1

Calculus in Locally Convex Spaces

1.1 Introduction

It is well known that ‘multidimensional calculus’, aka ‘Fréchet calculus’, carries over to the realm of Banach spaces and Banach manifolds (see e.g. Lang, 1999). As we have seen in the Preface, Banach spaces are often not sufficient for our purposes. To generalise derivatives we will, as a minimum, need vector spaces with an amenable topology (which need not be induced by a norm).

1.1 Definition Consider a vector space E. A topology $\mathcal{T}$ on E making addition $+\colon E \times E \to E$ and scalar multiplication $\cdot\colon \mathbb{R} \times E \to E$ continuous is called a *vector topology* (where $\mathbb{R}$ carries the usual norm topology). We then say that $(E, \mathcal{T})$ (or E for short) is a *topological vector space* (or *TVS* for short).

1.2 Example (a) Every normed space and, in particular, every finite-dimensional vector space is a topological vector space.

(b) For a more interesting example, fix $0 < p < 1$. Two measurable functions $\gamma, \eta\colon [0,1] \to \mathbb{R}$ are equivalent $\gamma \sim \eta$ if and only if $\int_0^1 |\gamma(s) - \eta(s)| \mathrm{d}s = 0$. Denote by $L^p[0,1]$ the vector space of all equivalence classes $[\gamma]$ of functions such that $\int_0^1 |\gamma(s)|^p \mathrm{d}s < \infty$. Topologise $L^p[0,1]$ via the metric topology induced by

$$d([\gamma],[\eta]) := \int_0^1 |\gamma(s) - \eta(s)|^p \mathrm{d}s.$$

In a metric space, we can test continuity of the vector space operations using sequences. For this, pick $\lambda_n \to \lambda \in \mathbb{R}$ and $[\gamma_n] \to [\gamma], [\eta_n] \to [\eta]$ (with respect to d) and use the triangle inequality to obtain:

https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press

$$d(\lambda_n[\gamma_n] + [\eta_n], \lambda[\gamma] + [\eta]) \\ \leq |\lambda_n - \lambda|^p d([\gamma_n],[0]) + |\lambda|^p d([\gamma_n],[\gamma]) + d([\eta_n],[\eta]).$$

This shows that the vector space operations are continuous, that is, $L^p[0,1]$ is a TVS.

In topological vector spaces, differentiable curves can be defined as follows:

1.3 Definition Let E be a topological vector space. A continuous mapping $\gamma\colon I \to E$ from a non-degenerate interval[1] $I \subseteq \mathbb{R}$ is called a C^0-curve. A C^0-curve is called a C^1-*curve* if the limit

$$\gamma'(s) := \lim_{t\to 0} \frac{1}{t}(\gamma(s+t) - \gamma(s))$$

exists for all $s \in I^\circ$ (interior of I) and extends to a continuous map $\frac{\mathrm{d}}{\mathrm{d}t}\gamma := \gamma'\colon I \to E$, $s \mapsto \gamma'(s)$. Recursively for $k \in \mathbb{N}$, we call γ a C^k-*curve* if γ is a C^{k-1}-curve and $\frac{\mathrm{d}^{k-1}}{\mathrm{d}t^{k-1}}\gamma$ is a C^1-curve. Then $\frac{\mathrm{d}^k}{\mathrm{d}t^k}\gamma := \left(\frac{\mathrm{d}^{k-1}}{\mathrm{d}t^{k-1}}\gamma\right)'$. If γ is a C^k-curve for every $k \in \mathbb{N}_0$, we also say that γ is *smooth* or of class C^∞.

Unfortunately, calculus on topological vector spaces is, in general, ill behaved. The next exercise shows that derivatives may fail to give us meaningful information.

Exercises

1.1.1 Given $0 < p < 1$ we let $L^p[0,1]$ be the topological vector space from Example 1.2(b). Recall that the topology on $L^p[0,1]$ is induced by the metric $d([\gamma],[\eta]) := \int_0^1 |\gamma(s) - \eta(s)|^p \mathrm{d}s$. For a set $A \subseteq [0,1]$ write $\mathbf{1}_A$ for the characteristic function and define

$$\beta\colon [0,1] \to L^p[0,1], \quad \beta(t) := [\mathbf{1}_{[0,t[}].$$

Show that β is an injective C^1-curve with $\beta'(t) = 0$, for all $t \in [0,1]$.

Obviously we would like to avoid this defect, and so we have to strengthen the assumptions on our vector spaces.

[1] That is, I has more than one point. In the following, we will always assume this when talking about intervals.

https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press

1.2 Curves in Locally Convex Spaces

Calculus in topological vector spaces exhibits pathologies that can be avoided by strengthening the requirements on the underlying space. This leads to locally convex spaces, whose topology is induced by so-called seminorms. See also Appendix A for more information on locally convex spaces.

1.4 Definition Let E be a vector space. A map $p\colon E \to [0,\infty[$ is called a *seminorm* if it satisfies the following:

(a) $p(\lambda x) = |\lambda| p(x), \forall \lambda \in \mathbb{R}, x \in E$,
(b) $p(x+y) \leq p(x) + p(y)$.

Note that, in contrast with the definition of a norm, we did not require that $p(x) = 0$ if and only if $x = 0$. The next definition uses the notion of an initial topology, which we recall for the reader's convenience in Appendix B.

1.5 Definition A topological vector space $(E,\mathcal{T})$ is called a *locally convex space* if there is a family $\{p_i : E \to [0,\infty[\mid i \in I\}$ of continuous seminorms for some index set I such that

(a) $\mathcal{T}$ is the initial topology with respect to the canonical projections $\{q_i : E \to E/p_i^{-1}(0)\}_{i\in I}$ onto the normed spaces $E/p_i^{-1}(0)$.
(b) If $x \in E$ with $p_i(x) = 0$ for all $i \in I$, then $x = 0$. Thus the seminorms separate the points, that is, $\mathcal{T}$ has the Hausdorff property.[2]

We then say that the topology $\mathcal{T}$ is *generated by the family of seminorms* $\{p_i\}_{i\in I}$ and call this family a *generating family of seminorms*. Usually we suppress $\mathcal{T}$ and write $(E,\{p_i\}_{i\in I})$ or simply E instead of $(E,\mathcal{T})$.

Alternative to (a) We will see in Appendix A that equivalent to (a), we can define $\mathcal{T}$ to be the unique vector topology determined by the basis of 0-neighbourhoods given by (finite) intersections of the balls $B_{i,\varepsilon}(0) = \{x \in E \mid p_i(x) < \varepsilon\}$, where p_i runs through a generating family of seminorms. These balls are all convex, thus justifying the name locally convex space.

A locally convex space $(E,\{p_i\}_{i\in\mathbb{N}})$ with a countable system of seminorms is *metrisable* (i.e. its topology is induced by a metric; see Exercise 1.2.1) and if E is complete, it is called *Fréchet space*.

1.6 Example (a) Every normed space $(E,\|\cdot\|)$ is a locally convex space, where the family of seminorms consists only of the norm $\|\cdot\|$.

[2] Some authors do not require separation of points, whence our locally convex spaces are Hausdorff locally convex spaces in their terminology.

(b) Consider the space $C^\infty([0,1],\mathbb{R})$ of all smooth functions from the interval $[0,1]$ to $\mathbb{R}$ (with pointwise addition and scalar multiplication). This space is not naturally a normed space.[3] We define a family of seminorms on it via

$$\|f\|_n := \sup_{0\leq k\leq n}\left\|\frac{\mathrm{d}^k}{\mathrm{d}t^k}f\right\|_\infty = \sup_{0\leq k\leq n}\sup_{t\in[0,1]}\left|\frac{\mathrm{d}^k}{\mathrm{d}t^k}f(t)\right|, n\in\mathbb{N}_0.$$

The topology generated by the seminorms is called the *compact-open C^∞-topology* and turns $C^\infty([0,1],\mathbb{R})$ into a locally convex space, which is even a Fréchet space (Exercise 1.2.2).

Locally convex spaces have many good properties, for example, they admit enough continuous linear functions to separate the points, that is, the following holds.

1.7 Theorem (Hahn–Banach (Meise and Vogt, 1997, Proposition 22.12)) *For a locally convex space E the continuous linear functionals separate the points, that is, for each pair $x, y \in E$ there exists a continuous linear $\lambda\colon E \to \mathbb{R}$ such that $\lambda(x) \neq \lambda(y)$.*

1.8 Definition Let E be a locally convex space, then we denote by $E' = L(E,\mathbb{R})$ the continuous linear maps from E to $\mathbb{R}$. The space E' is the so-called *dual space* of E. There are several ways to turn E' into a locally convex space (Rudin, 1991, p. 63f) but, in general, we will not need a topology beyond the special case if E is a Banach space and E' carries the operator norm topology.

With the help of the Hahn–Banach theorem, we can avoid the pathologies observed for topological vector spaces. To this end, we need the notion of a weak integral.

1.9 Definition Let $\gamma\colon I \to E$ be a C^0-curve in a locally convex space E and $a, b \in I$. If there exists $z \in E$ such that

$$\lambda(z) = \int_a^b \lambda(\gamma(t))\mathrm{d}t, \quad \forall\lambda\in E',$$

then $z \in E$ is called the *weak integral of* γ from a to b and denoted $\int_a^b \gamma(t)\mathrm{d}t := z$.

Note that weak integrals (if they exist) are uniquely determined due to the Hahn–Banach theorem.

[3] For any normed topology, the differential operator $D\colon C^\infty([0,1],\mathbb{R}) \to C^\infty([0,1],\mathbb{R})$, $D(f) = f'$ must be discontinuous (which is certainly undesirable). To see this, recall that a continuous linear map on a normed space has bounded spectrum, but D has arbitrarily large eigenvalues (consider $f_n(t) := \exp(nt)$, $n \in \mathbb{N}$).

1.10 Proposition (First part of the fundamental theorem of calculus) *Let $\gamma\colon I \to E$ be a C^1-curve in a locally convex space E and $a, b \in I$, then*

$$\gamma(b) - \gamma(a) = \int_a^b \gamma'(t)\mathrm{d}t.$$

Proof Let $\lambda \in E'$. It is easy to see that $\lambda \circ \gamma\colon I \to \mathbb{R}$ is a C^1-curve with $(\lambda \circ \gamma)' = \lambda \circ (\gamma')$. The standard fundamental theorem of calculus yields

$$\lambda(\gamma(b) - \gamma(a)) = \lambda(\gamma(b)) - \lambda(\gamma(a)) = \int_a^b (\lambda \circ \gamma)'(s)\mathrm{d}s = \int_a^b \lambda(\gamma'(s))\mathrm{d}s.$$

Hence $z = \gamma(b) - \gamma(a)$ satisfies the defining property of the weak integral. □

Note that Proposition 1.10 implies that $L^p[0,1]$ cannot be a locally convex space for $0 < p < 1$; see Rudin (1991, 1.47) for an elementary proof of this fact.

1.11 Remark Also the second part of the fundamental theorem of calculus is true in our setting. Thus if $\gamma\colon I \to E$ is a C^0-curve, $a \in I$ and the weak integral

$$\eta(t) := \int_a^t \gamma(s)\mathrm{d}s$$

exists for all $t \in I$. Then $\eta\colon I \to E$ is a C^1-curve in E, and $\eta' = \gamma$.

The proof, however, needs more techniques based on convex sets which we do not wish to go into (see Glöckner and Neeb, forthcoming).

The reader may wonder now, when do weak integrals of curves exist? One can prove that weak integrals of continuous curves always exist *in the completion of a locally convex space*. The key point is that the integrals can be defined using Riemann sums, but these do not necessarily converge in the space itself (Kriegl and Michor, 1997, Lemma 2.5). Thus weak integrals exist for suitably complete spaces. To avoid getting bogged down with the discussion of completeness properties, we define the following:

1.12 Definition A locally convex space E is *Mackey complete* if for each smooth curve $\gamma\colon [0,1] \to E$ there exists a smooth curve $\eta\colon [0,1] \to E$ with $\eta' = \gamma$.

Due to the fundamental theorem of calculus this implies that $\eta(s) - \eta(0) = \int_0^s \gamma(t)\mathrm{d}t$. Thus the weak integral of smooth curves exists in Mackey complete spaces.

1.13 Remark Mackey completeness is a very weak completeness condition, in particular, sequential completeness (i.e. Cauchy sequences converge in the space) implies Mackey completeness. This is evident from the

https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press

alternative characterisation of Mackey completeness using sequences; see Definition A.1. Note, however, that it is not entirely trivial to find examples of Mackey complete but not sequentially complete spaces. We mention here that the space $\mathcal{K}(E,F)$ of compact operators between two (infinite-dimensional) Banach spaces E,F with the strong operator topology is not sequentially complete but Mackey complete (see Voigt, 1992).

However, in metrisable locally convex spaces (e.g. in normed spaces) Mackey completeness is equivalent to completeness; see Jarchow (1981, 10.1.4). We refer to Kriegl and Michor (1997, I.2) for more information on Mackey completeness. In particular, Kriegl and Michor (1997, Theorem 2.14) show that integrals exist for C^1-curves in Mackey complete spaces.

So far we have defined differentiable curves with values in locally convex spaces. The next step is to consider differentiable mappings between locally convex spaces. Here a different notion of calculus is needed. It turns out that (even on Fréchet spaces) there are many generalisations of Fréchet calculus (see Keller, 1974) without a uniquely preferable choice. In the next section, we present a simple and versatile notion called Bastiani calculus. Another popular approach to calculus in locally convex spaces, the so-called convenient calculus, is discussed in Appendix A.7.

Exercises

1.2.1 Let $(E,\{p_n\}_{n\in\mathbb{N}})$ be a locally convex space whose topology is generated by a countable set of seminorms. Prove that

$$d(x,y) := \sum_{n\in\mathbb{N}} 2^{-n} \frac{p_n(x-y)}{p_n(x-y)+1}$$

is a metric on E and the metric topology coincides with the locally convex topology.

1.2.2 Consider $C^\infty([0,1],\mathbb{R})$ with the compact open C^∞-topology (see Example 1.6).

(a) Show that a sequence $(f_k)_{k\in\mathbb{N}}$ converges to f in this topology if and only if for all $\ell \in \mathbb{N}_0$ $\left(\frac{\mathrm{d}^\ell}{\mathrm{d}t^\ell} f_k\right)_k$ converges uniformly to $\frac{\mathrm{d}^\ell}{\mathrm{d}t^\ell} f$.

Hint: The uniform limit of a sequence of continuous functions is continuous. If a function sequence and the sequence of (first) derivatives converges, the limit of the sequence is differentiable.

(b) Deduce that every Cauchy sequence in the compact open C^∞-topology converges to a smooth function. As $C^\infty([0,1],\mathbb{R})$ is

https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press

a metric space by Exercise 1.2.1, this implies that the space is complete, that is, a Fréchet space.

(c) Show that the differential operator

$$D\colon C^\infty([0,1],\mathbb{R}) \to C^\infty([0,1],\mathbb{R}), \qquad f \mapsto f'$$

is continuous linear. *Hint:* Lemma A.5.

1.2.3 Let $(E,\{p_i\}_I)$ be a locally convex space whose topology is generated by a *finite* set of seminorms. Show that $p(x) = \max_{i\in I} p_i(x)$ defines a norm on E, which induces the same topology as the family $\{p_i\}$. In this case we call E normable.

1.2.4 Establish the following properties of weak integrals:

(a) If the weak integrals of $\gamma,\eta\colon [\mathrm{a},\mathrm{b}] \to E$ from a to b exist and $s \in \mathbb{R}$, then also the weak integral of $\gamma + s\eta$ exists and $\int_a^b (\gamma(t) + s\eta(t))\mathrm{d}t = \int_a^b \gamma(t)\mathrm{d}t + s\int_a^b \eta(t)\mathrm{d}t$.

(b) If $\gamma\colon [\mathrm{a},\mathrm{b}] \to E$ is constant, $\gamma(t) \equiv K$, then $\int_a^b \gamma(t)\mathrm{d}t$ exists and equals $(b-a)K$.

(c) $\int_a^c \gamma(t)\mathrm{d}t = \int_a^b \gamma(t)\mathrm{d}t + \int_b^c \gamma(t)\mathrm{d}t$ (if the integrals exist).

1.2.5 Let $\gamma\colon I \to E$ be a C^k-curve ($k \in \mathbb{N}$) and $\lambda\colon E \to F$ be continuous linear for E,F locally convex. Show that $\lambda \circ \gamma$ is C^k such that $\frac{\mathrm{d}^\ell}{\mathrm{d}t^\ell}(\lambda\circ\gamma) = \lambda \circ \left(\frac{\mathrm{d}^\ell}{\mathrm{d}t^\ell}\gamma\right)$, $1 \leq \ell \leq k$.

1.2.6 Endow a vector space E with a topology $\mathcal{T}$ generated by seminorms as in Definition 1.5. Show that $(E,\mathcal{T})$ is a topological vector space (and so requiring that locally convex spaces are topological vector spaces was superfluous).

1.3 Bastiani Calculus

Bastiani calculus (also called Keller's C_c^k-theory; Keller, 1974), introduced in Bastiani (1964), builds a calculus around directional derivatives and their continuity. It is the basis of our investigation as this calculus works in locally convex spaces beyond the Banach setting.

1.14 Definition Let E,F be locally convex spaces, $U \subseteq_{\circ} E$, $f\colon U \to F$ a map and $r \in \mathbb{N}_0 \cup \{\infty\}$. If it exists, we define for $(x,h) \in U \times E$ the *directional derivative*

$$df(x;h) \coloneqq D_h f(x) \coloneqq \lim_{\mathbb{R}\setminus\{0\}\ni t\to 0} t^{-1}(f(x+th) - f(x)).$$

https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press

We say that f is C^r if the *iterated directional derivatives*

$$d^k f(x; y_1, \dots, y_k) := (D_{y_k} D_{y_{k-1}} \cdots D_{y_1} f)(x)$$

exist for all $k \in \mathbb{N}_0$ such that $k \leq r$, $x \in U$ and $y_1, \dots, y_k \in E$ and define continuous maps $d^k f\colon U \times E^k \to F$ (where $d^0 f := f$). If f is C^k for all $k \in \mathbb{N}_0$ we say that f is *smooth* or C^∞. Note that $df = d^1 f$ and for curves $c\colon I \to E$ we have $c'(t) = dc(t; 1)$.

1.15 Remark Note that the iterated directional derivatives are only taken with respect to the first variable (i.e. of the map $x \mapsto df(x; v)$, where v is supposed to be fixed). One can alternatively define iterated differentials to derivate with respect to all variables, but this leads to the same differentiability concept (see Glöckner, 2002 for a detailed explanation). The following observations are easily proved from the definitions:

(a) $d^2 f(x; v, w) = \lim_{t\to 0} t^{-1}(df(x + tw; v) - df(x; v))$.
(b) $d^k f(x; v_1, \dots, v_k) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} d^{k-1} f(x + tv_k; v_1, \dots, v_{k-1})$.
(c) f is C^k if and only if f is C^{k-1} and $d^{k-1} f$ is C^1. Then $d^k f = d(d^{k-1} f)$.

Finally, there is a version of the *Schwarz theorem* which states that the order of directions $v_1, \dots, v_k$ in $d^k f(x; v_1, \dots, v_k)$ is irrelevant (see Exercise 1.3.3).

1.16 Example Let $A\colon E \to F$ be a continuous linear map between locally convex spaces. Then A is C^1, as we can exploit

$$dA(x; v) = \lim_{t\to 0} t^{-1}(A(x + tv) - A(x)) = \lim_{t\to 0} A(v) = A(v).$$

In particular, since A is continuous, so is the first derivative and we see that A is a C^1-map. Computing the second derivative, we use that the first derivative is constant in x (but not in v!) to obtain

$$\begin{aligned} d^2 A(x; v, w) &= D_w(dA(x; v)) = \lim_{t\to 0} t^{-1}(dA(x + tw; v) - dA(x; v)) \\ &= \lim_{t\to 0} t^{-1}(A(v) - A(v)) = 0. \end{aligned}$$

In conclusion A is a C^2-map (obviously even a C^∞-map) whose higher derivatives vanish.

1.17 Lemma *Let $f\colon E \supseteq U \to F$ be a C^1-map. Then $df(x; \cdot)$ is homogeneous, that is, $df(x; sv) = s\,df(x; v)$ for all $x \in U, v \in E$ and $s \in \mathbb{R}$.*

Proof As $df(x; 0v) = df(x; 0) = 0 = 0\,df(x; v)$, we may assume that $s \neq 0$ and thus $df(x; sv) = \lim_{t\to 0} t^{-1}(f(x + tsv) - f(x)) = s \lim_{t\to 0} (st)^{-1}(f(x + tsv) - f(x)) = s\,df(x; v)$. □

1.18 Proposition (Mean value theorem on locally convex spaces) *Let E, F be locally convex spaces and $f : U \to F$ a C^1-map on $U \Subset E$. Then*

$$f(y) - f(x) = \int_0^1 df(x + t(y - x); y - x)\mathrm{d}t \tag{1.1}$$

for all $x, y \in U$ such that U contains the line segment $\overline{xy} := \{tx + (1-t)y \mid t \in [0,1]\}$.

Proof Note that the curve $\gamma : [0,1] \to F, \gamma(t) := f(x + t(y - x))$ is differentiable at each $t \in [0,1]$. Its derivative is

$$\gamma'(t) = \lim_{s \to 0} s^{-1}(\gamma(t + s) - \gamma(t)) = df(x + t(y - x), y - x),$$

whence γ' is continuous (as df is) and thus a C^1-curve. Apply now the Fundamental theorem 1.10 to γ' to obtain (1.1). □

On a locally convex space, every point has arbitrarily small convex neighbourhoods. Convex neighbourhoods contain all line segments between points in the neighbourhood, whence Proposition 1.18 is available on these neighbourhoods. As a consequence we obtain the following.

1.19 Corollary *If $f : U \to F$ is a C^1-map with $df \equiv 0$, then f is locally constant.*

Proof For $x \in U$ choose a convex neighbourhood $x \in V \subseteq U$ (see Appendix A). For each $y \in V$ the line segment connecting x and y is contained in V, and so the vanishing of the derivative with (1.1) implies $f(x) = f(y)$ and f is constant on V. □

1.20 Proposition (Rule on partial differentials) *Let E_1, E_2, F be locally convex spaces, $U \Subset E_1 \times E_2$ and let $f : U \to F$ be continuous. Then f is C^1 if and only if the limits*

$$d_1 f(x, y; v_1) := \lim_{t \to 0} t^{-1}(f(x + tv_1, y) - f(x, y)),$$
$$d_2 f(x, y; v_2) := \lim_{t \to 0} t^{-1}(f(x, y + tv_2) - f(x, y))$$

exist for all $(x, y) \in U$ and $(v_1, v_2) \in E_1 \times E_2$ and extend to continuous mappings $d_i f : U \times E_i \to F, i = 1, 2$. In this case,

$$df(x, y; v_1, v_2) = d_1 f(x, y; v_1) + d_2 f(x, y; v_2), \quad \forall (x, y) \in U, (v_1, v_2) \in E_1 \times E_2. \tag{1.2}$$

Proof If f is C^1 the mappings $d_i f$ clearly exist and are continuous. Conversely, let us assume that the mappings $d_i f$ exist and are continuous. For

https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press

$(x,y) \in U$, $(v_1,v_2) \in E \times F$, we fix $\varepsilon > 0$ such that $(x,y) + t(v_1,v_2) \in U$ whenever $|t| < \varepsilon$. Now if we fix the ith component of f we obtain a C^1-mapping (by hypothesis, the derivative is $d_i f$). Therefore Proposition 1.18 with Lemma 1.17 yields

$$\begin{aligned}
&\frac{f((x,y)+t(v_1,v_2)) - f(x,y)}{t} \\
&\quad = \frac{f(x+tv_1, y+tv_2) - f(x+tv_1,y)}{t} + \frac{f(x+tv_1,y) - f(x,y)}{t} \\
&\quad = \int_0^1 d_2 f(x+tv_1, y+stv_2; v_2)\mathrm{d}s + \int_0^1 d_1 f(x+stv_1, y; v_1)\mathrm{d}s. \qquad (1.3)
\end{aligned}$$

The integrals (1.3) make sense also for $t = 0$, whence they define maps $I_i\colon]-\varepsilon,\varepsilon[\to H$. Due to continuous dependence on the parameter t,[4] the right-hand side of (1.3) converges for $t \to 0$. We deduce that the limit df exists and satisfies (1.2) which is continuous, whence f is C^1. □

The following alternative characterisation of C^1-maps will turn the proof of the chain rule into a triviality. However, we shall only sketch the proof to avoid discussing convergence issues of the weak integral involved.

1.21 Lemma *A map $f\colon E \supseteq U \to F$ is of class C^1 if and only if there exists a continuous mapping, the difference quotient map,*

$$f^{[1]}\colon U^{[1]} := \{(x,v,s) \in U \times E \times \mathbb{R} \mid x + sv \in U\} \to F$$

such that $f(x+sv) - f(x) = sf^{[1]}(x,v,s)$ for all $(x,v,s) \in U^{[1]}$.

Proof Let us assume first that $f^{[1]}$ exists and is continuous. Note that $U^{[1]} \Subset U \times E \times \mathbb{R}$. Then $df(x;v) = f^{[1]}(x,v,0)$ exists and is continuous as a partial map of $f^{[1]}$. So f is C^1. Conversely, if f is C^1, the map

$$f^{[1]}(x,v,s) := \begin{cases} s^{-1}(f(x+sv) - f(x)), & (x,v,s) \in U^{[1]}, s \neq 0, \\ df(x;v), & (x,v,s) \in U^{[1]}, s = 0 \end{cases}$$

is continuous on the open set $U^{[1]} \setminus \{(x,v,s) \in U^{[1]} \mid s = 0\}$. That $f^{[1]}$ extends to a continuous map on all of $U^{[1]}$ follows from continuity of parameter-dependent weak integrals; see Bertram et al. (2004, Proposition 7.4) for details. □

1.22 Lemma *If $f\colon E \supseteq U \to F$ is C^1, then $df(x;\cdot)\colon E \to F$ is a continuous linear map for each $x \in U$.*

[4] We are cheating here; the continuous dependence of weak integrals on parameters has not been established in this book. See Hamilton (1982, I Theorem 2.1.5) for a proof that carries over to our setting.

Proof Fix $x \in U$ and note that $df(x;\cdot)$ is continuous as a partial map of the continuous df. We have already seen in Lemma 1.17 that $df(x;sv) = sdf(x;v)$ for all $s \in \mathbb{R}$, and so we only have to prove additivity. Choosing $v,w \in E$ we compute

$$\begin{aligned} &t^{-1}(f(x+t(v+w)-f(x)) \\ &\quad = t^{-1}(f(x+t(v+w)-f(x+tv)) + t^{-1}(f(x+tv)-f(x)) \\ &\quad = f^{[1]}(x+tv,w,t) + f^{[1]}(x,v,t). \end{aligned}$$

The right-hand side also makes sense for $t = 0$ and is continuous, whence passing to the limit we get $df(x;v+w) = df(x;v) + df(x;w)$. □

1.23 Proposition (Chain rule) *Let $f\colon E \supseteq U \to F$ and $g\colon F \supseteq W \to K$ be C^1-maps with $f(U) \subseteq W$. Then $g \circ f$ is a C^1-map with derivative given by*

$$d(g \circ f)(x;v) = dg(f(x);df(x,v)) \quad (\text{i.e.}(g \circ f)'(x) = g'(f(x)) \circ f'(x)).$$

Proof We use the notation from Lemma 1.21 and write for $(x,y,t) \in U^{[1]}$ with $t \neq 0$,

$$\begin{aligned} t^{-1}(g(f(x+ty)) - g(f(x))) &= t^{-1}\left(g\left(f(x) + t\frac{f(x+ty)-f(x)}{t}\right) - g(f(x))\right) \\ &= g^{[1]}(f(x), f^{[1]}(x,y,t),t). \end{aligned} \tag{1.4}$$

The function $h\colon U^{[1]} \to K, h(x,y,t) := g^{[1]}(f(x),f^{[1]}(x,y,t),t)$ is continuous and extends the right-hand side of (1.4). Hence Lemma 1.21 shows that $g \circ f$ is C^1, with

$$(g \circ f)^{[1]}(x,y,t) = g^{[1]}(f(x),f^{[1]}(x,y,t),t) \text{ for all } (x,y,t) \in U^{[1]}.$$

Thus

$$\begin{aligned} d(g \circ f)(x;y) &= (g \circ f)^{[1]}(x,y,0) = g^{[1]}(f(x),f^{[1]}(x,y,0),0) \\ &= dg(f(x);df(x;y)). \end{aligned}$$ □

The chain rule is the basis to transport concepts from differential geometry such as manifolds, tangent spaces and so on to our setting. Later chapters will define these objects.

1.24 Lemma *Let $E,(F_i)_{i\in I}$ be locally convex spaces and $f\colon E \supseteq U \to \prod_{i\in I} F_i$ a map on an open subset. Let $k \in \mathbb{N}_0 \cup \{\infty\}$ and set $f_i : \mathrm{pr}_i \circ f : U \to F_i$ (where pr_i is the ith* coordinate projection*). Then f is C^k if and only if every $f_i, i \in I$ is C^k and*

$$df(x;v) = (df_i(x;v))_{i\in I}, \quad x \in U, v \in E. \tag{1.5}$$

Proof If f is C^k, we note that every $f_i = \mathrm{pr}_i \circ f$ is C^k by the chain rule as the projections are continuous linear, hence smooth. Further, $df_i = \mathrm{pr}_i \circ df$ which establishes (1.5).

For the converse, note that it suffices to assume that $k \in \mathbb{N}$. We argue by induction starting with $k = 1$. Then $t^{-1}f(x + tv) - f(x)) = (t^{-1}(f_i(x + tv) - f_i(x))_{i \in I}$. Now the components converge to $df_i(x; v)$ for $t \to 0$, whence the difference quotient converges to the limit in (1.5). Since every df_i is continuous, $df = (df_i)_{i \in I} : U \times E \to \prod_{i \in I} F_i$ is also continuous and f is C^1. For the induction step we notice that by induction, $\mathrm{pr}_i \circ df = df_i$ is C^{k-1}. By induction, df is C^{k-1} and thus f is C^k. □

A subset $Y \subseteq X$ of a topological space X is called *sequentially closed* if $\lim_{n\to\infty} x_n \in A$ for each sequence $(x_n)_{\mathbb{N}} \subseteq A$ which converges in X. The following lemma will be useful in the discussion of submanifolds in Section 1.5.

1.25 Lemma *Let $f: E \supseteq U \to F$ be a continuous map from an open subset of a locally convex space to a locally convex space, and $F_0 \subseteq F$ a sequentially closed vector subspace such that $f(U) \subseteq F_0$. Let $k \in \mathbb{N} \cup \{\infty\}$, then f is C^k if and only if the corestriction $f|^{F_0} : U \to F_0$ is C^k.*

Proof If $f|^{F_0}$ is C^k, so is $f = \iota \circ f|^{F_0}$, where $\iota : F_0 \to F$ is the continuous linear (hence smooth) inclusion.

Conversely, we argue by induction and assume first that f is C^1. For $x \in U$ and $v \in E$, pick a sequence $t_n \to 0$ such that $x + t_n v \in U$ for each $n \in \mathbb{N}$. Then $df(x; y) = \lim_{n\to\infty}(f(x + t_n v) - f(x)) \in F_0$ by sequential closedness. Hence the limit exists in F_0. Further, as a map $U \times E \to F_0$, $(x, v) \mapsto df(x; v)$ is continuous. We conclude that $f|^{F_0}$ is C^1. If f is C^k, $d(f|^{F_0}) = (df)|^{f_0}$ is C^{k-1} by induction and hence f is C^k. □

In Example A.32 we will see that in Lemma 1.25 sequential closedness is a necessary assumption for the validity of the statement.

Exercises

1.3.1 Check the details for Remark 1.15 (for Schwarz' theorem see below).

1.3.2 Let E_1, E_2, F be locally convex spaces and $\beta: E_1 \times E_2 \to F$ be a continuous bilinear map. Show that β is C^1 with first derivative $d\beta(x_1, x_2; y_1, y_2) = \beta(x_1, y_2) + \beta(y_1, x_2)$. Compute all higher derivatives of β and show that β is smooth.

1.3.3 *Schwarz Theorem*: If $f: E \supseteq U \to F$ is a C^k-map, and $x \in U$, prove that $d^r f(x; \cdot) : E^r \to F$ is symmetric for all $2 \leq r \leq k$ (i.e. the order of arguments is irrelevant to the function value).

Hint: By the Hahn–Banach theorem, it suffices to consider $\lambda \circ d^r f(x;\cdot) = d^r \lambda \circ f(x;\cdot)$, whence without loss of generality $F = \mathbb{R}$. Now for fixed $v_1,\dots,v_r$, the property can be checked in the finite-dimensional subspace generated by $v_1,\dots,v_r$.

1.3.4 The continuous functions $C([0,1],\mathbb{R})$ form a Banach space with respect to the supremum norm $\|\cdot\|_\infty$ (the resulting topology is the compact open topology).

(a) Let $f\colon \mathbb{R} \to \mathbb{R}$ be continuous and $(\gamma_n)_{n\in\mathbb{N}} \subseteq C([0,1],\mathbb{R})$ be a uniformly convergent sequence of functions with limit γ. Exploit that f is uniformly continuous on each ball and show that $f \circ \gamma_n \to f \circ \gamma$ uniformly. Deduce that the pushforward $f_*\colon C([0,1],\mathbb{R}) \to C([0,1],\mathbb{R}), \eta \mapsto f \circ \eta$ is continuous

Assume that $f \in C^1(\mathbb{R},\mathbb{R})$. Our aim will be to see that f_* is then C^1.

(b) Assume that the limit

$$df_*(\gamma;\eta) := \lim_{t\to 0} t^{-1}(f_*(\gamma + t\eta) - f_*(\gamma)) \tag{1.6}$$

exists. The point evaluation $\mathrm{ev}_x\colon C([0,1],\mathbb{R}) \to \mathbb{R}, \eta \mapsto \eta(x)$ is continuous linear for each $x \in [0,1]$. Apply ev_x to both sides of (1.6) and find the only possible candidate $\psi(\gamma,\eta)$ for $df_*(\gamma;\eta)$.

(c) Use point evaluations to verify that

$$t^{-1}(f_*(\gamma + t\eta) - f_*(\gamma)) = \int_0^1 \psi(\gamma + st\eta,\eta)\mathrm{d}s.$$

(d) Show that $\psi(\gamma,\eta)$ from (b) is indeed the directional derivative $df_*(\gamma;\eta)$.

(e) Verify that df_* is continuous, hence f_* is C^1.

1.3.5 Let E,F,H be locally convex spaces, $U \Subset E, V \Subset H$, $f\colon U \to F$ a C^2-map and let $g\colon V \to U$ and $h\colon V \to E$ be C^1. Prove that the differential of the C^1-map $\phi := df \circ (g,h)\colon V \to F, \phi(x) = df(g(x);h(x))$ is given by

$$d\phi(x;y) = d^2 f(g(x);h(x),dg(x;y)) + df(g(x);dh(x;y)), \tag{1.7}$$

for all $x \in V$, $y \in H$.

1.4 Bastiani versus Fréchet Calculus on Banach Spaces

On Banach spaces one usually defines differentiability in terms of the so-called Fréchet derivative. We briefly recall the definitions that should be familiar from basic courses on calculus. While Bastiani calculus is somewhat weaker than Fréchet calculus, the gap between those two can be quantified. We collect these results in the present section.

1.26 Definition A map $f\colon E \supseteq U \to F$ from an open subset of a normed space $(E, \|\cdot\|_E)$ to a normed space $(F, \|\cdot\|_F)$ is *continuous Fréchet differentiable* (or FC^1) if, for each $x \in U$, there exists a continuous linear map $A_x \in L(E,F)$ such that

$$f(x+h) - f(x) = A_x \cdot h + R_x(h) \text{ with } \lim_{h\to 0} \|R_x(h)\|_F / \|h\|_E = 0$$

and the mapping $Df\colon U \to L(E,F)$, $x \mapsto A_x$ is continuous (where the right-hand side carries the operator norm). Inductively we define for $k \in \mathbb{N}$ that f is a *k-times continuous Fréchet differentiable* map (or FC^k-map) if it is FC^1 and Df is FC^{k-1}. Moreover, f is (Fréchet-)smooth or FC^∞ if f is an FC^k map for every $k \in \mathbb{N}$.

The reader may wonder why the notion of Fréchet differentiability cannot be generalised beyond the setting of normed spaces. The reason for this is that the continuity of the derivative cannot be formulated as there is no suitable topology on the spaces $L(E,F)$. Indeed, there is no locally convex topology making evaluation and composition on the spaces $L(E,F)$ continuous. We refer to Proposition A.19 for an example of this pathology in the context of dual spaces.

From the definition, it is apparent (Exercise 1.4.2) that if f is FC^1, then $Df(x)(h) = df(x;h)$ and thus every FC^1-map is automatically C^1 in the Bastiani sense. However, one learns in basic calculus courses that existence of directional derivatives is weaker than the existence of derivatives in the Fréchet sense. The next example exhibits this.

1.27 Example (Bastiani C^1 is weaker than Fréchet C^1 (Milnor 1982)) Consider the Banach space

$$\ell^1 = \left\{ (x_n)_{n\in\mathbb{N}} \mid x_n \in \mathbb{R}, \text{ for all } n \in \mathbb{N}, \|(x_n)\|_{\ell^1} = \sum_{n\in\mathbb{N}} |x_n| < \infty \right\}.$$

Let $\varphi(u) := \log(1+u^2)$ and $\psi(u) := \frac{\mathrm{d}}{\mathrm{d}u}\varphi(u) = \frac{2u}{1+u^2}$. We observe that $|\psi(u)| \leq 1$, whence $|\varphi(u)| \leq |u|$ and we obtain a well-defined map

$$f\colon \ell^1 \to \mathbb{R}, f((x_n)) := \sum_{n\in\mathbb{N}} \frac{\varphi(nx_n)}{n}.$$

Observe that $|f((x_n))| \leq \|(x_n)\|_{\ell^1}$. In Exercise 1.4.3, we will show that f is C^1 with differential $df((x_n);(v_n)) = \sum_{n\in\mathbb{N}} v_n \psi(nx_n)$ but not Fréchet differentiable.

1.28 (Bastiani vs. Fréchet calculus) While Bastiani C^k is weaker than FC^k, Walter (2012, Appendix A.3) shows that there is only a mild loss of differentiability. In particular, $\mathrm{FC}^\infty = C^\infty$. The proofs are somewhat technical induction arguments involving the operator norm. Hence we only summarise the relation between the calculi *on Banach spaces* in the following diagram (arrows denote implications between conditions):

$$C^{k+1} \Longrightarrow \mathrm{FC}^k \underset{\dim E < \infty}{\overset{\Longrightarrow}{\Longleftarrow}} C^k$$

Exercises

1.4.1 Let $(E, \|\cdot\|_E), (F, \|\cdot\|_F)$ be normed spaces and the space $L(E,F)$ of continuous linear maps be endowed with the operator norm $\|A\|_{\mathrm{op}} = \sup_{x \in E\setminus\{0\}} \frac{\|A(x)\|_F}{\|x\|_E}$. Show that the evaluation map $\varepsilon \colon L(E,F) \times E \to F, \varepsilon(A,x) = A(x)$ is continuous.

1.4.2 Let $f \colon U \to F$ be an FC^1-map on $U \Subset E$, where E, F are Banach spaces.

(a) Show that the Fréchet derivative satisfies $Df(x)(h) = df(x;h)$ for every $x \in U$, $h \in E$ and deduce that f is C^1 in the Bastiani sense.

(b) Use induction to prove that every FC^k-map is already C^k by showing that the kth-Fréchet derivative gives rise to the kth derivative in the Bastiani sense.

1.4.3 We fill in the details for Example 1.27. Notation is as in the example. Prove that

(a) $|\psi(u)| \leq 1$ for all $u \in \mathbb{R}$.

(b) $df((x_n);(v_n)) = \sum_{n \in \mathbb{N}} v_n \psi(n x_n)$, hence continuous and thus f is C^1.

(c) $\|df((x_n);\cdot)\|_{\mathrm{op}}$ equals 0 if $x = 0$ but is ≥ 1 if $x_n = 1/n$ for some $n \in \mathbb{N}$.

(d) For $\delta_n := \begin{cases} 0 & \text{if } m \neq n, \\ 1/n & \text{if } m = n, \end{cases}$ the expression $\|df(\delta_n;\cdot)\|_{\mathrm{op}}$ does not converge to 0. Deduce that f is not FC^1.

https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press

1.5 Infinite-Dimensional Manifolds

In this section, we recall the basic notions of manifolds modelled on locally convex spaces. Most of these definitions should be very familiar from the finite-dimensional setting.

1.29 In this section we will write $g \circ f$ as a shorthand for $g \circ f|_{f^{-1}(A)}\colon f^{-1}(A) \to B$ if it helps to avoid clumsy notation.

1.30 Definition (Charts and atlas) Let M be a Hausdorff topological space. A *chart* for M is a homeomorphism $\varphi\colon U_\varphi \to V_\varphi$ from $U_\varphi \Subset M$ onto $V_\varphi \Subset E_\varphi$, where E_φ is a locally convex space. Let $r \in \mathbb{N}_0 \cup \{\infty\}$. A C^r*-atlas* for M is a set $\mathcal{A}$ of charts for M satisfying the following:

(a) $M = \bigcup_{\varphi \in \mathcal{A}} U_\varphi$.
(b) For all $\varphi, \psi \in \mathcal{A}$ the *change of charts* $\varphi \circ \psi^{-1}$ (which are mappings between open subsets of locally convex spaces) are C^r.[5]

Two C^r-atlases $\mathcal{A}, \mathcal{A}'$ for M are *equivalent* if their union $\mathcal{A} \cup \mathcal{A}'$ is a C^r-atlas for M. This is an equivalence relation.

1.31 Definition A C^r manifold $(M, \mathcal{A})$ is a Hausdorff topological space with an equivalence class of C^r-atlases $\mathcal{A}$. (If the equivalence class $\mathcal{A}$ is clear, we simply write M.)

1.32 Remark In contrast with the finite-dimensional case, we do not require manifolds to be paracompact or second countable (as topological spaces).

In general, the manifolds we are interested in will not be modelled on a single locally convex space. For a C^1-atlas, the locally convex spaces in which charts take their image are necessarily isomorphic on each connected component. However, some examples we will encounter later on have a huge number of connected components. For each of these connected components the locally convex model spaces will, in general, not be isomorphic.

1.33 Example Every locally convex space E is a manifold with global chart given by the identity id_E. Similarly, every $U \Subset E$ is a manifold with global chart given by the inclusion $U \to E$.

1.34 Example (Hilbert sphere) For a Hilbert space $(H, \langle \cdot, \cdot \rangle)$ the unit sphere $S_H := \{x \in H \mid \langle x, x \rangle = 1\}$ is a C^∞-manifold. To construct charts, we define the Hilbert space $H_{x_0} = \{y \in H \mid \langle y, x_0 \rangle = 0\} \subseteq H$ for $x_0 \in S_H$. (If dim

[5] Formally, if the charts do not intersect, the change of charts is the empty map $\emptyset \to \emptyset$ which is C^r.

$H < \infty$, H_{x_0} is a proper subspace of H, but if H is infinite-dimensional, it is isomorphic to H; see Dobrowolski, 1995). Then define the sets $U_{x_0} := \{x \in S_H \mid \langle x, x_0 \rangle > 0\}$ and $V_{x_0} := \{y \in H_{x_0} \mid \langle y, y \rangle < 1\}$. We can now define a chart

$$\varphi_{x_0} \colon U_{x_0} \to V_{x_0}, \quad x \mapsto x - \langle x, x_0 \rangle x_0$$

(its inverse is given by the formula $\varphi_{x_0}^{-1}(y) = y + \sqrt{1 - \langle y, y \rangle} x_0$). Applying these formulae, we see that the change of charts map for $x_0, z_0 \in S_H$ is a smooth map between open (possibly empty) subsets of Hilbert spaces:

$$\varphi_{z_0} \circ \varphi_{x_0}^{-1}(y) = (y - \langle y, z_0 \rangle z_0) + \sqrt{1 - \langle y, y \rangle}(x_0 - \langle x_0, z_0 \rangle z_0).$$

1.35 Definition Let M be a C^r-manifold and together with a sequentially closed vector subspace $F_\varphi \subseteq E_\varphi$ for each chart $\varphi \colon U_\varphi \to V_\varphi \underset{\circ}{\subseteq} E_\varphi$. A *($C^r$-)submanifold of M* is a subset $N \subseteq M$ such that for each $x \in N$, there exists a chart $\phi \colon U_\phi \to V_\phi$ of M around x such that $\phi(U_\phi \cap N) = V_\phi \cap F_\varphi$. Then $\phi_N := \phi|_{U_\phi \cap N}^{V_\phi \cap F_\varphi}$ is a chart for N, called a *submanifold chart*. Thanks to Lemma 1.25, the submanifold charts form a C^r-atlas for N.

If N is a submanifold of M such that all the sequentially closed subspaces F_ϕ are complemented subspaces of E_ϕ (see §1.7), we call N a *split submanifold* of M.

1.36 Definition Let $(M, \mathcal{A})$ and $(N, \mathcal{B})$ be C^r manifolds. Then the product $M \times N$ becomes a C^r-manifold using the atlas $C := \{\phi \times \psi \mid \phi \in \mathcal{A}, \psi \in \mathcal{B}\}$. We call the resulting C^r-manifold the *(direct) product of M and N*.

1.37 Definition Let $r \in \mathbb{N}_0 \cup \{\infty\}$ and M, N be C^r-manifolds. A map $f \colon M \to N$ is called C^r if f is continuous and, for every pair of charts ϕ, ψ, the map

$$\psi \circ f \circ \phi^{-1} \colon E_\phi \supseteq \phi(f^{-1}(U_\psi) \cap U_\phi) \to F_\psi$$

is a C^r-map. We write $C^r(M, N)$ for the set of all C^r-maps from M to N.

1.38 Remark Let $f \colon M \to N$ be a continuous map between C^r-manifolds. Assume that for some charts (U_φ, φ) and (U_ψ, ψ) the composition $\varphi \circ f \circ \psi$ is $C^r, r \in \mathbb{N} \cup \{\infty\}$. Then for any other pair of charts (U_κ, κ) and (U_λ, λ) with $f(U_\lambda \cap U_\psi) \subseteq U_\varphi \cap U_\kappa$ we have on $U_\lambda \cap U_\psi$ that

$$\kappa \circ f \circ \lambda^{-1}|_{\lambda(U_\lambda \cap U_\psi)} = (\kappa \circ \varphi^{-1}) \circ (\varphi \circ f \circ \psi^{-1}) \circ (\psi \circ \lambda^{-1}),$$

where the mapping in the middle is C^r by assumption and the other mappings are change of charts (which are C^r by M, N being C^r-manifolds). Hence $\kappa \circ f \circ \lambda^{-1}|_{\lambda(U_\lambda \cap U_\psi)}$ is also C^r by the chain rule, Proposition 1.23. This argument is called 'insertion of charts' and we leave it from now on to the reader. With the insertion of charts argument, it is easy to see that:

(a) It suffices to test the C^r-property with respect to any atlas of M and N.
(b) If $f\colon M \to N$ and $g\colon N \to L$ are C^r-maps, so is $g \circ f\colon M \to L$.
(c) Using Lemma 1.24: If M, N_1, N_2 are C^r-manifolds and $f_i\colon M \to N_i$, $i = 1,2$ are mappings. Then $f := (f_1, f_2)\colon M \to N_1 \times N_2$ is C^r if and only if f_1, f_2 are C^r.

1.39 Lemma *Let N be a (C^r-)submanifold of the C^r-manifold M. Then the inclusion $\iota\colon N \to M$ is C^r. Further, $f\colon P \to N$ is C^r if and only if $\iota \circ f$ is C^r.*

Proof Thanks to Remark 1.38, it suffices to check the C^r-property of ι in charts that cover N. Thus we may choose charts (ϕ, U_ϕ) of M that induce submanifold charts ϕ_N as in Definition 1.35. But then $\phi \circ \iota \circ \phi_N^{-1}$ is the inclusion map $V_\phi\colon F \to V$ which is C^r (as restriction of a continuous linear map).

If f is C^r, so is $\iota \circ f$ by Remark 1.38. Conversely, let $\iota \circ f$ be a C^r-map and ϕ, ϕ_N as before and ψ a chart for P. Then $\phi \circ \iota \circ f \circ \psi^{-1}\colon \psi(U_\psi \cap f^{-1}(U_\phi)) \to E$ is C^r with values in the sequentially closed subspace F and thus $(\phi \circ \iota \circ f \circ \psi)|^F = \phi_N \circ f \circ \psi^{-1}$ is C^r by Lemma 1.25. We conclude that f is C^r. □

Exercises

1.5.1 Let $f\colon M \to N$ be a C^r-map between C^r-manifolds. Show that the graph$(f) := \{(m, f(m)) \mid m \in M\}$ is a split submanifold of $M \times N$.
Hint: Use the description of the graph to construct submanifold charts by hand.

1.5.2 Verify the details of Example 1.34: Check that the charts make sense as mappings from U_{x_0} to V_{x_0}. Show that the change of charts $\varphi_{x_0} \circ \varphi_{z_0}^{-1}$ is smooth for all $x_0, z_0 \in S_H$ such that $U_{x_0} \cap U_{z_0} \neq \emptyset$.

1.5.3 Let M be a manifold and $U \Subset M$. Let $\mathcal{A}$ be an atlas of M. Endow U with the subspace topology and show that $\mathcal{A}_U := \{\phi|_{U \cap U_\phi} \mid (\phi, U_\phi) \in \mathcal{A}\}$ is a manifold atlas for U turning it into a submanifold of M.

1.5.4 Check that the set C in Definition 1.36 is a C^r-atlas for the product manifold.

1.5.5 Show that a compact manifold must be modelled on a finite-dimensional space.
Hint: Proposition A.3.

1.6 Tangent Spaces and the Tangent Bundle

1.40 Definition Let M be a C^r-manifold ($r \geq 1$) and $p \in M$. We say that a C^1-curve γ *passes through* p if $\gamma(0) = p$. For two such curves γ, η we define the relation

$$\gamma \sim \eta \Leftrightarrow (\phi \circ \gamma)'(0) = (\phi \circ \eta)'(0) \tag{1.8}$$

for some chart ϕ of M around p. By the chain rule, (1.8) holds for every chart around p and defines an equivalence relation on the set of all curves passing through p. The equivalence class $[\gamma]$ is called the *(geometric) tangent vector* (of M at p). Define the *(geometric) tangent space of M at p* as the set T_pM of all geometric tangent vectors at p.

We recall now that the tangent space at a point can be turned into a locally convex space isomorphic to the modelling space at that point.

1.41 Lemma (a) *Let ϕ be a chart of M around p, set $p_\phi := \phi(p)$. Then*

$$h_\phi \colon E_\phi \to T_pM, \quad h_\phi(y) := [t \mapsto \phi^{-1}(p_\phi + ty)]$$

is a bijection with inverse $h_\phi^{-1} \colon T_pM \to E_\phi, [\gamma] \mapsto (\phi \circ \gamma)'(0)$.

(b) *For all charts ϕ, ψ around p, we have $h_\psi^{-1} \circ h_\phi = d(\psi \circ \phi^{-1})(p_\phi; \cdot)$ which is an automorphism of the topological vector space E.*

(c) *T_pM admits a unique locally convex space structure such that h_ϕ is an isomorphism of locally convex spaces for some (and hence all) charts ϕ of M around p.*

Proof (a) Note that h_ϕ and h_ϕ^{-1} are well defined and h_ϕ^{-1} is injective. For $y \in E_\phi$, $h_\phi^{-1} \circ h_\phi(y)) = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \phi(\phi^{-1}(p_\phi + ty)) = y$. Thus h_ϕ^{-1} is surjective and the inverse of h_ϕ.

(b) Compute for $y \in E$: $h_\psi \circ h_\phi^{-1}(y) = \frac{\mathrm{d}}{\mathrm{d}t}\big|_{t=0} \psi(\phi^{-1}(p_\phi + ty)) = d(\psi \circ \phi^{-1})(p_\phi; \cdot)$.

(c) This follows directly from the definition of the vector space structure. □

1.42 Let $U \Subset E$, E a locally convex space and $f \colon U \to F$ a C^1-map. We define the mapping

$$Tf \colon U \times E \to F \times F, \quad (x, v) \mapsto (f(x), df(x; v))$$

and call this mapping the *tangent map* of f. Note that the chain rule, Proposition 1.23, can now be written as $T(f \circ g) = Tf \circ Tg$.

1.43 Definition (Tangent bundle) Let $(M, \mathcal{A})$ be a C^r-manifold with $r \geq 1$. We call $TM := \bigcup_{p \in M} T_pM$ the *tangent bundle* of M. Then $\pi_M \colon TM \to M$,

https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press

$T_pM \ni v \mapsto p$ is called the *bundle projection*. We equip TM with the final topology with respect to the family $(T\phi^{-1})_{\phi\in\mathcal{A}}$ of mappings

$$T\phi^{-1}\colon V_\phi \times E_\phi \to TM, \quad (x,v) \mapsto [t \mapsto \phi^{-1}(x+ty)] \in T_{\phi^{-1}(x)}M.$$

Note that $TU_\phi = \pi_M^{-1}(U_\phi)$ is open in TU for all $\phi \in \mathcal{A}$, and

$$T\phi := (T\phi^{-1})^{-1}\colon TU_\phi \to V_\phi \times E$$

is a homeomorphism. Moreover, $\mathcal{B} := \{T\phi \mid \phi \in \mathcal{A}\}$ is a C^{r-1}-atlas for TM. Thus TM becomes a C^{r-1}-manifold and $\pi_M\colon TM \to M$ a C^{r-1}-map.

1.44 Lemma *We check the details in Definition* 1.43. *Let* $\varphi,\psi \in \mathcal{A}$*:*

(a) *We have* $T(\psi\circ\phi^{-1})\circ T\phi = T\psi$.
(b) TU_φ *is open in* TM *and* $T\varphi\colon TU_\varphi \to V_\varphi \times E_\varphi$ *is a homeomorphism.*
(c) $\mathcal{B} = \{T\phi \mid \phi \in \mathcal{A}\}$ *is a* C^{r-1}*-atlas, if* M *is Hausdorff, so is* TM.
(d) π_M *is a* C^{r-1}*-map.*
(e) TM *induces on each tangent space* T_pM *its natural topology.*

Proof (a) This follows from Lemma 1.41(b); we leave the details to the reader.

(b) If $\varphi,\psi \in \mathcal{A}$, we have $(T\psi^{-1})^{-1}(TU\varphi) = T\psi(TU_\varphi \cap TU_\psi) = \psi(U_\varphi\cap U_\psi)\times E_\psi \Subset V_\psi \times E_\psi$. By the definition of the final topology, TU_φ is open in TM.

By definition of the final topology $T\varphi^{-1}$ is continuous, and so $T\varphi$ is open for every $\varphi \in \mathcal{A}$. For continuity, pick $U \Subset V_\varphi \times E_\varphi$ and let $\psi \in \mathcal{A}$. Now $W := U \cap \varphi(U_\varphi \cap U_\psi)\times E_\varphi \Subset V_\varphi \times E_\varphi$, whence a quick computation shows

$$(T\psi^{-1})^{-1}((T\varphi)^{-1}(U)) = T(\psi\circ\varphi^{-1})(W) \Subset V_\psi \times E_\psi$$

as $T(\psi\circ\varphi^{-1})$ is a homeomorphism between open subsets of $V_\varphi \times E_\varphi$ and $V_\psi \times E_\psi$. We deduce that $(T\varphi)^{-1}(U)$ is open in TM, and so $T\varphi$ is continuous.

(c) By (b) each $T\phi$ is a homeomorphism from an open subset of TM onto an open subset of $E_\phi \times E_\phi$, whence it is a chart. Clearly, the TU_φ cover TM and by (a) the transition maps are $T(\psi\circ\varphi^{-1}) = (\psi\circ\varphi^{-1}, d(\psi\circ\varphi^{-1}))$ whence C^{r-1}. The Hausdorff property is left as Exercise 1.6.1.

(d) In every chart we have $\varphi\circ\pi_M = \mathrm{pr}_1 \circ T\varphi$, where $\mathrm{pr}_1\colon V_\varphi \times E_\varphi \to V_\varphi$ is the canonical projection. As the charts conjugate π_M to a smooth map, it is of class C^{r-1}.

(e) By the definition of the vector topology in Lemma 1.41 (c) this follows from (b). □

1.45 Remark In finite-dimensional differential geometry (and on Banach manifolds), one often introduces the *dual bundle* or *cotangent bundle*

$$T^*M := L(TM,\mathbb{R}) := \bigcup_{x\in M} (T_x M)' = \bigcup_{x\in M} L(T_x M,\mathbb{R}),$$

where $(T_x M)' := L(T_x M,\mathbb{R})$ is the space of continuous linear mappings $T_x M \to \mathbb{R}$ (with the operator norm topology if $T_x M$ is Banach; see Klingenberg, 1995 or Lang, 1999, III.1 for the bundle structure). In our setting, there is, *in general*, no canonical manifold structure which turns T^*M into a vector bundle.

This poses a problem for theories employing the dual bundle, for example, symplectic geometry, and also differential forms cannot be defined as sections of the dual bundle. However, in the case of differential forms, one can circumvent this problem and obtain a theory similar to finite-dimensional differential forms without the dual bundle. We briefly discuss this in Appendix E.

We will now introduce tangent mappings of differentiable mappings between manifolds (see Definition 1.42 for the case of an open subset of a locally convex space).

1.46 Definition (Tangent maps) Let $f\colon M \to N$ be a C^r-map between C^r-manifolds for $r \geq 1$. Define the mappings

$$T_p f\colon T_p M \to T_{f(p)}N, \quad [\gamma] \mapsto [f\circ\gamma], \quad p \in M.$$

Then we define the *tangent map* $Tf\colon TM \to TN$, $T_pM \ni [\gamma] \mapsto T_p f([\gamma])$. Note that by construction $\pi_N \circ Tf = f \circ \pi_M$. Moreover, for each pair of charts ψ of N and ϕ of M such that $f(U_\phi) \subseteq U_\psi$, the following diagram is commutative:

$$\begin{array}{ccc} TU_\phi & \xrightarrow{Tf|_{TU_\phi}^{TU_\psi}} & TU_\psi \\ \downarrow{\scriptstyle T\phi} & & \downarrow{\scriptstyle T\psi} \\ V_\phi \times E_\phi & \xrightarrow{T(\psi\circ f\circ\phi^{-1})} & V_\psi \times F_\psi . \end{array}$$

Hence the tangent map Tf is a C^{r-1}-map if f is a C^r-map.

1.47 Lemma (Chain rule on manifolds) *Let M,N,L be C^r-manifolds and $f\colon M \to N$, $g\colon N \to L$ be C^r-maps with $r \geq 1$. Then $T(g\circ f) = Tg \circ Tf$.*

Note that we can of course iterate the tangent construction and form the higher tangent manifolds $T^k M := \underbrace{(T(T(\cdots(T}_{k \text{ times}} M)\cdots)$ if M is a C^ℓ-manifold and $k \leq \ell$. Similarly one defines higher tangent maps $T^k f := \underbrace{(T(T(\cdots(T}_{k \text{ times}} f)\cdots)$.

https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press

For later use, we set a notation for derivatives of manifold valued curves.

1.48 Definition Let M be a manifold and $c\colon J \to M$ be a C^1-map from some interval $J \subseteq \mathbb{R}$. Then we identify $T_t J = \mathbb{R}$ and define the mapping

$$\dot{c}\colon J \to TM, \quad t \mapsto Tc(t,1).$$

Note that $\dot{c}$ is the manifold version of the curve differential c', in particular, if (U,φ) is a chart of M, we have for each $t \in c^{-1}(U)$ the relation $T\varphi(\dot{c}(t)) = (\varphi \circ c, (\varphi \circ c)')$.

Exercises

1.6.1 Verify the details in Definition 1.46 and supply a proof for Lemma 1.47. Show, in particular, that

(a) Tf is well defined and defines a C^{r-1}-map with the claimed properties.

(b) Check that for $M = U$ an open subset of a locally convex space, both definitions of tangent mappings coincide.

(c) The manifold TM is a Hausdorff topological space.
Hint: Consider two cases for $v, w \in TM$: $\pi_M(v) = \pi_M(w)$ and $\pi_M(v) \neq \pi_M(w)$.

1.6.2 Show inductively, that

(a) If E is a locally convex space and $U \Subset E$, then $T^k U \cong U \times E^{2^k-1}$.

(b) For $U \Subset E, V \Subset F$ in locally convex spaces, we have

$$d^k f(x; v_1, \ldots, v_k) = \mathrm{pr}_{2^k}\left(T^k f(x, w_1, \ldots, w_{2^k-1})\right),$$

where pr_{2^k} is the projection onto the 2^kth component and $w_{2^i+1} = v_{i+1}$ for $0 \leq i \leq k-1$ and $w_i = 0$ else.

1.6.3 Establish a manifold version of the rule on partial differentials Proposition 1.20, that is, show that: If $f\colon M_1 \times M_2 \to N$ is a C^1-map (between C^1-manifolds) and $p_i\colon M_1 \times M_2 \to M_i$ are the canonical projections, then for $p = (x,y) \in M_1 \times M_2$,

$$T_p f(v) = T_x f(\cdot, y)(Tp_1(v)) + T_y f(x, \cdot)(Tp_2(v)).$$

Identifying $T(M_1 \times M_2)$ with $TM_1 \times TM_2$ and $v = (v_x, v_y)$, this formula becomes

$$T_p f = T_x f(\cdot, y)(v_x) + T_y f(x, \cdot)(v_y).$$

https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press

1.7 Elements of Differential Geometry: Submersions and Immersions

Immersions and submersions are among the first tools students encounter in courses on differential geometry when asked to construct (sub-)manifolds. They can still serve this purpose in infinite dimensions if one chooses one's definitions carefully. In this section, we follow Hamilton (1982, Section 4.4)[6] and the discussion requires the concept of complemented subspaces of a locally convex space (see Appendix A.3).

1.49 Definition Let M,N be smooth manifolds and $\phi\colon M \to N$ smooth. We say that ϕ is

(a) an *immersion* if for every $x \in M$ there are manifold charts κ,ψ around x and $\phi(x)$ such that $F_\psi \cong E_\kappa \times H$ (as locally convex spaces; see Appendix A.3), the local representative of ϕ in these charts is the inclusion $E_\kappa \to E_\kappa \times H \cong F_\psi$,

(b) an *embedding* if f is an immersion and a topological embedding (i.e. a homeomorphism onto its image),

(c) a *submersion* if for every $x \in M$ there are manifold charts κ,ψ around x and $\phi(x)$ such that $E_\kappa \cong F_\psi \times H$ and the local representative of ϕ in these charts is the projection $E_\kappa \cong F_\psi \times H \to F_\psi$.

1.50 Lemma *If $f\colon M \to N$ is an immersion, then for every $x \in M$ there is an open neighbourhood W_x such that $f|_{W_x}$ is an embedding.*

Proof Pick immersion charts around x and $f(x)$, that is, charts $(U,\varphi), x \in U$ and (W,ψ) such that $f(x) \in W$ and $\psi \circ f \circ \varphi^{-1} = j$, where $j\colon E_\varphi \to E_\psi \cong E_\varphi \times F$ is the inclusion of the complemented subspace E_φ of E_ψ. Note that j is a topological embedding onto its image, whence $f|_{U_\varphi} = \psi^{-1} \circ j \circ \varphi$ is a topological embedding (and thus an embedding). □

There are several alternative characterisations of submersions and immersions known in the finite-dimensional setting. Many of these turn out to be weaker in our setting.

1.51 Lemma *A smooth map $f\colon M \to N$ is a submersion if and only if for each $x \in M$ there are (U,φ) of M and (V,ψ) of N with $x \in U$ and $f(U) \subseteq V$ such that $\psi \circ f \circ \varphi^{-1} = \pi|_{\varphi(U)}$ for a continuous linear map π with continuous linear right inverse σ (i.e. $\pi \circ \sigma = \mathrm{id}$).*

[6] The smoothness assumption conveniently shortens the exposition but can of course be replaced by finite orders of differentiability; see Glöckner (2016).

https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press

Proof If f is a submersion and $x \in M$, we can pick submersion charts around x, that is, charts $\varphi\colon U_\varphi \to U_F \times U_H \subseteq F \times H \cong E$ with $U_F \Subset F$ and $U_H \Subset E$ and $\psi\colon V_\psi \to U_F$ such that $\psi \circ f \circ \varphi^{-1}(x,y) = x$, for all $(x,y) \in U_F \times U_H$. Obviously, the projection is continuous linear with right inverse given by the inclusion $F \to F \times H \cong E$.

Let us conversely assume that around $x \in M$ there are charts such that $\psi \circ f \circ \varphi^{-1} = \pi|_{\varphi(U)}$ holds for a continuous linear map $\pi\colon E \to F$ with continuous right inverse $\sigma\colon F \to E$. Then $\pi|_{\sigma(F)}\sigma(F) \to F$ is an isomorphism of locally convex spaces with inverse σ. We obtain a new chart $\nu := \sigma \circ \psi$ of N. Now $\kappa := \sigma \circ \pi\colon E \to E$ is continuous linear and satisfies $\kappa \circ \kappa = \kappa$. We deduce from Lemma A.21 that $\sigma(F)$ is a complemented subspace of E and can identify $E \cong \sigma(F) \times \ker(\pi)$ such that κ becomes the projection onto $\sigma(F)$. Shrinking U and V if necessary we may assume $\varphi(U) = A \times B$ and $\nu(V) = A$ for open sets $A \Subset \sigma(F)$ and $B \Subset \ker\pi$. Then

$$\nu \circ f \circ \varphi^{-1} = \sigma \circ \psi \circ f \circ \varphi^{-1} = \sigma \circ \pi|_{A\times B} = \kappa|_{A\times B}. \qquad \square$$

1.52 Remark The alternative characterisation of submersions from Lemma 1.51 implies (see Exercise 1.7.5) that a submersion admits local smooth sections. If the manifolds M and N are Banach manifolds (i.e. modelled on Banach spaces), the inverse function theorem implies that the existence of local smooth sections is even equivalent to being a submersion (see Margalef-Roig and Domínguez, 1992, Proposition 4.1.13).

Finally, we mention that *if the submersion φ is surjective*, we obtain the following result: *a map $f\colon N \to L$ is C^r if and only if $f \circ \varphi$ is a C^r-map* for $r \in \mathbb{N}_0 \cup \{\infty\}$ (see Exercise 1.7.6).

In finite-dimensional differential geometry, the above definitions are usually not the definitions of submersions/immersions but one deduces them from ‘easier conditions’ involving the tangent mappings, such as the following.

1.53 Definition Let M, N be smooth manifolds and $\phi\colon M \to N$ smooth. We say that ϕ is

(a) *infinitesimally injective (or surjective)* if the tangent map $T_x\phi\colon T_xM \to T_{\phi(x)}N$ is injective (or surjective, respectively) for every $x \in M$,
(b) a *naïve immersion* if for every $x \in M$ the tangent map $T_x\phi\colon T_xM \to T_{\phi(x)}N$ is a topological embedding onto a complemented subspace of $T_{\phi(x)}N$,
(c) a *naïve submersion* if for every $x \in M$ the map $T_x\phi\colon T_xM \to T_{\phi(x)}N$ has a continuous linear right inverse.

In infinite dimensions, none of the naïve, infinitesimal versions or the properties from Definition 1.49 are equivalent as the following (counter-)examples show.

https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press

1.54 (Infinitesimal properties are weaker than naïve properties) Consider the Banach space c_0 of all (real) sequences converging to 0 as a subspace of the Banach space ℓ^∞ of all bounded real sequences. We obtain a *short exact sequence*[7] of Banach spaces

$$0 \longrightarrow c_0 \xhookrightarrow{\quad i \quad} \ell^\infty \xrightarrow{\quad q \quad} \ell^\infty/c_0 \longrightarrow 0. \tag{1.9}$$

As i and q are continuous linear we have $T_x i(y) = di(x,y) = i(y)$ and $T_x q(y) = q(y)$, whence i is infinitesimally injective and q is infinitesimally surjective. Since c_0 is not complemented, Example A.23, i is not a naïve immersion. This implies that (1.9) does not split, that is, q does not admit a continuous linear right inverse and thus cannot be a naïve submersion.

A submersion as in Definition 1.49 turns out to be an open map (Glöckner, 2016, Lemma 1.7). In finite-dimensional differential geometry, this is a consequence of the inverse function theorem (when applied to a naïve submersion). Going beyond Banach manifolds, the submersion property is stronger than the naïve notion.

1.55 (Naïve submersions need not be submersions) Consider the space $A := C(\mathbb{R},\mathbb{R})$ of continuous functions from the reals to the reals. The pointwise operations and the compact open topology (see Appendix B.2) turn A into a locally convex space (by Lemma B.7, one can even show that it is a Fréchet space). Then the map $\exp_A : A \to A,\ f \mapsto e^f$ is continuous by Lemma B.8. Recall that the point evaluations $\mathrm{ev}_x(f) := f(x)$ are continuous linear on A, whence we can use them to find a candidate for the derivative of $\exp_A$. Then the finite-dimensional chain rule yields the only candidate for the derivative of $\exp_A$ to be $d\exp_A(f;g)(x) = g(x)\cdot\exp_A(f)(x)$. Computing in the seminorms, one can show that indeed $d\exp_A(f;g) = g\cdot\exp_A(f)$ and by induction $\exp_A$ is smooth. Moreover, at the constant zero function $\mathbf{0}$ we have $d\exp_A(\mathbf{0};\cdot) = \mathrm{id}_A$, so $\exp_A$ is a naïve submersion. However, $\exp_A$ takes values in $C(\mathbb{R},]0,\infty[)$ which do not contain a neighbourhood of $\exp_A(\mathbf{0}) = \mathbf{1}$. Thus $\exp_A$ cannot be a submersion (as all submersions are open mappings by Glöckner (2016, Lemma 1.7).

[7] In the category of locally convex spaces, a sequence

$$0 \longrightarrow A \xrightarrow{i} B \xrightarrow{q} C \longrightarrow 0$$

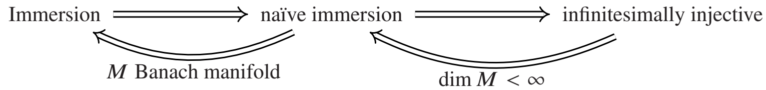
of continuous linear maps is *exact* if it satisfies both of the following conditions:
(a) *algebraically exact*, that is, images of maps coincide with kernels of the next map;
(b) *topologically exact*, that is, i and q are open mappings onto their images.
If A, B and C are Fréchet (or Banach) spaces, topological exactness follows from algebraic exactness by virtue of the open mapping theorem (Rudin, 1991, I. 2.11); for general locally convex spaces this is not the case.

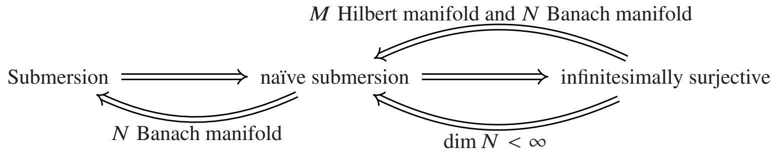
https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press

This example also shows that the Inverse Function Theorem fails in this case (Eells, 1966); see also Appendix A.5.

However, as Glöckner (2016) shows, the following relations do hold.

1.56 (Submersions, immersions vs. the naïve and infinitesimal concepts) Consider a map $\phi\colon M \to N$ between manifolds modelled on locally convex spaces. Then:





The reason one really would like the strong notions of submersions and immersions is that these notions are strong enough to carry over the usual statements on submersions and immersions to the setting of infinite-dimensional manifolds. For example, one can prove several useful statements on split submanifolds. Again we refer to Glöckner (2016) for more general results on submersions and immersions in infinite dimensions.

1.57 Definition Let $f\colon M \to N$ be smooth and $S \subseteq N$ be a split submanifold. Then f is *transversal over* S if for each $m \in f^{-1}(S)$ and submanifold chart $\psi\colon V \to V_1 \times V_2$ with $\psi(f(m)) = (0,0)$ and $\psi(S) \subseteq V_1 \times \{0\}$, there exists an open m-neighbourhood U with $f(U) \subseteq V$ and

$$U \xrightarrow{f} V \xrightarrow{\psi} V_1 \times V_2 \xrightarrow{\mathrm{pr}_2} V_2 \tag{1.10}$$

is a submersion.

Note that due to the fact that compositions of submersions are again submersions (see Exercise 1.7.1), if f is a submersion, (1.10) is always a submersion.

1.58 Proposition *Let $\phi\colon M \to N$ be a smooth map. If $S \subseteq N$ is a split submanifold*[8] *of N such that f is transversal over S, then $\phi^{-1}(S)$ is a submanifold of M.*

[8] A more involved proof works for every submanifold (not only for split ones), see Glöckner (2016, Theorem C).

https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press

Proof The map (1.10) is a submersion. Shrinking U, V_2, there are charts $\varphi \colon U \to U_1 \times U_2$ and $\kappa \colon V_2 \to U_2$ such that $\varphi(m) = (0,0)$ and $\kappa(0) = 0$ and the following commutes:

$$\begin{array}{ccccccc} U & \xrightarrow{f} & V & \xrightarrow{\psi} & V_1 \times V_2 & \xrightarrow{\mathrm{pr}_2} & V_2 \\ \downarrow{\scriptstyle \varphi} & & & & & & \downarrow{\scriptstyle \kappa} \\ U_1 \times U_2 & & & \xrightarrow{\mathrm{pr}_2} & & & U_2. \end{array}$$

Now we will prove that φ is a submanifold chart for $f^{-1}(S)$, that is, $\varphi(U \cap f^{-1}(S)) = \varphi(U) \cap (U_1 \times \{0\})$. To see this, note that since ψ is a submanifold chart, we have for $x \in U$ that $f(x) \in S$ if and only if $\mathrm{pr}_2(\psi(f(x))) = 0$. Now the commutativity of the diagram shows that this is the case if and only if $\varphi(x) \in \mathrm{pr}_2^{-1}(0) = U_1 \times \{0\}$. □

1.59 Corollary *If $f \colon M \to N$ is a submersion, $f^{-1}(n)$ is a split submanifold for $n \in N$.*

1.60 Lemma *Let $f \colon M \to P$ and $g \colon N \to P$ be smooth maps and g be a submersion. Then the* fibre product $M \times_P N := \{(m,n) \in M \times N \mid f(m) = g(p)\}$ *is a split submanifold of $M \times N$ and the projection* $\mathrm{pr}_1 \colon M \times_P N \to M$ *is a submersion.*

Proof Let $(m,n) \in M \times_P N$ and pick submersion charts $\psi \colon U_\psi \to V_\psi \Subset E_N$, $\kappa \colon U_\kappa \to V_\kappa \Subset E_P$ for g with $n \in U_\psi$. Recall that for the submersion charts $\kappa \circ g \circ \psi^{-1} = \pi$ for a continuous projection $\pi \colon E_N = E_P \times F \to E_P$ (where F is the subspace complement) and we may assume that $V_\psi = V_\kappa \times \tilde{V}$. Finally, we pick a chart (U_φ, φ) of M such that $m \in U_\varphi$ and $f(U_\varphi) \subseteq U_\kappa$. Hence we obtain a commutative diagram

$$\begin{array}{ccccccc} & & U_\psi & \xrightarrow{\psi} & V_\psi & = & V_\kappa \times \tilde{V} \\ & & \downarrow{\scriptstyle g} & & & \searrow{\scriptstyle \pi} & \downarrow{\scriptstyle \mathrm{pr}_{V_\kappa}} \\ U_\varphi & \xrightarrow{f} & U_\kappa & & \xrightarrow{\kappa} & & V_\kappa. \end{array}$$

Denote by $\mathrm{pr}_{\tilde{V}}$ the projection onto $\tilde{V}$. Then we construct a smooth map for $(m,n) \in U_\varphi \times U_\psi$ via

$$\delta(m,n) := (\varphi(m), (\mathrm{pr}_{V_\kappa}(\psi(n)) - \kappa(f(m)), \mathrm{pr}_{\tilde{V}}(\psi(n)))) \in V_\varphi \times ((V_\kappa - V_\kappa) \times \tilde{V}).$$

This mapping is smoothly invertible with inverse given by

$$\delta^{-1}(x,(y,z)) = (\varphi^{-1}(x), \psi^{-1}(y + \kappa(f(\varphi^{-1}(x))), z)).$$

We leave it again as an exercise to work out that the domain of the inverse is open. Note that due to the commutative diagram we see that $(m,n) \in M \times_P N$ if and only if (m,n) maps under δ to $E_M \times \{0\} \times F$ which is a complemented subspace of $E_M \times E_P \times F \cong E_M \times E_N$. Thus $M \times_P N$ is a split submanifold of $M \times N$. Since the projection $\mathrm{pr}_1 : M \times N \to N$ is smooth, so is its restriction to $M \times_P N$ by Lemma 1.39. As $\mathrm{pr}_1 : M \times_P N \to M$ is conjugated by δ and φ to the projection pr_{V_κ}, it is a submersion. □

Finally, there is a close connection between embeddings and split submanifolds.

1.61 Lemma *Let $f : M \to N$ be smooth. The following conditions are equivalent:*

(a) *f is an embedding.*
(b) *$f(M)$ is a split submanifold of N and $f|^{f(M)} : M \to f(M)$ is a diffeomorphism.*

Proof Let E, F be the modelling spaces of M and N, respectively.

(a) ⇒ (b): By assumption, f is a topological embedding and a smooth immersion. Consider $y \in f(M)$ and $x \in M$ with $f(x) = y$ and pick charts $\varphi_x : U_x \to V_x \subseteq E$ and $\varphi_y : U_y \to V_y \subseteq F$ such that $x \in U_x$, $f(U_x) \subseteq U_y$ and $\varphi_y \circ f \circ \varphi_x^{-1} = j|_{V_x}$ for a linear topological embedding $j : E \to F$ onto a complemented subspace $j(E) \times H = F$. Since $j(V_x)$ is relatively open, we may adjust choices such that $j(V_x) = V_y \cap j(E)$. A quick computation then shows that φ_y restricts to a submanifold chart and $f(M)$ becomes a split submanifold of N. Moreover, in the (sub)manifold charts we have $j|_{V_x}^{V_y \cap j(E)} = \varphi_y|_{f(M) \cap U_y} \circ f \circ \varphi_x^{-1}$ and this map is a diffeomorphism. Thus $f|^{f(M)} : M \to f(M)$ is a local diffeomorphism and a homeomorphism, hence a diffeomorphism.

(b) ⇒ (a) Let $\iota : f(M) \to N$ be the inclusion map. Since $f|^{f(M)}$ is a diffeomorphism, $\iota \circ f|^{f(M)}$ is a topological embedding. Since $f(M)$ is a split submanifold, there is an isomorphism $\alpha : E \to \alpha(E) \subseteq F$ of locally convex spaces such that $\alpha(E)$ is complemented in F. Pick charts $\varphi_x : U_x \to V_x$ and $\varphi_{f(x)} : U_{f(x)} \to V_{f(x)}$ with $x \in U_x$ and $f(U_x) \subseteq U_{f(x)}$. We may assume that $V_{f(x)} = P \times Q$ for $P \Subset \alpha(E)$ and $\varphi_{f(x)}(U_x \cap f(M)) = V_{f(x)} \cap \alpha(E) = P$. Set now $W := \alpha^{-1}(P)$. Then it is easy to see that $\theta := (\varphi_{f(x)} \circ f \circ \varphi_x^{-1})^{-1} \circ \alpha|_W : W \to V_x$ makes sense and is a smooth diffeomorphism, whence $\theta^{-1} \circ \varphi_x : U_x \to W$ is a chart for M. By construction, $\varphi_{f(x)} \circ f \circ (\theta^{-1} \circ \varphi_x)^{-1} = \alpha|_W$ is a linear topological embedding onto $\alpha(E)$. This shows that f is an immersion. □

1.7.1 Exercises

1.7.1 Show that if $f\colon M \to N$ and $g\colon N \to L$ are submersions, so is $g \circ f\colon M \to L$.
Hint: The submersion property is local. Try constructing small enough neighbourhoods around $m \in U \subseteq M, f(m) \in V \subseteq N$ and $g(f(m)) \in W \subseteq L$ and charts such that:

$$\begin{array}{ccccc} U & \xrightarrow{\quad f \quad} & V & \xrightarrow{\quad g \quad} & W \\ \downarrow & & \downarrow & & \downarrow \\ F \times X & \xrightarrow{\mathrm{pr}_F} & F \cong Y \times Z & \xrightarrow{\mathrm{pr}_Y} & Y. \end{array}$$

1.7.2 Let $f\colon M \to N$ and $g\colon K \to L$ be submersions (immersions). Show that then $f \times g\colon M \times K \to N \times L$, $(m,k) \mapsto (f(m), g(k))$ is also a submersion (immersion).

1.7.3 Let $p\colon M \to N$ be a submersion and $n \in N$. From Corollary 1.59 we obtain a submanifold $P := p^{-1}(n)$. Show that for $x \in P$ one can identify the tangent space of the submanifold as $T_x P = \ker T_x p = \{v \in T_x M \mid T_x p(v) = 0\}$.

1.7.4 Let $(H, \langle \cdot, \cdot \rangle)$ be a Hilbert space. Prove that $\psi\colon H \setminus \{0\} \to \mathbb{R}$, $x \mapsto \langle x, x \rangle$ is a submersion and deduce that the Hilbert sphere $S_H = \psi^{-1}(1)$ is a submanifold of H. Then show that $T_x S_H = \{v \in H \mid \langle v, x \rangle = 0\}$ for all $x \in S_H$ (that is, the tangent space is the orthogonal complement to the base point x).

1.7.5 Let $\varphi\colon M \to N$ be a smooth submersion. Show that φ admits smooth local sections, that is, for every $x \in M$ there is $\varphi(x) \in U \Subset N$ and a smooth map $\sigma\colon U \to M$ with $\sigma(\varphi(x)) = x$ and $\varphi \circ \sigma = \mathrm{id}_U$. Deduce then that φ is an open map.
Hint: Use the characterisation from Remark 1.52.
Remark: If M, N are Banach manifolds, the existence of local sections is equivalent to φ being a submersion; see Margalef-Roig and Domínguez (1992, Proposition 4.1.13).

1.7.6 Let $\varphi\colon M \to N$ be a smooth surjective submersion. Show that $f\colon N \to L$ is C^r if and only if $f \circ \varphi$ is C^r for $r \in \mathbb{N}_0 \cup \{\infty\}$.
Hint: Use Exercise 1.7.5.

1.7.7 Show that if $f\colon M \to N$ and $g\colon N \to L$ are immersions (embeddings), so is $g \circ f$.

1.7.8 Work out the details omitted in the proof of Lemma 1.61.

https://doi.org/10.1017/9781009091251.002 Published online by Cambridge University Press