = V_m \subseteq \mathbb{R}^d$, and a diffeomorphism $h = h_m : U \to V$ such that

$$h(U \cap M) = V \cap (\mathbb{R}^q \times \{0\}).$$

By theorem 5-2 in Spivak's book [Spi65] this is equivalent to the existence of a local "differentiable parametrization" $f : W \to U$, also called coordinate system, where $W \subseteq \mathbb{R}^q$ is an open set and $U \subseteq \mathbb{R}^d$ is open and contains $m \in M$, satisfying

1. $f(W) = M \cap U$,
2. $f'(y)$ has full rank on W,
3. $f^{-1} : f(W) \to W$ is continuous.

Through $\psi : \mathbb{R}^q \setminus E \to \mathbb{R}^d$ we have not only local but even global knowledge of the q-dimensional submanifold $\psi(\mathbb{R}^q \setminus E)$ in terms of a global coordinate system. Unfortunately, in order to construct the parametrization $F \circ \psi$ of some $f \in \mathcal{A}(\psi)$, we require additional knowledge of the map given in the definition above. More precisely we require accurate estimates for the norms $\|\!|h|\!\|_{K,n}$ and $\|\!|h^{-1}|\!\|_{K,n}$ of the diffeomorphisms stated in the definition above as well as a certain stability condition for the derivatives, namely that the inverse $\Phi = h^{-1}$ satisfies $\Phi'(y,0)[\{0\} \times \mathbb{R}^{d-q}] \perp \psi'(y)[\mathbb{R}^q]$. This means that Φ is an extension of the coordinate map to a diffeomorphism on $\mathbb{R}^d$ such that the derivative $\Phi'(y,0)$ maps the last $d-q$ unit vectors to a basis of the space of normal vectors at $\psi(y)$.
The construction will require two steps. In the first step, we will obtain a coordinate system satisfying the required norm estimates. In the second step, we will use this parametrization to generate another coordinate system that also satisfies the stability conditions for the derivatives.

Lemma 6.4 (First step, norm domination)**.**
Consider some q-dimensional manifold M in $\mathbb{R}^d$. For every coordinate system $\psi : U \to M$ and $x \in U$, we can find open neighborhoods U_x of $x \in \mathbb{R}^q$, U_0 of $0 \in \mathbb{R}^{d-q}$, and V_x of $\psi(x) \in \mathbb{R}^d$ as well as a diffeomorphism $\Phi = \Phi_x : U_x \times U_0 \to V_x$ such that:

(i) $\psi(y) = \Phi(y,0)$ for all $y \in U_x$.

(ii) *For every* $r \in \mathbb{R}^q$ *and* $s \in \mathbb{R}^{d-q}$ *we have* $\Phi'(x,0)[(r, 0_{d-q})] \perp \Phi'(x,0)[(0_q, s)]$.

(iii) $|||\Phi|||_{(x,0),n} = |||\psi|||_{x,n}$ *as well as* $|||(\Phi^{-1})'(\psi(x))||| \leq 2 \cdot |||\psi'(x)^{-1}|||$, *where* $\psi'(x)^{-1}$ *denotes the inverse of* $\psi'(x) : \mathbb{R}^q \to \mathrm{Range}(\psi'(x))$.

Proof. Prior to the proof, we give a sketch of our construction in the case of a very simple 2-dimensional submanifold of $\mathbb{R}^3$. The diffeomorphism Φ_x can be viewed as an extension of ψ describing the manifold. The blue hyperplane on the left visualizes the definition area $\mathbb{R}^q$ of ψ as a subspace of $\mathbb{R}^d$, whereas the blue structure on the right is the image of $\psi(\cdot) = \Phi_x(\cdot, 0)$. The derivative $\Phi'_x(x,0)$ then maps directions along the hyperplane on the left to tangential directions on the right (both blue) and the remaining orthogonal vector to the hyperplane to one that is orthogonal to the manifold (red). Note that the orthogonality only holds at $(x,0)$ and generally fails on the rest of $U_x \times U_0$.

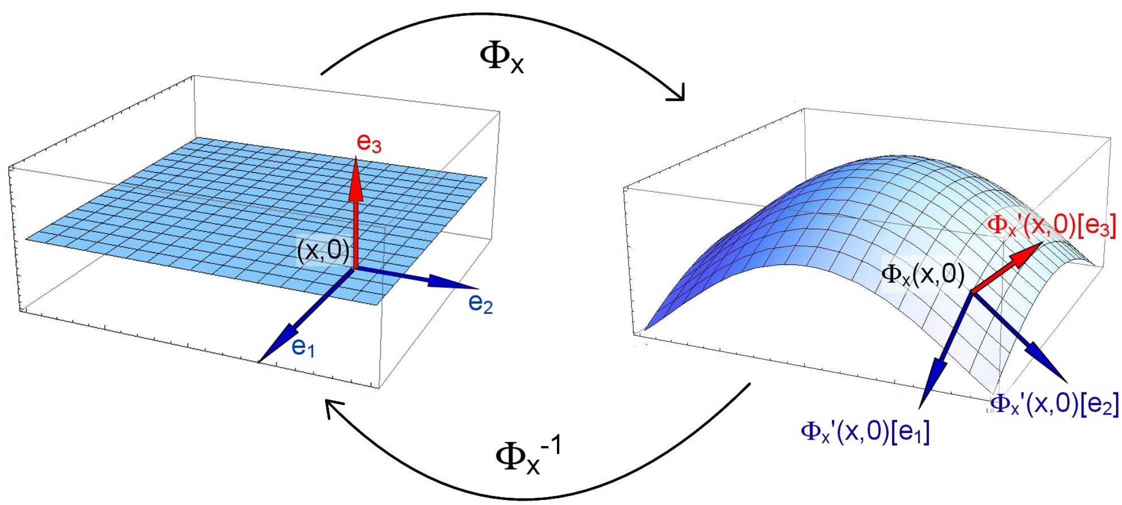


Figure 6.1: Φ_x mapping the hyperplane $\mathbb{R}^2 \times \{0\}$ to the manifold.

Let us now proceed with the detailed proof. Fix $x \in U$. Since ψ is a coordinate system, its derivative $\psi'(x)$ has full rank, hence $\mathrm{Range}(\psi'(x))$ is a q-dimensional subspace of $\mathbb{R}^d$. We can therefore find an orthogonal basis $\{\gamma_1, ..., \gamma_{d-q}\}$ of the $(d-q)$-dimensional subspace $\mathrm{Range}(\psi'(x))^\perp$. Without loss of generality we can even demand that $\|\gamma_j\| = |||\psi'(x)|||$. Let us now define the linear map $\Gamma : \mathbb{R}^{d-q} \to \mathbb{R}^d$ by

$$\Gamma(t) = \sum_{j=1}^{d-q} \gamma_j t_j$$

as well as

$$\Phi_x : U \times \mathbb{R}^{d-q} \to \mathbb{R}^d,\ (x,t) \mapsto \psi(x) + \Gamma(t).$$

Obviously, Φ_x is a smooth map on the open set $U \times \mathbb{R}^{d-q}$.
Computing its partial derivatives, we see that $\Phi'_x(x,0)[(r,0)] = \psi'(x)[r]$ for all $r \in \mathbb{R}^q$ as well as $\Phi'_x(x,0)[(0,s)] = \sum_{j=1}^{d-q} s_j \cdot \gamma_j$ for all $s \in \mathbb{R}^{d-q}$. We obtain that Φ'_x has full rank in

$(x, 0)$ and the inverse function theorem implies that it is a diffeomorphism onto its range once restricted to an adequate open neighborhood of $(x, 0)$. Without loss of generality we can suppose this neighborhood to have product structure $U_x \times U_0$, where $U_x \subseteq \mathbb{R}^q$ and $U_0 \subseteq \mathbb{R}^{d-q}$ are respective neighborhoods of $x \in \mathbb{R}^q$ and $0 \in \mathbb{R}^{d-q}$. The set $W = \Phi_x(U_x \times U_0)$ is open and we have $\Phi_x(U_x \times \{0\}) = \psi(U_x) = M \cap W$. Since Γ is a linear map and therefore $\Gamma'(t)[s] = \Gamma(s)$, we obtain the explicit representation

$$\Phi_x'(y,t)[(r,s)] = \psi'(y)[r] + \Gamma'(t)[s] = \psi'(y)[r] + \Gamma(s).$$

Using the orthogonality of $\psi'(x)[r]$ and $\Gamma(s)$, we can deduce the identity

$$\|\Phi_x'(x,0)[r,s]\|^2 = \|\psi'(x)[r]\|^2 + \|\Gamma(s)\|^2 = \|\psi'(x)\|^2 \cdot (\|r\|^2 + \|s\|^2) = \|\psi'(x)\|^2 \cdot \|(r,s)\|^2.$$

This implies $\|\Phi'(x,0)\| = \|\psi'(x)\|$ for the first derivative when taking the supremum over all $\|(r,s)\| \leq 1$.
For $k \geq 2$ the linearity of Γ implies $\Gamma^{(k)} = 0$. Applying formula 3.2 to compute the k-derivative of the composition $\psi \circ \pi$, where $\pi : \mathbb{R}^d \to \mathbb{R}^q$ denotes the projection $(y,t) \mapsto y$, we have

$$\begin{aligned}
\Phi_x^{(k)}(y,t)\left[\boxtimes_{j=1}^{k}(r_j,s_j)\right] &= (\psi \circ \pi)^{(k)}(y,t)\left[\boxtimes_{j=1}^{k}(r_j,s_j)\right] + \Gamma^{(k)}(y,t)\left[\boxtimes_{j=1}^{k}(r_j,s_j)\right] \\
&= \sum_{\mathcal{P} \in \mathfrak{P}(k)} \psi^{(|\mathcal{P}|)}(\pi(y,t))\left[\boxtimes_{P \in \mathcal{P}} \pi^{(|P|)}(y,t)\left[\boxtimes_{j \in P}(r_j,s_j)\right]\right] + 0.
\end{aligned}$$

Using the fact that π is also a linear map and hence $\pi^{(j)} = 0$ for all $j \geq 2$, every evaluation of the multilinear map $\psi^{(|\mathcal{P}|)}(x)$ in the sum on the right hand side vanishes whenever some $P \in \mathcal{P}$ satisfies $|P| \geq 2$. This means that all summands are zero with exception of the one belonging to $\mathcal{P} = \{\{1\}, ..., \{k\}\}$ which implies

$$\begin{aligned}
\Phi_x^{(k)}(y,t)\left[\boxtimes_{j=1}^{k}(r_j,s_j)\right] &= \psi^{(k)}(\pi(y,t))\left[\boxtimes_{j=1}^{k}\pi'(y,t)[r_j,s_j]\right] \\
&= \psi^{(k)}(y)\left[\boxtimes_{j=1}^{k} r_j\right].
\end{aligned}$$

Taking the supremum over all $\|r_j\| \leq 1$ we obtain $\|\Phi_x^{(k)}(y,t)\| \leq \|\psi^{(k)}(y)\|$. Combining it with the estimate for the first derivative leads to the required $\|\Phi_x\|_{(x,0),n} \leq \|\psi\|_{x,n}$.
We can also give an estimate for the inverse map. Since $\mathbb{R}^d$ can be represented as the direct sum $\mathrm{Range}(\psi'(x)) \oplus \mathrm{Range}(\psi'(x))^\perp$, we can find unique directions $r \in \mathbb{R}^q$ and $s \in \mathbb{R}^{d-q}$ such that $w = \psi'(x)[r] + \Gamma(s) = \Phi'(x,0)[r,s]$. The orthogonality of the two summands implies $\|w\|^2 = \|\psi'(x)[r]\|^2 + \|\Gamma(s)\|^2$ and thus we have both

$$\|s\| = \|\Gamma^{-1}(\Gamma(s))\| \leq \|\Gamma^{-1}\| \cdot \|\Gamma(s)\| \leq \|\Gamma^{-1}\| \cdot \|w\|$$

as well as

$$\|r\| = \|\psi'(x)^{-1}[\psi'(x)[r]]\| \leq \|\psi'(x)^{-1}\| \cdot \|w\|.$$

For $\|w\| \le 1$ we obtain

$$\|(\Phi^{-1})'(\psi(x))[w]\| = \|\Phi'(x,0)^{-1}[w]\| = \|(r,s)\| \le |||\psi'(x)^{-1}||| + |||\Gamma^{-1}|||.$$

hence $|||(\Phi^{-1})'(\psi(x))||| \le |||\psi'(x)^{-1}||| + |||\Gamma^{-1}|||$ by taking the supremum over all $\|w\| \le 1$. To estimate $|||\Gamma^{-1}|||$ let us recall that we have chosen $\{\gamma_j : j \le d-q\}$ to be an orthogonal system satisfying $\|\gamma_j\| = |||\psi'(x)|||$. This orthogonality implies

$$\|\Gamma[t]\|^2 = \left\| \sum_{j=1}^{d-q} t_j \gamma_j \right\|^2 = \sum_{j=1}^{d-q} |t_j|^2 \cdot |||\psi'(x)|||^2 = |||\psi'(x)|||^2 \cdot \|t\|^2,$$

which gives us $\|\Gamma^{-1}[\Gamma[t]]\| = \|t\| = \frac{\|\Gamma(t)\|}{|||\psi'(x)|||}$. We recall that $\Gamma^{-1} : \mathrm{Range}(\Gamma) \to \mathbb{R}^{d-q}$ and that its norm is defined by $|||\Gamma^{-1}||| = \sup\{\|\Gamma^{-1}(\Gamma(r))\| : \|\Gamma(r)\| \le 1\}$. We therefore obtain $|||\Gamma^{-1}||| = \frac{1}{|||\psi'(x)|||} \le |||\psi'(x)^{-1}|||$. Inserting this into the estimate above leads to the required

$$|||(\Phi^{-1})'(\psi(x))||| \le 2 \cdot |||\psi'(x)^{-1}|||.$$

□

With this intermediate construction, we can modify the diffeomorphism Φ_x so that $\Phi_x'(y,0)$ maps the first q unit vectors to a basis of the tangent space $\mathrm{Range}(\psi'(y))$ at $\psi(y)$ and the last remaining $d-q$ unit vectors to a basis of the space $\mathrm{Range}(\psi'(y))^\perp$ of normal directions at $\psi(y)$. We will call this property orthogonally stable. Note that this is an improvement of condition (ii) in the previous lemma as we are not only restricted to the orthogonality of the partial derivatives at $(x,0)$ but can extend this to all $(y,0) \in U_x \times \{0\}$. As visualization, let us consider the same submanifold as in figure 6.1. The hyperplane is still mapped to the manifold but this time the orthogonal stability holds in every point of the manifold. To better visualize this, we have sketched several orthogonal vectors to the hyperplane that are mapped to vectors orthogonal to the manifold in their respective images.

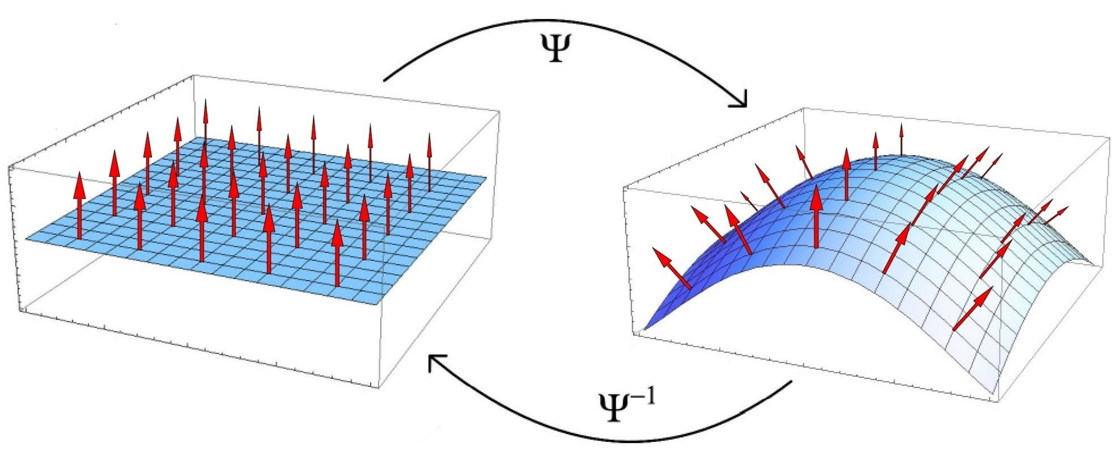


Figure 6.2: The red arrows are orthogonal to the hyperplane (left) and manifold (right).

Proposition 6.5 (Second step, orthogonal stability)**.**
Consider some q-dimensional manifold M in $\mathbb{R}^d$. For every coordinate system $\psi : U \to M$ and $x \in U$ we can find open neighborhoods U_x of $x \in \mathbb{R}^q$, U_0 of $0 \in \mathbb{R}^{d-q}$ and V_x of $\psi(x) \in \mathbb{R}^d$ as well as a diffeomorphism $\Psi = \Psi_x : U_x \times U_0 \to V_x$ such that:

(i) $\psi(y) = \Psi(y, 0)$ *for all* $y \in U_x$.

(ii) $\Psi'(y, 0)$ *is a bijection between* $\mathbb{R}^q \times \{0_{d-q}\}$ *and* $\psi'(y)[\mathbb{R}^d]$ *for all* $y \in U_x$.

(iii) $\Psi'(y, t)$ *is a bijection between* $\{0_q\} \times \mathbb{R}^{d-q}$ *and* $\psi'(y)[\mathbb{R}^d]^\perp$ *for all* $(y, t) \in U_x \times U_0$.

(iv) $\Psi''(y, t)[s_1, s_2] = 0$ *whenever* $s_1, s_2 \in \{0\} \times \mathbb{R}^{d-q}$ *and* $(y, t) \in U_x \times U_0$.

(v) There are constants $C_n, \alpha_n > 0$ *depending only on* n *such that*

$$|||\Psi|||_{(x,0),n} \leq C_n \cdot (1 + |||\psi|||_{x,n})^{\alpha_n} \cdot (1 + |||\psi'(x)^{-1}|||)^{\alpha_n}.$$

(vi) There are constants $c_n, \beta_n > 0$ *depending only on* n *such that*

$$|||\Psi^{-1}|||_{\psi(x),n} \leq c_n \cdot (1 + |||\psi|||_{x,n})^{\beta_n} \cdot (1 + |||\psi'(x)^{-1}|||)^{\beta_n}.$$

Proof. For $x \in U$, fix U_x, U_0, V_x and $\Phi_x : U_x \times U_0 \to V_x$ as given by the previous lemma 6.4. We define $\Psi = \Psi_x : U_x \times U_0 \to \mathbb{R}^d$ by

$$\Psi_x(y, t) = \psi(y) + \left(\Phi_x'(y, 0)^{-1}\right)^* [(0, t)],$$

where T^* denotes the adjoint of a linear map T with respect to the usual scalar product on $\mathbb{R}^d$.

The map Ψ is smooth.
The function $(y, t) \mapsto \psi(y)$ is obviously smooth and it remains to prove that this is also the case for $(y, t) \mapsto (\Phi'(y, 0)^{-1})^* [(0, t)]$. Since Φ is a diffeomorphism, the function defined by $(y, t) \mapsto \Phi'(y, t)^{-1} = (\Phi^{-1})' (\Phi(y, t))$ is smooth with values in the finite-dimensional Banach space $X = L(\mathbb{R}^d, \mathbb{R}^d)$ of linear maps on $\mathbb{R}^d$. The map $T \mapsto T^*$ is also smooth as a linear map on X and we obtain that $(y, t) \mapsto (\Phi'(y, 0)^{-1})^*$ is also smooth with values in X. The evaluation map $z \mapsto \delta_z$, where $\delta_z : X \to \mathbb{R}^d,\ T \mapsto T(z)$ is linear and hence also smooth with values in $L(X, \mathbb{R}^d)$. As a composition of smooth maps, $(y, t) \mapsto (\Phi'(y, 0)^{-1})^* [(0, t)]$ must be smooth itself.

The map Ψ satisfies the conditions (i) through (iv)
Let us first compute the partial derivatives to show the tangential stability conditions (ii) and (iii). Fix a direction $(r, s) \in \mathbb{R}^q \times \mathbb{R}^{d-q}$. Consider the linear map $\tau : \mathbb{R}^q \to \mathbb{R}^d$ defined by $\tau(y) = (y, 0)$. We obtain $\psi = \Psi \circ \tau$ and applying the chain rule leads to

$$\psi'(y)[r] = (\Psi \circ \tau)'(y)[r] = \Psi'(\tau(y))[\tau'(y)[r]] = \Psi'(y, 0)[(r, 0)],$$

where we have deduced $\tau'(y)[r] = \tau(r) = (r, 0)$ from the fact that τ is linear. Since $\psi'(y)$ is injective for all $y \in U_x$, the subspace $\Psi'(y, 0)[\mathbb{R}^q \times \{0\}] = \psi'(y)[\mathbb{R}^q]$ has dimension q.

Therefore $\Psi'(y,0)$ is a linear bijection between $\mathbb{R}^q \times \{0\}$ and $\mathrm{Range}(\psi'(y))$ which proves (ii). To show (iii) we use the identity

$$\begin{aligned}\Psi'(y,t)[(0,s)] &= \lim_{0\neq\varepsilon\to 0} \frac{\Psi(y,t+\varepsilon\cdot s)-\Psi(y,t)}{\varepsilon}\\ &= \lim_{0\neq\varepsilon\to 0} \frac{\left(\Phi'(y,0)^{-1}\right)^*[(0,\varepsilon\cdot s)]}{\varepsilon} = \left(\Phi'(y,0)^{-1}\right)^*[(0,s)].\end{aligned}$$

Since Φ is a diffeomorphism, the linear map $\Phi'(y,0)^{-1}$ and hence also its adjoint are invertible and $\Psi'(y,t)[\{0_q\}\times\mathbb{R}^{d-q}]$ has dimension $d-q$. It now suffices to show that $\Psi'(y,t)[(0,s)] \perp \Psi'(y,0)[(r,0)]$ for all $r\in\mathbb{R}^q$ and $s\in\mathbb{R}^{d-q}$. This relation is given by the fact that $\Psi'(y,0)[r,0]=\psi'(y)[r]=\Phi'(y,0)[r,0]$ and the computation

$$\begin{aligned}\langle\Psi'(y,t)[(0,s)],\Psi'(y,0)[(r,0)]\rangle &= \left\langle\left(\Phi'(y,0)^{-1}\right)^*[(0,s)]\ ,\ \Phi'(y,0)[(r,0)]\right\rangle\\ &= \left\langle(0,s)\ ,\ \Phi'(y,0)^{-1}\left[\Phi'(y,0)[(r,0)]\right]\right\rangle\\ &= \langle(0,s),(r,0)\rangle = 0,\end{aligned}$$

where we have used the defining property $\langle T^*x,y\rangle = \langle x,Ty\rangle$ of the adjoint map. This implies that $\Psi'(y,t)$ is a bijection between $\{0\}\times\mathbb{R}^{d-q}$ and $\mathrm{Range}(\psi'(y))^\perp$, hence (iii).
We also obtain that Ψ' has full rank on $U_x\times\{0\}$ and, by shrinking U_x, U_0 and V_x, the inverse function theorem states that Ψ is a diffeomorphism, hence bijective. As a result we have $\Psi(U_x\times\{0\})=\psi(U_x)$ and (i).
To prove (iv), we only need to take a closer look at our computation of the derivative. For $s\in\mathbb{R}^{d-q}$, the directional derivative $\Psi'(y,t)[(0,s)]=\Psi'(y,0)[(0,s)]$ does not depend on t. Therefore the second order directional derivative $\Psi''(y,t)[(0,s)\boxtimes(0,w)]$ vanishes on $U_x\times U_0$ and so must every multidirectional derivative $\Psi^{(k)}(y,t)[r_1,...,r_k]=0$ whenever two different components r_ℓ and r_j are contained in the subspace $\{0^q\}\times\mathbb{R}^{d-q}$. We will now proceed and estimate the norms of Ψ and Ψ^{-1}.
The estimate (v) for $|||\Psi|||_{(x,0),n}$ holds.
We recall that by definition we have

$$\Psi(y,t)=\psi(y)+\left(\Phi'(y,0)^{-1}\right)^*[(0,t)].$$

The first summand $\Theta:\mathbb{R}^q\times\mathbb{R}^{d-q}\to\mathbb{R}^d, (y,t)\mapsto\psi(y)$ can be written as $\Theta=\psi\circ\pi$ where π denotes the projection $(y,t)\mapsto y$ onto the first q coordinates. Let us show that this implies the inequality $|||\Theta|||_{(x,0),n}\leq(1+|||\psi|||_{x,n+1})\cdot(1+|||\psi'(x)^{-1}|||)$. Using the fact that π is a linear map and therefore $\pi'(y,t)=\pi$ and $\pi^{(k)}=0$ for all $k\geq 2$, we can compute the directional derivatives of Θ, namely

$$\Theta^{(k)}(y,t)\left[\mathop{\boxtimes}_{j=1}^{k}(r_j,s_j)\right]=(\psi\circ\pi)^{(k)}(y,t)\left[\mathop{\boxtimes}_{j=1}^{k}(r_j,s_j)\right]=\psi^{(k)}(\pi(y,t))\left[\mathop{\boxtimes}_{j=1}^{k}\pi(r_j,s_j)\right].$$

This obviously implies

$$|||\Theta|||_{(x,0),n}\leq|||\psi|||_{x,n}\leq(1+|||\psi|||_{x,n})^{\alpha_n}\cdot(1+|||\psi(x)^{-1}|||)^{\alpha_n}$$

for every $\alpha_n \in \mathbb{N}$ by taking the supremum over all $\|(r_j, s_j)\| \leq 1$.
It remains to show that the required estimate also holds for the map

$$\varphi : U_x \times \mathbb{R}^{d-q}, (y,t) \mapsto (\Phi'(y,0)^{-1})^*[0,t].$$

To do this we will decompose $\varphi(y,t) = M(t) \circ F(y)$, where $M(t) : L(\mathbb{R}^d, \mathbb{R}^d) \to \mathbb{R}^d$ is defined by $G \mapsto G^*[(0,t)]$ and $F(y) = \Phi'(y,0)^{-1}$, to show that for every $\alpha \in \mathbb{N}_0^d$ and $\beta \in \mathbb{N}_0^{d-q}$ the partial derivative satisfies

$$D^{(\alpha,\beta)}\varphi(y,t) = (D^\beta M(t)) \circ (D^\alpha F(y)).$$

For all $t \in \mathbb{R}^{d-q}$, the map $M(t)$ is linear and continuous on $L(\mathbb{R}^d, \mathbb{R}^d)$ with respect to the operator norm. For all $j \leq q$, computing the respective partial derivative yields

$$D^{e_j}\varphi(y,t) = M(t) \circ D^{e_j}F(y),$$

which, by induction, leads to $D^{(\alpha,0)}(M(t) \circ F(y)) = M(t) \circ (D^\alpha F(y))$ for all $\alpha \in \mathbb{N}_0^q$. Furthermore the linearity of $t \mapsto M(t)$ also implies that the partial derivatives $D^{(0,\beta)}$ of the map $(y,t) \mapsto M(t) \circ L(y)$ satisfy $D^{(0,\beta)}\,(M(t) \circ L(y)) = D^\beta M(t) \circ L(y)$. This leads to the required

$$D^{(\alpha,\beta)}\varphi(y,t) = D^{(0,\beta)}D^{(\alpha,0)}M(t)\circ F(y) = D^{(0,\beta)}(M(t)\circ(D^\alpha F(y))) = (D^\beta M(t))\circ(D^\alpha F(y)).$$

The linearity of M also implies that every partial derivative $D^\beta M(t)$ vanish when the length of β exceeds one. Moreover we have $\|\!|M(t)|\!\| \leq \|t\|$ and since M is linear also $D^{e_j}M(t) = M(e_j) \leq 1$. For all $\|t\| \leq 1$ this implies $\|D^\beta M(t)\| \leq 1$, hence the estimate

$$\|D^{(\alpha,\beta)}\varphi(y,t)\| \leq \|\!|D^\beta M(t)|\!\| \cdot \|\!|D^\alpha F(y)|\!\| \leq \|\!|D^\alpha F(y)|\!\| \leq \|F\|_{y,n},$$

where $\|F\|_{y,n} = \sup\{\|\!|D^\alpha F(y)|\!\| : |\alpha| \leq n\}$. The equivalence of the norms proved in proposition 6.1 states that $\|\!|\varphi|\!\|_{(y,0),n} \leq \sqrt{d^n} \cdot \|\varphi\|_{(y,0),n} \leq \sqrt{d^n} \cdot \|F\|_{y,n}$. We can use the identity $F(y) \circ \Phi'(y,0) = Id$ to compute the partial derivatives of F. It follows from lemma 6.6 (i) that

$$0 = D^{e_j}(F(y) \circ \Phi(y,0)) = D^{e_j}F(y) \circ \Phi'(y,0) + F(y) \circ D^{e_j}\Phi'(y,0)$$

hence $D^{e_j}F(y) = -F(y) \circ D^{e_j}\Phi'(y,0) \circ F(y)$ and lemma 6.6 (ii) implies

$$\|D^{e_j}F\|_{y,n} \leq 4^n \cdot \|F\|_{y,n}^2 \cdot \|\!|D^{e_j}\Phi(\cdot,0)|\!\|_{y,n}.$$

Dominating $\|\!|D^{e_j}\Phi(\cdot,0)|\!\|_{y,n} \leq \|\!|\Phi|\!\|_{(y,0),n+1}$ and taking the supremum over all unit vectors $e_j \in \mathbb{R}^q$, we obtain the fundamental estimate

$$\|F\|_{y,n+1} \leq 4^n \cdot (1 + \|F\|_{y,n})^2 \cdot (1 + \|\!|\Phi|\!\|_{(y,0),n+1}). \qquad (*)$$

This will imply the required $\|F\|_{y,n} \leq C_n \cdot (1+\|\!|\psi|\!\|_{x,n+1})^{\alpha_n} \cdot (1+\|\!|\psi'(x)^{-1}|\!\|)^{\alpha_n}$ by induction.

For $n = 0$ we can use the estimates $\|F\|_{x,0} = \|\!|F(x)|\!\| = \|\!|\Phi'(x,0)^{-1}|\!\| \leq 2 \cdot \|\!|\psi'(x)^{-1}|\!\|$ as well as $\|\!|\Phi|\!\|_{(x,0),n+1} \leq \|\!|\psi|\!\|_{x,n+1}$ from the previous proposition 6.4 (iii) and the equivalence of the norms $\|\!| \cdot |\!\|_{K,n}$ and $\| \cdot \|_{K,n}$ to see that

$$\begin{aligned}
\|F\|_{x,1} &\leq (1 + \|F\|_{x,0})^2 \cdot (1 + \|\!|\Phi|\!\|_{(y,0),1}) \\
&\leq (1 + 2\|\!|\psi'(x)^{-1}|\!\|)^2 \cdot (1 + \|\!|\psi|\!\|_{x,1}) \\
&\leq 2 \cdot (1 + \|\!|\psi'(x)^{-1}|\!\|)^2 \cdot (1 + \|\!|\psi|\!\|_{x,1})^2.
\end{aligned}$$

To prove the induction step, suppose $\|F\|_{x,n} \leq C_n \cdot (1 + \|\!|\psi'(x)^{-1}|\!\|)^{\alpha_n} \cdot (1 + \|\!|\psi|\!\|_{x,1})^{\alpha_n}$ to be true. Using the fundamental estimate $(*)$ we obtain

$$\begin{aligned}
\|F\|_{x,n+1} &\leq 4^n \cdot (1 + \|F\|_{x,n})^2 \cdot (1 + \|\!|\Phi|\!\|_{(x,0),n+1}) \\
&\leq 4^n \cdot \Big(1 + C_n \cdot (1 + \|\!|\psi'(x)^{-1}|\!\|)^{\alpha_n} \cdot (1 + \|\!|\psi|\!\|_{x,n+1})^{\alpha_n}\Big)^2 \cdot (1 + \|\!|\psi|\!\|_{x,n+1}) \\
&\leq 4^n \cdot (1 + C_n)^2 \cdot (1 + \|\!|\psi'(x)^{-1}|\!\|)^{2\alpha_n} \cdot (1 + \|\!|\psi|\!\|_{x,n+1})^{2\alpha_n+1} \\
&\leq C_{n+1} \cdot (1 + \|\!|\psi'(x)^{-1}|\!\|)^{\alpha_{n+1}} (1 + \|\!|\psi|\!\|_{x,n+1})^{\alpha_{n+1}},
\end{aligned}$$

where $\alpha_{n+1} = 2\alpha_n + 1$ and $C_{n+1} = 4^n \cdot (1 + C_n)^2$.
Let us now proceed with the estimate for the norm of the inverse.
By corollary 3.5 we have that

$$\|\!| \left(\Psi_x^{-1}\right)^{(n)} (\psi(x))|\!\| \leq C_n \cdot (1 + \|\!|\Psi_x|\!\|_{x,n})^{\gamma(n-1)} \cdot (1 + \|\!|\Psi_x'(x,0)^{-1}|\!\|)^{\gamma(n)},$$

where $\gamma(j) = \frac{j(j+1)}{2}$. The monotonicity of the terms on the right hand side implies that the estimate stays true if we replace $\|\!| \left(\Psi_x^{-1}\right)^{(n)} (\psi(x))|\!\|$ by $\|\!|\Psi_x^{-1}|\!\|_{\psi(x),n}$ and $\gamma(n-1)$ by $\gamma(n)$. We then obtain

$$\|\!|\Psi_x^{-1}|\!\|_{\psi(x),n} \leq C_n \cdot (1 + \|\!|\Psi_x|\!\|_{x,n})^{\gamma(n)} \cdot (1 + \|\!|\Psi_x'(x,0)^{-1}|\!\|)^{\gamma(n)}.$$

We can estimate $1 + \|\!|\Psi_x|\!\|_{x,n}$ using property (v) to obtain

$$\begin{aligned}
1 + \|\!|\Psi_x|\!\|_{x,n} &\leq 1 + \tilde{C}_n \cdot (1 + \|\!|\psi|\!\|_{x,n})^{\alpha_n} \cdot (1 + \|\!|\psi'(x)^{-1}|\!\|)^{\alpha_n} \\
&\leq 2 \cdot \tilde{C}_n \cdot (1 + \|\!|\psi|\!\|_{x,n})^{\alpha_n} \cdot (1 + \|\!|\psi'(x)^{-1}|\!\|)^{\alpha_n}.
\end{aligned}$$

It remains to show that $\|\!|\Psi_x'(x,0)^{-1}|\!\| \leq \|\!|\psi'(x,0)|\!\|^{-1}$. To this end we recall our remark 3.7 stating the relation

$$\lambda(T) = \inf\{\|T(r)\| : \|r\| = 1\} = \frac{1}{\|\!|T^{-1}|\!\|}$$

between T and the norm of its inverse T^{-1}.
By computing the partial derivatives we have already seen that

$$\Psi_x'(x,0)[(r,t)] = \psi'(x)[r] + (\Phi_x'(x,0)^{-1})^*[(0,t)]$$

and we can use the orthogonality of both components to further estimate

$$\begin{aligned}\|\Psi'(x,0)[(r,t)]\|^2 &= \|\psi'(x)[r]\|^2 + \|(\Phi'_x(x,0)^{-1})^*[(0,t)]\|^2 \\ &\geq \lambda(\psi'(x))^2 \cdot \|r\|^2 + \lambda((\Phi'_x(x,0)^{-1})^*)^2 \cdot \|t\|^2 \\ &\geq \min\left\{\lambda(\psi'(x))^2, \lambda((\Phi'_x(x,0)^{-1})^*)^2\right\} \cdot (\|r\|^2 + \|t\|^2).\end{aligned}$$

Taking the infinum over all $\|(r,t)\| = 1$ we get $\lambda(\Psi'(x,0)) \geq \min\{\lambda(\psi'(x)), \lambda((\Phi'_x(x,0)^{-1})^*)\}$. Applying the identity from remark 3.7 to $T = (\Phi'_x(x,0)^{-1})^*$ we obtain

$$\lambda((\Phi'_x(x,0)^{-1})^*) = \lambda((\Phi'_x(x,0)^*)^{-1}) = \frac{1}{|||\Phi'_x(x,0)^*|||} = \frac{1}{|||\Phi'_x(x,0)|||}.$$

Therefore we have

$$\begin{aligned}|||\Psi'(x,0)^{-1}||| = \frac{1}{\lambda(\Psi'(x,0))} &= \max\left\{\frac{1}{\lambda(\psi'(x))}, |||\Phi'_x(x,0)|||\right\} \\ &= \max\left\{|||\psi'(x)^{-1}|||, |||\Phi'_x(x,0)|||\right\} = \max\left\{|||\psi'(x)^{-1}|||, |||\psi'(x)|||\right\} \\ &\leq (1 + |||\psi'(x)^{-1}|||) \cdot (1 + |||\psi'(x)|||).\end{aligned}$$

Inserting these estimates we finally get (vi) namely

$$|||\Psi_x^{-1}|||_{\psi(x),n} \leq c_n \cdot (1 + |||\psi'(x)^{-1}|||)^{\beta_n} \cdot (1 + |||\psi'(x)|||_{x,n})^{\beta_n},$$

where $\beta_n = \gamma(n) \cdot (1 + \alpha_n)$.

□

Lemma 6.6.
We consider the ring $(Y, +, *)$ of smooth maps from $\mathbb{R}^n$ to $L(\mathbb{R}^d, \mathbb{R}^d)$ together with the pointwise addition $(G+H)(z) = G(z)+H(z)$ and multiplication $(G*H)(z) = G(z) \circ H(z)$. For two smooth maps $G, H : \mathbb{R}^n \to L(\mathbb{R}^d, \mathbb{R}^d)$ and $k \in \mathbb{N}$ the following two statements are true:

(i) The higher order derivatives of $G * H$ in the multidirection $\mathfrak{r} = (r_1, ..., r_k)$ are given by

$$(G * H)^{(k)}(x)[r_1, ..., r_k] = \sum_{A \subseteq \{1,...,k\}} G^{(|A|)}(x)[\mathfrak{r}_A] \circ H^{(|A^c|)}(x)[\mathfrak{r}_{A^c}].$$

(ii) The following estimate holds:

$$|||G * H|||_{x,n} \leq 2^n \cdot |||G|||_{x,n} \cdot |||H|||_{x,n}.$$

Proof. We can view the first statement as a version of the Leibniz formula for multidirections instead of the usual partial derivatives. We will proceed by induction. For $k = 1$

computing the difference quotients gives us

$$\begin{aligned}(G*H)'(x)[r] &= \lim_{0\neq t\to 0}\frac{G(x+tr)\circ H(x+tr)-G(x)\circ H(x)}{t}\\ &= \lim_{0\neq t\to 0}\frac{G(x+tr)\circ(H(x+tr)-H(x))}{t}+\frac{(G(x+tr)-G(x))\circ H(x)}{t}\\ &= G(x)\circ H'(x)[r]+G'(x)[r]\circ H(x).\end{aligned}$$

To simplify the computation of the induction step we use the notation $\partial_r G(x) = G'(x)[r]$ as well as $\partial^k_{\mathfrak{r}} G(x) = G^{(k)}(x)\,[\mathfrak{r}] = G^{(k)}(x)[r_1, ..., r_k]$ for the higher order directional derivatives in the multidirection $\mathfrak{r} = [r_1, ..., r_k]$. This way we obtain

$$\begin{aligned}\partial^k_{\mathfrak{r}}(G*H) &= \partial_{r_k}\partial^{k-1}_{r_1,\dots,r_{k-1}}(G*H) = \partial_{r_k}\sum_{A\subseteq\{1,\dots,k-1\}}\partial^{|A|}_{\mathfrak{r}_A}G*\partial^{|A^c|}_{\mathfrak{r}_{A^c}}H\\ &= \sum_{A\subseteq\{1,\dots,k-1\}}\big(\partial_{r_k}\partial^{|A|}_{\mathfrak{r}_A}G\big)*\partial^{|A^c|}_{\mathfrak{r}_{A^c}}H+\partial^{|A|}_{\mathfrak{r}_A}G*\big(\partial_{r_k}\partial^{|A^c|}_{\mathfrak{r}_{A^c}}H\big)\\ &= \sum_{\substack{A\subseteq\{1,\dots,k\}\\ k\in A}}\partial^{|A|}_{\mathfrak{r}_A}G*\partial^{|A^c|}_{\mathfrak{r}_{A^c}}H+\sum_{\substack{A\subseteq\{1,\dots,k\}\\ k\in A^c}}\partial^{|A|}_{\mathfrak{r}_A}G*\partial^{|A^c|}_{\mathfrak{r}_{A^c}}H\\ &= \sum_{A\subseteq\{1,\dots,k\}}\partial^{|A|}_{\mathfrak{r}_A}G*\partial^{|A^c|}_{\mathfrak{r}_{A^c}}H.\end{aligned}$$

For the second part, let us recall the definition

$$\begin{aligned}|||G|||_{x,n} &= \sup\{|||G^{(k)}(x)[r_1,\dots,r_k]||| : k\le n,\ \|r_1\|,\dots,\|r_k\|\le 1\}\\ &= \sup\{|||\partial^k_{r_1,\dots,r_k}G(x)||| : k\le n,\ \|r_1\|,\dots,\|r_k\|\le 1\},\end{aligned}$$

where $|||\cdot|||$ simply denotes the operator norm on $L(\mathbb{R}^d, \mathbb{R}^d)$. For some multidirection $\mathfrak{r} = (r_1, ..., r_k)$, where all components satisfy $\|r_j\| \le 1$, we can use the formula given by (i) and the submultiplicativity $|||S * T||| \le |||S||| \cdot |||T|||$ to obtain

$$|||\partial^k_{\mathfrak{r}}(G*H)(x)||| \le \sum_{A\subseteq\{1,\dots,k\}}|||\partial^{|A|}_{\mathfrak{r}_A}G|||\cdot|||\partial^{|A^c|}_{\mathfrak{r}_{A^c}}H||| \le 2^k\cdot|||G|||_{x,k}\cdot|||H|||_{x,k}.$$

The estimate for $|||(G * H)^{(k)}(x)|||$ follows by taking the supremum over all multidirections $\mathfrak{r}$ in the product of the unit balls. The monotonicity with respect to k of the terms on the right hand side implies $|||G * H|||_{x,n} = \max\{|||G * H|||_{x,k} : k \le n\} \le 2^n \cdot |||G|||_{x,n} \cdot |||H|||_{x,n}$. $\square$

List of Symbols

$\lvert\lvert\lvert \cdot \rvert\rvert\rvert$	$\sup\{\lVert T[r_1, ..., r_n]\rVert : \lVert r_j\rVert \leq 1\}$	11
$\lvert\lvert\lvert \cdot \rvert\rvert\rvert_{K,n}$	$\sup\{\lvert\lvert\lvert f^{(k)}(x)\rvert\rvert\rvert : x \in K, k \leq n\}$	84
$\boxtimes$	Juxtaposition of multidirections	56
$\mathcal{A}(\psi)$	$\{F \circ \psi : F \in \mathcal{E}(\mathbb{R}^d, \mathbb{R})\}$	11
C_ψ	Composition operator $F \mapsto F \circ \psi$	11
$E(\psi)$	Critical set of ψ	26
$\mathcal{E}(\mathbb{R}^q, \mathbb{R}^d)$	Smooth functions from $\mathbb{R}^q$ to $\mathbb{R}^d$	11
$\mathcal{E}(M, \mathbb{R}^q)$	Smooth germs on M	85
$\mathcal{E}(\psi)$	$\mathcal{E}(\psi(\mathbb{R}^q))$	85
$\mathcal{E}_0(\psi)$	$\mathcal{R}(\psi)/\mathcal{I}(\psi)$	86
$\mathcal{I}(M, \mathbb{R}^d)$	Flat functions on M	85
$\mathcal{I}(E)$	$\mathcal{I}(E(\psi), \mathbb{R})$	85
$\mathcal{I}(\psi)$	$\mathcal{I}(\psi(\mathbb{R}^q), \mathbb{R})$	85
$\lambda(T)$	$\inf\{\lVert T[r]\rVert : \lVert r\rVert = 1\}$	64
$\mathcal{R}(\psi)$	Orthogonally flat functions with respect to ψ	86
$\mathfrak{P}(n)$	Set of partitions of $\{1, ..., n\}$	55
$\mathcal{P}$	Partition	55
$\mathfrak{r}, \mathfrak{s}$	Multidirections	55
T_x^∞	Taylor map	20
X_d	Space of formal power series in $\mathbb{R}^d$	20

Index

Composite function property, 35
Composition
  property, 36
  of formal power series, 22
Coordinate system, 118
Critical point
  for a function of one variable, 26
  for a function of several variables, 68
  of finite order, 37
Critical set
  for a function of one variable, 26
  for a function of several variables, 68

Diffeomorphism, 118
  norm dominated, 118
  orthogonally stable, 122

Finite order, 37
Flat
  in a point, 21
  on a set, 85
  othogonally, 85
Formal power series, 20

Lower bound, 64
Lower distance estimate, 77
  for derivatives, 83

Orthogonal stability, 121
Orthogonally flat, 85

Regularly situated, 105

Semiproper, 36

Taylor
  map, 20

Whitney-regular set, 39

Bibliography

[AKOW98] Graham Allan, Grayson Kakiko, A. G. O'Farrell, and R. O. Watson. Finitely-generated algebras of smooth functions, in one dimension. *J. Funct. Anal.*, 158(2):458–474, 1998.

[AKOW01] G. R. Allan, G. Kakiko, A. G. O'Farrell, and R. O. Watson. Localness of $A(\Psi)$ algebras. *Math. Proc. R. Ir. Acad.*, 101A(1):61–70, 2001.

[BM88] Edward Bierstone and Pierre D. Milman. Semianalytic and subanalytic sets. *Inst. Hautes Études Sci. Publ. Math.*, (67):5–42, 1988.

[BM98] Edward Bierstone and Pierre D. Milman. Geometric and differential properties of subanalytic sets. *Ann. of Math. (2)*, 147(3):731–785, 1998.

[BMP96] Edward Bierstone, Pierre D. Milman, and Wiesław Pawłucki. Composite differentiable functions. *Duke Math. J.*, 83(3):607–620, 1996.

[Bou04] Nicolas Bourbaki. *Functions of a real variable.* Elements of Mathematics. Springer-Verlag, Berlin, 2004. Elementary theory, Translated from the 1976 French original [MR0580296] by Philip Spain.

[BS83] Edward Bierstone and Gerald W. Schwarz. Continuous linear division and extension of $\mathcal{C}^\infty$ functions. *Duke Math. J.*, 50(1):233–271, 1983.

[FdB59] François Faà di Bruno. *Théorie générale de l'élimination.* Librairie centrale des sciences, Paris, 1859.

[Gla58] Georges Glaeser. Étude de quelques algèbres tayloriennes. *J. Analyse Math.*, 6:1–124; erratum, insert to 6 (1958), no. 2, 1958.

[Gla63] Georges Glaeser. Fonctions composées différentiables. *Ann. of Math. (2)*, 77:193–209, 1963.

[Hör03] Lars Hörmander. *The analysis of linear partial differential operators. I.* Classics in Mathematics. Springer-Verlag, Berlin, 2003. Distribution theory and Fourier analysis, Reprint of the second edition (1990).

[KW11] Nicolas Kenessey and Jochen Wengenroth. Composition operators with closed range for smooth injective symbols $\mathbb{R} \to \mathbb{R}^d$. *J. Funct. Anal.*, 260(10):2997–3006, 2011.

[Mal67] B. Malgrange. *Ideals of differentiable functions*. Tata Institute of Fundamental Research Studies in Mathematics, No. 3. Tata Institute of Fundamental Research, Bombay, 1967.

[MV97] Reinhold Meise and Dietmar Vogt. *Introduction to functional analysis*, volume 2 of *Oxford Graduate Texts in Mathematics*. The Clarendon Press Oxford University Press, New York, 1997.

[PB87] Pedro Pérez Carreras and José Bonet. *Barrelled locally convex spaces*, volume 131 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam, 1987. Notas de Matemática, 113.

[Rud73] Walter Rudin. *Functional analysis*. McGraw-Hill Book Co., New York, 1973. McGraw-Hill Series in Higher Mathematics.

[Sch66] Laurent Schwartz. *Théorie des distributions*. Publications de l'Institut de Mathématique de l'Université de Strasbourg, No. IX-X. Nouvelle édition, entiérement corrigée, refondue et augmentée. Hermann, Paris, 1966.

[Spi65] Michael Spivak. *Calculus on manifolds. A modern approach to classical theorems of advanced calculus*. W. A. Benjamin, Inc., New York-Amsterdam, 1965.

[Tou71] J.-Cl. Tougeron. An extension of Whitney's spectral theorem. *Inst. Hautes Études Sci. Publ. Math.*, (40):139–148, 1971.

[Tou72] Jean-Claude Tougeron. *Idéaux de fonctions différentiables*. Springer-Verlag, Berlin, 1972. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 71.

[Trè67] François Trèves. *Topological vector spaces, distributions and kernels*. Academic Press, New York, 1967.

[Whi48] Hassler Whitney. On ideals of differentiable functions. *Amer. J. Math.*, 70:635–658, 1948.