Handbook of Global Analysis
Demeter Krupka and David Saunders

Some aspects of differential theories

Respectfully dedicated to the memory of Serge Lang

József Szilasi and Rezső L. Lovas

Contents

Introduction

Global analysis is a particular amalgamation of topology, geometry and analysis, with strong physical motivations in its roots, its development and its perspectives. To formulate a problem in global analysis exactly, we need

(1) a base manifold,

(2) suitable fibre bundles over the base manifold,

(3) a differential operator between the topological vector spaces consisting of the sections of the chosen bundles.

In quantum physical applications, the base manifold plays the role of space-time; its points represent the location of the particles. The particles obey the laws of quantum physics, which are encoded in the vector space structure attached to the particles in the mathematical model. A particle carries this vector space structure with itself as it moves. Thus we arrive at the intuitive notion of vector bundle, which had arisen as the 'repère mobile' in Élie Cartan's works a few years *before* quantum theory was discovered.

To describe a system (e.g. in quantum physics) in the geometric framework of vector bundles effectively, we need a suitably flexible differential calculus. We have to differentiate vectors which change smoothly together with the vector space carrying them. The primitive idea of differentiating the coordinates of the vector in a fixed basis is obviously not satisfactory, since there is no intrinsic coordinate system which could guarantee the invariance of the results. These difficulties were traditionally solved by classical tensor calculus, whose most sophisticated version can be found in Schouten's 'Ricci-Calculus'.

From there, the theory could only develop to the global direction, from the debauch of indices to totally index-free calculus. This may be well illustrated by A. Nijenhuis' activity.

It had become essentially clear by the second decade of the last century that, for a coordinate-invariant tensorial differential calculus, we need a structure establishing an isomorphism between vector spaces at different points. Such a structure is called a *connection*. A connection was first constructed by Levi-Civita in the framework of Riemannian geometry by defining a *parallel transport* between the tangent spaces at two points of the base manifold along a smooth arc connecting the given points. This makes it possible to form a difference quotient and to differentiate vector and tensor fields along the curve. The differentiation procedure so defined is *covariant differentiation*. Thus it also becomes clear that

connection, *parallel transport*, *covariant differentiation*

are essentially equivalent notions: these are the same object, from different points of view.

The history sketched in the foregoing is of course well-known, and technical details are nowadays available in dozens of excellent monographs and textbooks. Somewhat paradoxically, *purely historical* features (the exact original sources of main ideas, the evolution of main streamlines) are not clear in every detail, and they would be worth of a more profound study. In our present work, we would like to sketch some aspects of the rich theory of differentiation on manifolds which are less traditional and less known, but which definitely seem to be progressive. One of the powerful trends nowadays is the globalization of calculus on infinite-dimensional non-Banach topological vector spaces, i.e., its transplantation onto manifolds, and a formulation of a corresponding Lie theory. The spectrum becomes more colourful (and the theory less transparent) by the circumstance that we see contesting calculi even on a local level. One corner stone in this direction is definitely A. Kriegl and P. W. Michor's truly monumental monograph [18], which establishes the theory on calculus in so-called convenient vector spaces. Differentiation theory in convenient vector spaces is outlined in the contribution of J. Margalef-Roig and E. Outerelo Domínguez in this volume. Another promising approach is Michel – Bastiani differential calculus, which became known mainly due to J. Milnor and R. Hamilton. We have chosen this way, drawing much from H. Glöckner and K.-H. Neeb's monograph is preparation, which will contain a thorough and exhaustive account of this calculus.

We would like to emphasize that the fundamental notions and techniques of 'infinite dimensional analysis' can be spared neither by those who study analysis 'only' on finite-dimensional manifolds. The reason for this is very simple: even the vector space of smooth functions on an open subset of $\mathbb{R}^n$ is infinite-dimensional, and the most natural structure with which it may be endowed is a suitable multinorm which makes it a Fréchet space. Accordingly, we begin our treatment by a review of some basic notions and facts in connection with topological vector spaces. The presentation is organized in such a way that we may provide a non-trivial application by a detailed proof of Peetre's theorem. This famous theorem characterizes linear differential operators as support-decreasing $\mathbb{R}$-linear maps, and it is of course well-known, together its various proofs. The source of one of the standard proofs is Narasimhan's book [29]; Helgason takes over essentially his proof [15]. Although Narasimhan's proof is very clear in its main features, our experience is that understanding it in every fine detail requires serious intellectual efforts. Therefore we thought that a detailed treatment of Narasimhan's line of thought could be useful by filling in the wider logical gaps. Thus, besides presenting fundamental techniques, we shall also have a possibility to demonstrate hidden subtleties of these sophisticated (at first sight mysterious) constructions.

We discuss Peetre's theorem in a local framework, its transplantation to vector bundles, however, does not raise any difficulty. After climbing this first peak, we sketch the main steps towards the globalization of Michel – Bastiani differential calculus, on the level of basic notions and constructions, on a manifold modeled on a locally convex topological vector space.

In the last section, for simplicity's sake, we return to finite dimension, and we discuss a special problem about covariant derivatives. Due to Élie Cartan's activity, it has been known since the 1930s that a covariant derivative compatible with a Finsler structure cannot be constructed on the base manifold (more precisely, on its tangent bundle), and the velocity-dependent character of the objects makes it necessary to start from a so-called line element bundle. In a contemporary language: the introduction of a covariant derivative, analogous to Levi-Civita's, metrical with respect to the Finsler structure, is only possible in the pull-back of the tangent bundle over itself, or in some 'equivalent' fibre bundle. We have to note that it was a long and tedious way from Élie Cartan's intuitively very clear but conceptually rather obscure construction to today's strict formulations, and this way was paved, to a great extent, by the demand for understanding Élie Cartan. In the meantime, in 1943, another 'half-metrical' covariant derivative in Finsler geometry was discovered by S. S. Chern (rediscovered by Hanno Rund in 1951). This covariant derivative was essentially in a state of suspended animation until the 1990s. Then, however, came a turning point. Due to Chern's renewed activity and D. Bao and Z. Shen's work, Chern's covariant derivative became one of the most important tools of those working in this field. We wanted to understand which were the properties of Chern's derivative which could give priority to it over other covariant derivatives used in Finsler geometry (if there are such properties). As we have mentioned, Chern's connection is only half-metrical: covariant derivatives of the metric tensor arising from the Finsler structure in vertical directions do not vanish in general. 'In return', however, the derivative is 'vertically natural': it induces the natural parallelism of vector spaces on the fibres. This property makes it possible to interpret Chern's derivative as a covariant derivative *given on the base manifold*, parametrized locally by a nowhere vanishing vector field. We think that this possibility of interpretation does distinguish Chern's derivative in some sense. As for genuine applications of Chern's connection, we refer to T. Aikou and L. Kozma's study in this volume.

Notation

As usual, $\mathbb{R}$ and $\mathbb{C}$ will denote the fields of real and complex numbers, respectively. $\mathbb{N}$ stands for the 'half-ring' of natural numbers (integers $\geqq 0$). If $\mathbb{K}$ if one of these number systems, then $\mathbb{K}^* := \mathbb{K} \setminus \{0\}$. $\mathbb{R}_+ := \{z \in \mathbb{R} | z \geqq 0\}$, $\mathbb{R}_+^* := \mathbb{R}_+ \cap \mathbb{R}^*$. A mapping from a set into $\mathbb{R}$ or $\mathbb{C}$ will be called a *function*.

Discussing functions defined on a subset of $\mathbb{R}^n$, it will be convenient to use the multi-index notation $\alpha = (\alpha_1, \dots, \alpha_n)$, where $\alpha_i \in \mathbb{N}$, $1 \leqq i \leqq n$. We agree that

$$|\alpha| := \alpha_1 + \dots + \alpha_n, \quad \alpha! := \alpha_1! \dots \alpha_n!.$$

If $\beta := (\beta_1, \dots, \beta_n) \in \mathbb{N}^n$, and $\beta_j \leqq \alpha_j$ for all $j \in \{1, \dots, n\}$, we write $\beta \leqq \alpha$ and define

$$\alpha - \beta := (\alpha_1 - \beta_1, \ldots, \alpha_n - \beta_n), \quad \binom{\alpha}{\beta} := \frac{\alpha!}{(\alpha-\beta)!\beta!}.$$

If V and W are vector spaces over a field $\mathbb{F}$, then $\mathcal{L}_{\mathbb{F}}(V,W)$ or simply $\mathcal{L}(V,W)$ denotes the vector space of all linear mappings from V into W, and $V^* := \mathcal{L}(V,\mathbb{F})$ is the (algebraic) dual of V. If $k \in \mathbb{N}^*$, $\mathcal{L}^k(V,W)$ is the vector space of all k-multilinear mappings $\varphi : V \times \ldots \times V \to W$. We note that $\mathcal{L}\left(V, \mathcal{L}^k(V,W)\right)$ is canonically isomorphic to $\mathcal{L}^{k+1}(V,W)$. We use an analogous notation if, more generally, V and W are modules over the same commutative ring.

1 Background

Topology

We assume the reader is familiar with the rudiments of point set topology, so the meaning of such elementary terms as open and closed set, neighbourhood, connectedness, Hausdorff topology, (open) covering, first and second countability, compactness and local compactness, continuity, homeomorphism, ... does not demand an explanation. For the sake of definiteness, we are going to follow, as closely as feasible, the convention of Dugundji's Topology [7]. Thus by a *neighbourhood* of a point or a set in a topological space we shall always mean an *open* subset containing the point or subset, and we include in the definition of second countability, compactness and local compactness the requirement that the topology is Hausdorff. This subsection serves to fix basic terminology and notation, as well as to collect some more subtle topological ideas which may be beyond the usual knowledge of non-specialists.

We denote by $\mathcal{N}(p)$ the set of all neighbourhoods of a point p in a topological space. A subset $\mathcal{F} \subset \mathcal{N}(p)$ is said to be a *fundamental system* of p if for every $\mathcal{V} \in \mathcal{N}(p)$ there exists $\mathcal{U} \in \mathcal{F}$ such that $\mathcal{U} \subset \mathcal{V}$. If S is a topological space, and $A \subset S$, then $\overset{\circ}{A}$, $\bar{A}$ and ∂A denote the interior, the closure and the boundary of A, resp. If G is an Abelian group, and $f : S \to G$ is a mapping, then

$$\operatorname{supp}(f) := \overline{\{p \in S | f(p) \neq 0\}}$$

is the *support* of f.

Let M be a metric space with distance function $\varrho : M \times M \to \mathbb{R}$. If $a \in M$, and $r \in \mathbb{R}_+^*$, the set

$$B_r^{\varrho}(a) := \{p \in M | \varrho(a,p) < r\}$$

is called the *open ϱ-ball* of centre a and radius r. (We shall omit the distinguishing ϱ whenever the distance function is clear from the context.) By declaring a subset of M to be open if it is a (possibly empty) union of open balls, a topology is obtained on M, called the *metric topology* of M, or the topology *induced by the distance function* ϱ. Unless otherwise stated, a metric space will always be topologized by its metric topology. Conversely, if a topology $\mathcal{T}$ on a set S is induced by a distance function $\varrho : S \times S \to \mathbb{R}$, then $\mathcal{T}$ and ϱ are called *compatible*, and $\mathcal{T}$ is said to be *metrizable*.

By far the most important metric space is the *Euclidean n-space* $\mathbb{R}^n$, the set of all n-tuples $v = \left(\nu^1, \dots, \nu^n\right)$ endowed with the *Euclidean distance* defined by

$$\varrho_E(a, b) := \|a - b\| = \langle a - b, a - b\rangle^{1/2} \quad (a, b \in \mathbb{R}^n),$$

where $\langle\ ,\ \rangle$ is the canonical scalar product in $\mathbb{R}^n$, and $\|\cdot\|$ is *Euclidean norm* arising from $\langle\ ,\ \rangle$. The metric topology of $\mathbb{R}^n$ is called the *Euclidean topology* of $\mathbb{R}^n$. We shall assume that $\mathbb{R}^n$ (in particular, $\mathbb{R} = \mathbb{R}^1$) is topologized with the Euclidean topology.

Let $\mathcal{U}$ be an open subset of $\mathbb{R}^n$. A sequence $(K_i)_{i\in\mathbb{N}^*}$ of compact subsets of $\mathcal{U}$ is said to be a *compact exhaustion* of $\mathcal{U}$ if

$$K_i \subset \overset{\circ}{K}_{i+1} \quad (i \in \mathbb{N}^*) \quad \text{and} \quad \mathcal{U} = \bigcup_{i\in\mathbb{N}^*} K_i.$$

Lemma 1.1 *There does exist a compact exhaustion for any open subset of* $\mathbb{R}^n$.

Proof. Let a (nonempty) open subset $\mathcal{U}$ of $\mathbb{R}^n$ be given. For every positive integer m, define a subset K_m of $\mathcal{U}$ as follows:

$$K_m := \left\{ p \in \mathcal{U} \,\middle|\, \varrho_E(p, \partial\mathcal{U}) \geqq \frac{1}{m} \right\} \cap \overline{B_m(0)}$$

(by convention, $\varrho_E(p, \partial\mathcal{U}) := \infty$ if $\partial\mathcal{U} = \emptyset$). It is then clear that each set K_m is compact, $K_m \subset \overset{\circ}{K}_{m+1}$ ($m \in \mathbb{N}^*$), and $\cup_{i\in\mathbb{N}^*} K_i = \mathcal{U}$. □

Now we recall some more delicate concepts and facts of point set topology.

A Hausdorff space is said to be a *Lindelöf space* if each open covering of the space contains a *countable* covering. By a theorem of Lindelöf [7, Ch. VIII, 6.3] *all second countable spaces are Lindelöf* (but the converse is not true!).

A locally compact space is called *σ-compact* if it can be expressed as the union of a sequence of its compact subsets. It can be shown (see [7, Ch. XI, 7.2]) that *a topological space is σ-compact, if and only if, it is a locally compact Lindelöf space*.

Let S be a topological space. An open covering $(\mathcal{U}_\alpha)_{\alpha\in A}$ of S is said to be *locally finite* (or *nbd-finite*) if each point of S has a nbd $\mathcal{U}$ such that $\mathcal{U} \cap \mathcal{U}_\alpha \neq \emptyset$ for at most finitely many indices α. If $(\mathcal{U}_\alpha)_{\alpha\in A}$ and $(\mathcal{V}_\beta)_{\beta\in B}$ are two coverings of S, then $(\mathcal{U}_\alpha)$ is a *refinement* of $(\mathcal{V}_\beta)$ if for each $\alpha \in A$ there is some $\beta \in B$ such that $\mathcal{U}_\alpha \subset \mathcal{V}_\beta$. A Hausdorff space is said to be *paracompact* if each open covering of the space has an open locally finite refinement.

The following result has important applications in analysis.

Lemma 1.2 *Any σ-compact topological space, in particular any second countable locally compact topological space, is paracompact.*

Proof. The statement is a fairly immediate consequence of some general topological facts. Namely, the σ-compact spaces, as we have remarked above, are just the locally compact Lindelöf spaces. Locally compact spaces are regular: each point and closed set not containing the point have disjoint nbds. Since by a theorem of K. Morita [7, VIII, 6.5] *in Lindelöf spaces regularity and paracompactness are equivalent concepts*, we get the result. □

It is well-known that the classical concept of convergence of sequences is not sufficient for the purposes of analysis. Nets, which are generalizations of sequences, and their convergence provide an efficient tool to handle a wider class of problems. To define nets, let us first recall that a *directed set* is a set A endowed with a partial ordering $\leqq$ such that for any two elements $\alpha, \beta \in A$ there is an element $\gamma \in A$ such that $\gamma \geqq \alpha$ and $\gamma \geqq \beta$. A *net* in a set S is a family $(s_\alpha)_{\alpha\in A}$ of elements of S, i.e., a mapping $s : A \to S$, $\alpha \mapsto s(\alpha) =: s_\alpha$, where A is a directed set. Obviously, any sequence $s : n \in \mathbb{N} \mapsto s(n) =: s_n \in S$ is a net, since $\mathbb{N}$ is a directed set with its usual ordering.

A net $(s_\alpha)_{\alpha\in A}$ in a *topological space* S is said to *converge* (or to *tend*) to a point $p \in S$ if for every $\mathcal{U} \in \mathcal{N}(p)$ there is a $\beta \in A$ such that

$$s_\alpha \in \mathcal{U} \textit{ whenever } \alpha \geqq \beta.$$

Then we use the standard notation $p = \lim\limits_{\alpha\in A} s_\alpha$, and we say that $(s_\alpha)_{\alpha\in A}$ is *convergent* in S and has a *limit* $p \in S$.

Lemma 1.3 *A topological space is Hausdorff if and only if any two limits of any convergent net are equal.*

Indication of proof. Let S be a topological space. It can immediately be seen that if S is Hausdorff, then any convergent net in S has a unique limit. Conversely, suppose that S is not Hausdorff. Let $p, q \in S$ be two points which cannot be separated by open sets. Consider the directed set A whose elements are ordered pairs $\alpha = (\mathcal{U}, \mathcal{V})$ where $\mathcal{U} \in \mathcal{N}(p)$, $\mathcal{V} \in \mathcal{N}(q)$ with the partial ordering

$$(\mathcal{U}, \mathcal{V}) \geqq (\mathcal{U}_1, \mathcal{V}_1) :\Longleftrightarrow (\mathcal{U} \subset \mathcal{U}_1 \text{ and } \mathcal{V} \subset \mathcal{V}_1).$$

For any $\alpha = (\mathcal{U}, \mathcal{V})$ let s_α be some point of $\mathcal{U} \cap \mathcal{V}$. Then the mapping $s : \alpha \in A \mapsto s_\alpha \in S$ is a net, and it is easy to check that s converges to both p and q. $\square$

The next result shows that nets are sufficient to control continuity.

Lemma 1.4 *A mapping $\varphi : S \to T$ between two topological spaces is continuous if and only if for every net $s : A \to S$ converging to $p \in S$, the net $\varphi \circ s : A \to T$ converges to $\varphi(p)$.*

For a (quite immediate) proof see e.g. [5, Ch. I, 6.6].

If S and T are topological spaces, then the set of continuous mappings of S into T will be denoted by $C(S, T)$. In particular, $C(S) = C^0(S) := C(S, \mathbb{R})$, and

$$C_c(S) := \{f \in C(S) \,|\, \operatorname{supp}(f) \text{ is compact}\}.$$

Topological vector spaces

When Banach published his famous book ‘*Théorie des operations linéaires*’ in 1932, it was the opinion that normed spaces provide a sufficiently wide framework to comprehend all interesting concrete problems of analysis. It turned out, however, in a short time that this is an illusion: a number of (non-artificial) problems of analysis lead to infinite-dimensional vector spaces whose topology cannot be derived from a norm. In this subsection we recall the most basic definitions and facts concerning topological vector spaces, especially locally convex spaces. We restrict ourselves to real vector spaces, although vector spaces over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ can also be treated without any extra difficulties.

If V is a (real) vector space, then a subset H of V is said to be

convex if for each $t \in [0,1]$, $tH + (1-t)H \subset H$;

balanced if $\alpha H \subset H$ whenever $\alpha \in [-1,1]$;

absorbing if $\cup_{\lambda \in \mathbb{R}_+^*} \lambda H = V$, i.e., for all $v \in V$ there is a positive real number λ (depending on v) such that $v \in \lambda H$.

By a *topological vector space* (TVS) we mean a real vector space V endowed with a Hausdorff topology compatible with the vector space structure of V in the sense that the addition map $V \times V \to V$, $(u,v) \mapsto u+v$ and the scalar multiplication $\mathbb{R} \times V \to V$, $(\lambda, v) \mapsto \lambda v$ are continuous. (The product spaces are equipped with the product topology.) Such a topology is called a *linear topology* or *vector topology* on V.

Let V be a topological vector space. For each $a \in V$ the *translation* $T_a : v \in V \mapsto T_a(v) := a + v \in V$, and for each $\lambda \in \mathbb{R}^*$ the *homothety* $h_\lambda : v \in V \mapsto \lambda v \in V$ are homeomorphisms. This simple observation is very important: it implies, roughly speaking, that the linear topology *looks like the same* at any point. Thus, practically, in a TVS it is enough to define a local concept or to prove a local property only in a neighbourhood of the origin. We agree that in TVS context *a fundamental system will always mean a fundamental system of the origin*. Thus $\mathcal{F}$ is a fundamental system in V if $\mathcal{F} \subset \mathcal{N}(0)$, and every neighbourhood of 0 (briefly 0-neighbourhood) contains a member of $\mathcal{F}$.

In the category TVS the morphisms are the *toplinear isomorphisms*: linear isomorphisms which are homeomorphisms at the same time. As the following classical result shows, the structure of finite dimensional topological vector spaces is the simplest possible.

Lemma 1.5 (A. Tychonoff) *If V is an n-dimensional topological vector space, then there is a toplinear isomorphism from V onto the Euclidean n-space $\mathbb{R}^n$.*

The *idea of proof* is immediate: we choose a basis $(v_1, \dots, v_n)$ of V, and we show that the linear isomorphism

$$\left(\nu^1, \dots, \nu^n\right) \in \mathbb{R}^n \mapsto \sum_{i=1}^{n} \nu^i v_i \in V,$$

which is obviously continuous, is an *open* map. For details see e.g. [7, p. 413] or [36, p. 28].

Now we return to a generic topological vector space V. A subset H of V is said to be *bounded* if for every 0-neighbourhood $\mathcal{U}$ there is a positive real number ε such that $\varepsilon H \subset \mathcal{U}$. V is called *locally convex* if it has a fundamental system whose members are convex sets, i.e., every 0-neighbourhood contains a convex 0-neighbourhood. V has the *Heine – Borel property* if every closed and bounded subset of V is compact.

Lemma 1.6 *If a topological vector space V has a countable fundamental system, then V is metrizable. More precisely, there is a distance function ϱ on V such that*

(i) ϱ is compatible with the linear topology of V;

(ii) ϱ *is translation invariant, i.e.,* $\varrho(T_a(u), T_a(v)) = \varrho(u, v)$ *for all* $a, u, v \in V$*;*

(iii) *the open* ϱ*-balls centred at the origin are balanced.*

If, in addition, V *is locally convex, then* ϱ *can be chosen so as to satisfy* (i)–(iii)*, and also*

(iv) *all open* ϱ*-balls are convex.*

For a proof we refer to [33, I, 1.24].

A net $(v_\alpha)_{\alpha \in A}$ in a locally convex space V is said to be a *Cauchy net* if for every 0-neighbourhood $\mathcal{U}$ in V there is an $\alpha_0 \in A$ such that

$$v_\alpha - v_\beta \in \mathcal{U} \text{ whenever } \alpha, \beta \geqq \alpha_0.$$

In particular, a sequence in V is called a *Cauchy sequence* if it is a Cauchy net. Note that if the topology of V is compatible with a translation invariant distance function $\varrho : V \times V \to \mathbb{R}$, then a sequence $(v_n)_{n \in \mathbb{N}^*}$ in V is a Cauchy sequence if and only if it is a Cauchy sequence in metrical sense, i.e., for every $\varepsilon \in \mathbb{R}_+^*$ there is an integer n_0 such that $\varrho(v_m, v_n) < \varepsilon$ whenever $m > n_0$ and $n > n_0$. A locally convex space V is said to be *complete*, resp. *sequentially complete* if any Cauchy net, resp. Cauchy sequence is convergent in V. These completeness concepts coincide if V is metrizable, i.e., we have

Lemma 1.7 *A metrizable locally convex space is complete if and only if it is sequentially complete.*

As for the *proof*, the only technical difficulty is to check that sequential completeness implies completeness, i.e., the convergence of every Cauchy net. For a detailed reasoning we refer to [8, B.6.2].

It follows from our preceding remarks that if the topology of a locally convex space is induced by a translation invariant distance function, then the space is complete if and only if it is complete as a metric space. TVSs sharing these properties deserve an own name: a locally convex space is said to be a *Fréchet space* if its linear topology is induced by a complete, translation invariant distance function. We shall see in the next chapter that important examples of Fréchet spaces occur even in the context of classical analysis. In the rest of this chapter we are going to indicate how one can construct locally convex spaces, in particular Fréchet spaces, starting from a family of seminorms.

We recall that a *seminorm* on a real vector space V is a function $\nu : V \to \mathbb{R}$, satisfying the following axioms:

$\nu(u + v) \leqq \nu(u) + \nu(v)$ for all $u, v \in V$ (*subadditivity*);

$\nu(\lambda v) = |\lambda|\nu(v)$ for all $\lambda \in \mathbb{R}$, $v \in V$ (*absolute homogeneity*).

Then it follows that $\nu(0) = 0$, and $\nu(v) \geqq 0$ for all $v \in V$. If, in addition, $\nu(v) = 0$ implies $v = 0$, then ν is a *norm* on V. As in the case of metric spaces, given a point $a \in V$ and a positive real number r, we use the notation

$$B_r^\nu(a) := \{v \in V | \nu(v - a) < r\}$$

and the term 'open ν-ball with centre a and radius r'. It can be seen immediately that the 'open unit ν-ball'

$$B := B_1^\nu(0) = \{v \in V | \nu(v) < 1\}$$

is convex, balanced and absorbing.

A family $\mathcal{P} = (\nu_\alpha)_{\alpha \in A}$ of seminorms on V is said to be *separating* if for any point $v \in V \setminus \{0\}$ there is an index $\alpha \in A$ such that $\nu_\alpha(v) \neq 0$. A separating family of seminorms on a vector space is also called a *multinorm*. A *multinormed vector space* is a vector space endowed with a multinorm.

Lemma 1.8 *Suppose $\mathcal{P} = (\nu_\alpha)_{\alpha \in A}$ is a separating family of seminorms on a vector space V. For each $\alpha \in A$ and $n \in \mathbb{N}^*$, let*

$$\mathcal{V}(\alpha, n) := \left\{ v \in V \middle| \nu_\alpha(v) < \frac{1}{n} \right\}.$$

If $\mathcal{F}$ is the family of all finite intersections of the sets $\mathcal{V}(\alpha, n)$, then $\mathcal{F}$ is a fundamental system for a topology on V, which makes V into a locally convex TVS such that

(i) the members of $\mathcal{F}$ are convex balanced sets;

(ii) for each $\alpha \in A$, the function $\nu_\alpha : V \to \mathbb{R}$ is continuous;

(iii) a subset H of V is bounded if and only if every member of $\mathcal{P}$ is bounded on H.

For a proof we refer to Rudin's text [33].

Some comments to this important result seem to be appropriate.

Remark 1.9 Suppose V is a locally convex space, and let $\mathcal{T}$ be its linear topology. It may be shown that if $\mathcal{F}$ is a fundamental system of V consisting of convex balanced sets, then $\mathcal{F}$ generates a separating family $\mathcal{P}$ of seminorms on V. According to 1.8, $\mathcal{P}$ induces a topology $\mathcal{T}_1$ on V. Now it can easily be checked that $\mathcal{T}_1 = \mathcal{T}$.

Remark 1.10 Suppose that in Lemma 1.8 a *countable* family $\mathcal{P} = (\nu_n)_{n \in \mathbb{N}^*}$ of seminorms is given on V. Then the fundamental system arising from $\mathcal{P}$ is also countable, therefore, by Lemma 1.6, the induced topology is compatible with a translation invariant distance function. In addition, such a distance function can explicitly be constructed in terms of $\mathcal{P}$; for example, the formula

$$\varrho(u, v) := \sum_{n=1}^{\infty} \frac{2^{-n} \nu_n(u - v)}{1 + \nu_n(u - v)}; \quad (u, v) \in V \times V$$

defines a distance function which satisfies the desired properties.

As a consequence of our preceding discussion, we have the following useful characterization of Fréchet spaces.

Corollary 1.11 *A vector space is a Fréchet space, if and only if, it is a complete multinormed vector space $(V, (\nu_\alpha)_{\alpha \in A})$, where the set A is countable. Completeness is understood with respect to the* multinorm topology *whose subbasis is the family*

$$\left\{ B_r^{\nu_\alpha}(v) \subset V | \alpha \in A, r \in \mathbb{R}_+^*, v \in V \right\}.$$

The distance function described in 1.10 is compatible with the multinorm topology. □

2 Calculus in topological vector spaces and beyond

Local analysis in the context of topological vector spaces

In infinite-dimensional analysis there is a deep breaking between the case of (real or complex) Banach spaces and that of more general locally convex topological vector spaces which are not normable: depending on the type of derivatives used (Fréchet derivative, Gâteaux derivative, …) one obtains non-equivalent calculi. As a consequence, there are several theories of infinite-dimensional manifolds, Lie groups and differential geometric structures. Changing the real or complex ground field to a more general topological field or ring, even more general differential calculus, Lie theory and differential geometry may be constructed [3, 4]. In this subsection we briefly explain the approach to differential calculus originated by A. D. Michel [25] and A. Bastiani [2], and popularized by J. Milnor [27] and R. Hamilton [13]. For an accurately elaborated, detailed recent account we refer to [8].

For simplicity, and in harmony with the applications we are going to present, *we restrict ourselves to the real case*. We begin with some remarks concerning the differentiability of curves with values in a locally convex topological vector space. This is the simplest situation, without any difficulty in principle. However, some interesting new phenomena occur.

Definition 2.1 Let V be a locally convex vector space and $I \subset \mathbb{R}$ an interval containing more than one point. By a C^0*-curve* on I with values in V we mean a continuous map $\gamma : I \to V$.

(1) Let $\alpha, \beta \in \mathbb{R}$, $\alpha < \beta$. A C^0-curve $\gamma : [\alpha, \beta] \to V$ is called a *Lipschitz curve* if the set

$$\left\{ \frac{1}{s-t}(\gamma(s) - \gamma(t)) \in V \middle| s, t \in [\alpha, \beta], s \neq t \right\}$$

is bounded in V.

(2) Suppose I is an open interval. A C^0-curve $\gamma : I \to V$ is said to be a C^1*-curve* if the limit

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

exists for all $t \in I$, and the map

$$\gamma' : t \in I \mapsto \gamma'(t) \in V$$

is continuous. Given $k \in \mathbb{N}^*$, γ is called *of class* C^k if all its iterated derivatives up to order k exist and are continuous. γ is *smooth* if it is of class C^k for all $k \in \mathbb{N}$.

Remark 2.2 It may be shown (but not at this stage of the theory) that if $\gamma : I \to V$ is a C^1-curve and $[\alpha, \beta] \subset I$ is a compact interval, then $\gamma \restriction [\alpha, \beta]$ is Lipschitz.

Definition 2.3 Suppose V is a locally convex vector space, and let $\gamma : [\alpha, \beta] \to V$ be a C^0-curve. If there exists a vector $v \in V$ such that for every continuous linear form $\lambda \in V^*$ we have

$$\lambda(v) = \int_\alpha^\beta \lambda \circ \gamma \quad \text{(Riemann integral)},$$

then $v \in V$ is called the *weak integral* of γ from α to β, and the notation

$$v =: \int_\alpha^\beta \gamma = \int_\alpha^\beta \gamma(t)dt$$

is applied. V is said to be *Mackey-complete* if the weak integral $\int_\alpha^\beta \gamma$ exists for each *Lipschitz-curve* $\gamma : [\alpha, \beta] \to V$.

Remark 2.4 (1) If the weak integral of a C^0-curve $\gamma : [\alpha, \beta] \to V$ exists, then it is unique, since by a version of the Hahn – Banach theorem V^* separates points on V (see e.g. [33, 3.4, Corollary]).

(2) If V is a sequentially complete locally convex vector space, then it is also Mackey-complete: it may be shown that the weak integral of any C^0-curve $\gamma : [\alpha, \beta] \to V$ can be obtained as the limit of a sequence of Riemann sums. For details see [8]. The importance of the concept of Mackey-completeness lies in the fact that a number of important constructions of the theory depends only on the existence of some weak integrals. For an alternative definition and an exhaustive description of Mackey-completeness we refer to the monograph of A. Kriegl and P. Michor [18].

(3) Let V and W be locally convex spaces, $\varphi : V \to W$ a continuous linear mapping, and $\gamma : I \to V$ a C^1-curve. One can check by an immediate application of the definition that $\varphi \circ \gamma : I \to W$ is also a C^1-curve, and

$$(\varphi \circ \gamma)' = \varphi \circ \gamma'.$$

Lemma 2.5 (Fundamental theorem of calculus) *Let V be a locally convex vector space, and $I \subset \mathbb{R}$ be an open interval.*

(1) If $\gamma : I \to V$ is a C^1-curve and $\alpha, \beta \in I$, then

$$\gamma(\beta) - \gamma(\alpha) = \int_\alpha^\beta \gamma'(t)dt.$$

(2) If $\gamma : I \to V$ is a C^0-curve, $\alpha \in I$, and the weak integral

$$\eta(t) := \int_\alpha^t \gamma(s)ds$$

exists for all $t \in I$, then $\eta : I \to V$ is a C^1-curve, and $\eta' = \gamma$.

Proof of part 1. Let $\lambda \in V^*$ be a continuous linear form. By Remark 2.4(3) and the classical 'Fundamental theorem of calculus', it follows that

$$\lambda(\gamma(\beta) - \gamma(\alpha)) = \lambda\gamma(\beta) - \lambda\gamma(\alpha) = \int_\alpha^\beta (\lambda \circ \gamma)' = \int_\alpha^\beta \lambda \circ \gamma'.$$

It means that $\gamma(\beta) - \gamma(\alpha)$ is the weak integral of γ' from α to β. □

Remark 2.6 The second part of the lemma is proved in [13]; it needs a more sophisticated argument, and hence additional preparations.

Definition 2.7 Let V and W be locally convex topological vector spaces, $\mathcal{U} \subset V$ an open set, and $f : \mathcal{U} \to W$ a mapping.

(1) By the *derivative* of f at a point $p \in \mathcal{U}$ *in the direction* $v \in V$ we mean the limit

$$D_v f(p) = df(p, v) := \lim_{t \to 0} \frac{1}{t}(f(p + tv) - f(p))$$

whenever it exists. f is called *differentiable at* p if $df(p, v)$ exists for all $v \in V$.

(2) The map f is said to be *continuously differentiable* or *of class* C^1 (briefly C^1) on $\mathcal{U}$ if it is differentiable at every point of $\mathcal{U}$ and the map

$$df : \mathcal{U} \times V \to W, \quad (p, v) \mapsto df(p, v)$$

is continuous.

(3) Let $k \in \mathbb{N}$, $k \geqq 2$. f is called a C^k*-map* (or briefly C^k) if it is of class C^1 and $d^1 f := df$ is a C^{k-1}-map. Then the k*-th iterated differential* of f is $d^k f := d^{k-1}(df)$. If f is C^k for all $k \in \mathbb{N}$, then f is said to be C^∞ or *smooth*.

Notation We write $C^k(\mathcal{U}, W)$ for the set (in fact a vector space) of C^k-maps from $\mathcal{U}$ into W. When W is 1-dimensional, and hence $W \cong \mathbb{R}$, we usually just write $C^k(\mathcal{U})$.

Remark 2.8 (1) It is obvious from the definition that the derivative of a *linear mapping* exists at every point. Since there are linear mappings which are not continuous, it follows that *differentiability does not imply continuity*. However, if $\varphi : V \to W$ is a *continuous linear mapping*, then φ is smooth, and at every $(p, v) \in V \times V$ we have $d\varphi(p, v) = \varphi(v)$, while $d^k \varphi = 0$ for $k \geqq 2$.

(2) If is an important technicality that f is C^k ($k \geqq 1$), if and only if, it is C^{k-1}, and $d^{k-1} f$ is C^1. Then we also have $d^k f = d\left(d^{k-1} f\right)$.

(3) The concept of C^k-differentiability introduced here will occasionally be mentioned as the *Michel – Bastiani differentiability*. Observe that in the formulation of the definition the local convexity of the underlying vector spaces does not play any role. However, if one wants to build a 'reasonable' theory of differentiation (with 'expectable' rules for calculation), the requirement of local convexity is indispensable.

Lemma 2.9 *Let V and W be locally convex vector spaces, $\mathcal{U} \subset V$ an open subset, and $f : \mathcal{U} \to W$ a C^1-map.*

(1) The map

$$f'(p) : V \to W, \quad v \mapsto f'(p)(v) := df(p, v)$$

is a continuous linear map for each $p \in \mathcal{U}$, and f is continuous.

(2) Let $p \in \mathcal{U}$, $v \in V$, and suppose that $p + tv \in \mathcal{U}$ for all $t \in [0, 1]$. Define the C^0-curve $c : [0, 1] \to W$ by

$$c(t) := df(p + tv, v) = f'(p + tv)(v).$$

Then

$$f(p+v) = f(p) + \int_0^1 c,$$

therefore f *is locally constant, if and only if,* $df = 0$.

(3) (Chain rule) *Suppose* Z *is another locally convex vector space,* $\mathcal{V} \subset W$ *is an open subset, and* $h : \mathcal{V} \to Z$ *is a* C^1*-map. If* $f(\mathcal{U}) \subset \mathcal{V}$*, then* $h \circ f : \mathcal{U} \to Z$ *is also a* C^1*-map, and for all* $p \in \mathcal{U}$ *we have*

$$(h \circ f)'(p) = h'(f(p)) \circ f'(p).$$

(4) (Schwarz's theorem) *If* f *is of class* C^k $(k \geqq 2)$*, then*

$$f^{(k)}(p) : (v_1, \dots, v_k) \in V^k \mapsto f^{(k)}(p)(v_1, \dots, v_k) := d^k f(p, v_1, \dots, v_k)$$

is a continuous, symmetric k*-linear map for all* $p \in \mathcal{U}$.

(5) (Taylor's formula) *Suppose* f *is of class* C^k $(k \geqq 2)$*. Then, if* $p \in \mathcal{U}$*,* $v \in V$ *and the segment joining* p *and* $p+v$ *is in* $\mathcal{U}$*, we have*

$$f(p+v) = f(p) + f'(p)(v) + \ldots + \frac{1}{(k-1)!} f^{(k-1)}(p)(v, \dots, v) + \frac{1}{(k-1)!} \int_0^1 c_k,$$

where $c_k : [0,1] \to W$ *is a* C^0*-curve given by*

$$c_k(t) := (1-t)^{k-1} f^{(k)}(p+tv)(v, \dots, v), \quad t \in [0,1].$$

Indication of proof. The continuity of $f'(p)$ is obvious, since $f'(p) = df(p, \cdot)$, and df is continuous. An immediate application of the definition of differentiability leads to the homogeneity of $f'(p)$. To check the additivity of $f'(p)$ some further (but not difficult) preparation is necessary, see [8, 1.2.13, 1.2.14]. To prove the integral representation in (2), let

$$\gamma(t) := f(p+tv), \quad t \in [0,1].$$

Then γ is differentiable at each $t \in [0,1]$ in the sense of 2.1(2), namely

$$\gamma'(t) := \lim_{s \to 0} \frac{\gamma(t+s) - \gamma(t)}{s} = \lim_{s \to 0} \frac{1}{s}(f(p+tv+sv) - f(p+tv)) =: df(p+tv, v) = f'(p+tv)(v) =: c(t).$$

Thus $\gamma' = c$, and the fundamental theorem of calculus (2.5(1)) gives

$$f(p+v) - f(p) = \gamma(1) - \gamma(0) = \int_0^1 \gamma' = \int_0^1 c.$$

To see that f is continuous, choose a continuous seminorm $\nu : W \to \mathbb{R}$, and let ε be an arbitrary positive real number. Then there exists a balanced neighbourhood $\mathcal{U}_0$ of the origin in V such that $p + \mathcal{U}_0 \subset \mathcal{U}$, and for all $t \in [0,1]$, $v \in \mathcal{U}_0$ we have

$$\nu(c(t)) = \nu(f'(p+tv)(v)) \leqq \varepsilon.$$

Now it may be shown [8, 1.1.8] that

$$\nu\left(\int_0^1 c\right) \leqq \sup\{\nu(c(t)) \in \mathbb{R} | t \in [0,1]\},$$

therefore

$$\nu(f(p+v) - f(p)) = \nu\left(\int_0^1 c\right) \leqq \varepsilon,$$

and hence f is continuous. This concludes the sketchy proof of (1) and (2).

For a proof of the chain rule and Schwarz's theorem we refer to [13] and [8]. The latter reference also contains a detailed treatment of Taylor's formula. □

Remark 2.10 We keep the hypotheses and notations of the Lemma.

(1) The continuous linear map $f'(p) : V \to W$ introduced in 2.9(1) is said to be the *derivative* of f at p. Note that the symbol $f'(t)$ carries double meaning if $f : I \to W$ is a C^1-curve: by 2.1(2) $f'(t) \in W$, while by 2.9(1) $f'(t) \in \mathcal{L}(\mathbb{R}, W)$. Fortunately, this abuse of notation leads to no serious conflict since the vector spaces W and $\mathcal{L}(\mathbb{R}, W)$ can be canonically identified via the linear isomorphism

$$\gamma \in \mathcal{L}(\mathbb{R}, W) \mapsto \gamma(1) \in W.$$

(2) In Lemma 2.9 we have listed only the most elementary facts concerning Michel–Bastiani differentiation. It is a more subtle problem, for example, to obtain inverse (or implicit) function theorems in this (or a more general) context. The idea of generalization of the classical *inverse function theorem* for mappings between some types of Fréchet spaces is due to John Nash. Nash's inverse function theorem played an important role in his famous paper on isometric embeddings of Riemannian manifolds [30]. It was F. Sergeraert who stated the theorem explicitly in terms of a category of maps between Fréchet spaces [34]. In Moser's formulation [28] the theorem became an abstract theorem in functional analysis of wide applicability. Further generalizations have been given by Hamilton [14], Kuranishi [19], Zehnder [37], and more recently Leslie [22] and Ma [24]. In reference [10], inspired by Hiltunen's results [16], Glöckner proves *implicit function theorems* for mappings defined on topological vector spaces over valued fields. In particular, in the real and complex cases he obtains implicit function theorems for mappings from not necessarily locally convex topological vector spaces to Banach spaces.

To emphasize that the results of calculus in Banach spaces cannot be transplanted into the wider framework of topological vector spaces in general, finally we mention a quite typical pathology: the uniqueness and existence of solutions to ordinary differential equations are not guaranteed beyond Banach spaces. For a simple illustration of this phenomenon in Fréchet spaces see [13, 5.6.1].

(3) We briefly discuss the relation between the concept of Michel – Bastiani differentiability and the classical concept of Fréchet differentiability for mappings between Banach spaces. Recall that a continuous map from an open subset $\mathcal{U}$ of a Banach space V into a Banach space W is called *continuously Fréchet differentiable* or FC^1, if for all $p \in \mathcal{U}$ there exists a (necessarily unique) continuous linear map $f'(p) : V \to W$ such that

$$\lim_{v \to 0} \frac{f(p+v) - f(p) - f'(p)(v)}{\|v\|} = 0,$$

and the map $f' : \mathcal{U} \to \mathcal{L}(V, W)$, $p \mapsto f'(p)$ is continuous (with respect to the operator norm in $\mathcal{L}(V, W)$). Inductively, we define f to be FC^k ($k \geqq 2$) if f' is FC^{k-1}. Now it may be shown that *every FC^k-map is C^k (in the sense of Michel – Bastiani), and every C^{k+1}-map between open subsets of Banach spaces is FC^k*, so the two concepts coincide in the C^∞ case. For a proof we refer to [26] or [11].

(4) As an equally important approach to non-Banach infinite-dimensional calculus we have to mention the so-called *convenient calculus* elaborated in detail by A. Kriegl and P. W. Michor [18]. Let V and W be locally convex vector spaces and $\mathcal{U} \subset V$ an open subset. A mapping $f : \mathcal{U} \to W$ is said to be *conveniently smooth* if $f \circ \gamma : I \to W$ is a smooth curve for each smooth curve $\gamma : I \to \mathcal{U}$. By the chain rule 2.9(3) it is clear that if a mapping is Michel – Bastiani smooth, then it is conveniently smooth as well. The converse of this statement is definitely false: a conveniently smooth mapping need not even be continuous. However, *if V is a Fréchet space, then $f : \mathcal{U} \to W$ is conveniently smooth if and only if it is Michel – Bastiani smooth*. For a sketchy proof see [31, II, 2.10].

Remark 2.11 It should be noticed that all the difficulties arising in our preceding discussion disappear if the underlying vector spaces are finite dimensional. The first reason for this lies in the fact that Tychonoff's theorem (1.5) guarantees, roughly speaking, that the Euclidean topology of $\mathbb{R}^n$ is the only linear topology that an n-dimensional real vector space can have. This canonical topology is locally convex, locally compact and metrizable. In particular, if V and W are finite dimensional vector spaces, then $\mathcal{L}(V, W)$ also carries the canonical linear topology. This also leads to a significant difference between the calculus in finite or infinite dimensional Banach spaces on the one hand, and in non-Banachable spaces on the other hand. For example, if V and W are (non-Banach) Fréchet spaces, then the vector space of continuous linear maps between V and W is not necessarily a Fréchet space. In the following *we assume that all finite dimensional vector spaces are endowed with the canonical topology assured by Tychonoff's theorem*.

In the finite dimensional case the differentiability concepts introduced above result in the same class of C^k mappings. For lack of norms we shall always use the Michel – Bastiani definition. So if V and W are finite dimensional (real) vector spaces, $\mathcal{U} \subset V$ is an open set, then a mapping $f : \mathcal{U} \to W$ is C^1 if there is a continuous mapping $f' : \mathcal{U} \to \mathcal{L}(V, W)$ such that at each point $p \in \mathcal{U}$ and for all $v \in V$ we have

$$f'(p)(v) = \lim_{t \to 0} \frac{1}{t}(f(p+tv) - f(p)).$$

Higher derivatives and smoothness can be defined by induction, as in 2.7(3). Notice that if $f \in C^k(\mathcal{U}, W)$ ($k \geqq 2$), then its kth derivative at a point $p \in \mathcal{U}$ is denoted by

$f^{(k)}(p)$, and it is an element of

$$\mathcal{L}\left(V, \mathcal{L}^{k-1}(V,W)\right) \cong \mathcal{L}^k(V,W)$$

given by

$$f^{(k)}(p)(v_1,\ldots,v_k) := \lim_{t\to 0}\frac{1}{t}\Big(f^{(k-1)}(p+tv_1)(v_2,\ldots,v_k) - f^{(k-1)}(p)(v_2,\ldots,v_k)\Big)$$

for each $(v_1,\ldots,v_k) \in V^k$.

Smooth functions and differential operators on $\mathbb{R}^n$

In this subsection we have a closer look at the most important special case, when the domain of the considered functions is a subset of the Euclidean n-space $\mathbb{R}^n$, and we describe the linear differential operators acting on the spaces of these functions.

First we recall a basic existence result.

Lemma 2.12 *Let $\mathcal{U}$ be a nonempty open subset of $\mathbb{R}^n$, and $(\mathcal{U}_i)_{i\in I}$ an open covering of $\mathcal{U}$. There exists a family $(f_i)_{i\in I}$ of smooth functions on $\mathcal{U}$ such that*

(i) $0 \leqq f_i(p) \leqq 1$ *for all $i \in I$ and $p \in \mathcal{U}$;*

(ii) $\operatorname{supp}(f_i) \subset \mathcal{U}_i$ *for all $i \in I$;*

(iii) $(\operatorname{supp}(f_i))_{i\in I}$ *is locally finite;*

(iv) for each point $p \in \mathcal{U}$ we have $\sum_{i\in I} f_i(p) = 1$. □

The family $(f_i)_{i\in I}$ in the Lemma is said to be a *partition of unity* subordinate to the covering $(\mathcal{U}_i)_{i\in I}$. The heart of the proof consists of an application of the purely topological Lemma 1.2 and the construction of a smooth function $f : \mathbb{R}^n \to \mathbb{R}$, called a smooth bump function, with the following properties:

$$0 \leqq f(p) \leqq 1 \text{ for all } p \in \mathbb{R}^n; \quad f(q) = 1 \text{ if } q \in \overline{B_1(0)}; \quad \operatorname{supp}(f) \subset \overline{B_2(0)}$$

(the balls are taken with respect to the Euclidean distance).

More generally, let $\mathcal{U}$ and $\mathcal{V}$ be open subsets, K a closed subset of $\mathbb{R}^n$, and suppose that $K \subset \mathcal{V} \subset \mathcal{U}$. A smooth function $f : \mathcal{U} \to \mathbb{R}$ is said to be a *bump function for K supported in $\mathcal{V}$* if $0 \leqq f(p) \leqq 1$ for each $p \in \mathcal{U}$, $f(q) = 1$ if $q \in K$, and $\operatorname{supp}(f) \subset \mathcal{V}$. As an immediate consequence of Lemma 2.12, we have

Corollary 2.13 *If $\mathcal{U}$ and $\mathcal{V}$ are open subsets, K is a closed subset of $\mathbb{R}^n$, and $K \subset \mathcal{V} \subset \mathcal{U}$, then there exists a bump function for K supported in $\mathcal{V}$.*

Indeed, if $\mathcal{U}_0 := \mathcal{V}, \mathcal{U}_1 := \mathcal{U} \setminus K$, then $(\mathcal{U}_0, \mathcal{U}_1)$ is an open covering of $\mathcal{U}$, so by 2.12 there exists a partition of unity (f_0, f_1) subordinate to $(\mathcal{U}_0, \mathcal{U}_1)$. Then $f_1 \restriction K = 0$, and for each $q \in K$ we have $f_0(q) = (f_0 + f_1)(q) = 1$, therefore the function $f := f_0$ has the desired properties.

Remark 2.14 In the following the canonical basis of $\mathbb{R}^n$ will be denoted by $(e_i)_{i=1}^n$, and $(e^i)_{i=1}^n$ will stand for its dual. The family $(e^i)_{i=1}^n$ will also be mentioned as the canonical coordinate system for $\mathbb{R}^n$. If $\mathcal{U} \subset \mathbb{R}^n$ is an open subset, and $f \in C^\infty(\mathcal{U})$, then

$$D_i f : \mathcal{U} \to \mathbb{R}, \quad p \mapsto D_i f(p) := f'(p)(e_i) \quad (i \in \{1, \dots, n\})$$

is the *ith partial derivative* of f (with respect to the canonical coordinate system). For each multi-index $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ we write

$$D^\alpha f := D_1^{\alpha_1} \dots D_n^{\alpha_n} f; \quad D_i^{\alpha_i} f := \underbrace{D_i \dots D_i}_{\alpha_i \text{times}} f \quad (i \in \{1, \dots, n\})$$

with the convention $D^\alpha f := f$ if $|\alpha| = 0$. We say that D^α is an *elementary partial differential operator of order* $|\alpha|$. A *linear differential operator* is a linear combination $D = \sum_{\|\alpha\| \leqq m} a_\alpha D^\alpha$, where $m \in \mathbb{N}$, $a_\alpha \in C^\infty(\mathcal{U})$. Clearly, D maps $C^\infty(\mathcal{U})$ linearly into $C^\infty(\mathcal{U})$. It is not difficult to show (see e.g. [6, (8.13.1)], or [20, Ch. XI, §1]) that *if a linear differential operator is identically* 0 *on* $C^\infty(\mathcal{U})$*, then each of its coefficients is identically* 0 *on* $\mathcal{U}$. From this it follows that the coefficients of a linear diferential operator $D = \sum_{\|\alpha\| \leqq m} a_\alpha D^\alpha$ are uniquely determined; the highest value of $|\alpha|$ such that $a_\alpha \neq 0$ is called the *order* of D.

Lemma 2.15 (generalized Leibniz rule) *Let* $\mathcal{U} \neq \emptyset$ *be an open subset of* $\mathbb{R}^n$*. If* $f, h \in C^\infty(\mathcal{U})$ *and* $\alpha \in \mathbb{N}^n$ *is a multi-index, then*

$$D^\alpha(fh) = \sum_{\mu+\nu=\alpha} \binom{\alpha}{\nu} (D^\nu f)(D^\mu h).$$

□

Remark 2.16 We need some further notations. We shall denote by $C_c^\infty(\mathbb{R}^n)$ the subspace of $C^\infty(\mathbb{R}^n)$ consisting of smooth functions on $\mathbb{R}^n$ which have compact support. For any nonempty subset S of $\mathbb{R}^n$, $C_c^\infty(S)$ will stand for the space of functions in $C_c^\infty(\mathbb{R}^n)$ whose support lies in S. A function in $C_c^\infty(S)$ will be identified with its restriction to S. (Notice that L. Schwartz's notation $\mathcal{E}(\mathcal{U}) := C^\infty(\mathcal{U})$, $\mathcal{D}(\mathcal{U}) := C_c^\infty(\mathcal{U})$, $\mathcal{D}(K) := C_c^\infty(K)$ is widely used in distribution theory.)

Proposition 2.17 *Let a (nonempty) open subset* $\mathcal{U}$ *of* $\mathbb{R}^n$ *be given. For any compact set* K *contained in* $\mathcal{U}$ *and any multi-index* $\alpha \in \mathbb{N}^n$*, define the function* $\| \; \|_\alpha^K : C^\infty(\mathcal{U}) \to \mathbb{R}$ *by setting*

$$\|f\|_\alpha^K := \sup_{p \in K} |D^\alpha f(p)|, \quad f \in C^\infty(\mathcal{U}).$$

Then the family $\left(\| \; \|_\alpha^K\right)$ *is a multinorm on* $C^\infty(\mathcal{U})$ *which makes it into a Fréchet space having the Heine – Borel property, such that* $C_c^\infty(K)$ *is a closed subspace of* $C^\infty(\mathcal{U})$ *whenever* $K \subset \mathcal{U}$ *is compact.*

Proof. It is clear that the family $\left(\| \; \|_\alpha^K\right)$ is a multinorm, so it defines a locally convex topology on $C^\infty(\mathcal{U})$ according to 1.8. We claim that the TVS so obtained is metrizable and complete. In order to show this we construct another *countable* family of seminorms on $C^\infty(\mathcal{U})$ which is equivalent to the given one in the sense that the two families generate the same topology.

Let $(K_n)_{n\in\mathbb{N}^*}$ be the compact exhaustion of $\mathcal{U}$ described in the proof of 1.1. If for each $n \in \mathbb{N}^*$

$$\nu_n(f) := \sup_{p\in K_n, |\alpha|\leqq n} |D^\alpha f(p)|, \quad f \in C^\infty(\mathcal{U}),$$

then $(\nu_n)_{n\in\mathbb{N}^*}$ is a multinorm on $C^\infty(\mathcal{U})$, which defines a metrizable locally convex topology on $C^\infty(\mathcal{U})$ by 1.8 and 1.10. To prove the equivalence of the two seminorm-families, it is enough to check that each member of the first family is majorized by a finite linear combination of the second family, and conversely.

Now, on the one hand, it is obvious that for every positive integer n and smooth function f in $C^\infty(\mathcal{U})$ we have

$$\nu_n(f) \leqq \sum_{|\alpha|\leqq n} \|f\|_\alpha^{K_n} \left(\overset{2.18}{=} \|f\|_n^{K_n}\right),$$

thus ν_n is majorized by $\left(\| \, \|_\alpha^{K_n}\right)_{|\alpha|\leqq n}$. To prove the converse, choose a compact set $K \subset \mathcal{U}$, and define the following functions:

$$\begin{aligned}&\delta : K \to]0, \infty] , \quad p \mapsto \delta(p) := \varrho_E(p, \partial\mathcal{U});\\ &\Delta : K \to \mathbb{R}_+, \quad p \mapsto \Delta(p) := \varrho_E(p, 0) = \|p\|.\end{aligned}$$

Then both δ and Δ are continuous; δ attains its minimum $c \in \mathbb{R}_+^*$, and Δ attains its maximum $C \in \mathbb{R}_+$ on K. Now choose $n \in \mathbb{N}^*$ such that

$$\frac{1}{n} < c \quad \text{and} \quad n > C.$$

Then the member K_n of the compact exhaustion of $\mathcal{U}$ contains the compact set K. If, in addition, $n \geqq |\alpha|$, the seminorm ν_n is a majorant of the seminorm $\| \, \|_\alpha^K$.

For a proof of the remaining claims we refer to [33, 1.46]. □

Remark 2.18 Let $\mathcal{U}$ be an open subset of $\mathbb{R}^n$, S any subset of $\mathcal{U}$, and $m \in \mathbb{N}$. It is easy to check that the mapping

$$\| \, \|_m^S : f \in C_c^\infty(\mathcal{U}) \mapsto \|f\|_m^S := \sum_{|\alpha|\leqq m} \sup_{p\in S} |D^\alpha f(p)|$$

is also a seminorm on $C_c^\infty(\mathcal{U})$. Using this seminorm, in case of $S = \mathcal{U}$ we shall omit the superscript in the notation.

Now, following Narasimhan [29, 1.5.1], we introduce a purely technical term. We say that a function $f \in C^\infty(\mathcal{U})$ is m*-flat* at a point $p \in \mathcal{U}$ if $D^\alpha f(p) = 0$, whenever $|\alpha| \leqq m$. We shall find the following observation useful:

Lemma 2.19 *If a function $f \in C^\infty(\mathbb{R}^n)$ is m-flat at the origin, then for every $\varepsilon \in \mathbb{R}_+^*$ there exists a function $h \in C^\infty(\mathbb{R}^n)$, vanishing on a 0-neighbourhood, such that $\|h - f\|_m < \varepsilon$.*

Proof. Corollary 2.13 assures the existence of a function $\psi \in C^\infty(\mathbb{R}^n)$ such that

$$\forall p \in \mathbb{R}^n : 0 \leqq \psi(p) \leqq 1, \quad \psi(p) = 0 \text{ if } p \in \overline{B_{\frac{1}{2}}(0)}, \quad \psi(p) = 1 \text{ if } p \in \mathbb{R}^n \setminus B_1(0).$$

Given a positive real number δ, consider the homothety h_δ of $\mathbb{R}^n$, and let

$$h := \left(\psi \circ h_{1/\delta}\right) f.$$

Then, obviously, $h \in C^\infty(\mathbb{R}^n)$, and h vanishes on a 0-neighbourhood. So it is sufficient to show that if $|\alpha| \leqq m$,

$$\sup_{p \in \mathbb{R}^n} |D^\alpha h(p) - D^\alpha f(p)| \to 0 \quad \text{as } \delta \to 0. \tag{$*$}$$

Since $h(p) = f(p)$ if $\|p\| > \delta$, it follows that

$$\begin{aligned} \sup_{p \in \mathbb{R}^n} |D^\alpha h(p) - D^\alpha f(p)| &= \sup_{\|p\| \leqq \delta} |D^\alpha h(p) - D^\alpha f(p)| \\ &\leqq \sup_{\|p\| \leqq \delta} |D^\alpha h(p)| + \sup_{\|p\| \leqq \delta} |D^\alpha f(p)|. \end{aligned}$$

By our assumption, we have $D^\alpha f(0) = 0$ for $|\alpha| \leqq m$; thus

$$\sup_{\|p\| \leqq \delta} |D^\alpha f(p)| \to 0 \quad \text{as } \delta \to 0. \tag{$**$}$$

Now we consider the function $D^\alpha h$. Using 2.15, we obtain

$$\begin{aligned} D^\alpha h = D^\alpha\left(\left(\psi \circ h_{1/\delta}\right) f\right) &= \sum_{\mu+\nu=\alpha} \binom{\alpha}{\nu} D^\nu \left(\psi \circ h_{1/\delta}\right) D^\mu f \\ &= \sum_{\mu+\nu=\alpha} \binom{\alpha}{\nu} \delta^{-nu} \left((D^\nu \psi) \circ h_{1/\delta}\right) D^\mu f. \end{aligned}$$

Since ψ is constant outside $B_1(0)$, the function $D^\nu \psi$ is bounded. If

$$C_\nu := \sup_{p \in \mathbb{R}^n} |D^\nu \psi(p)|, \quad C := \max_\nu \binom{\alpha}{\nu} C_\nu,$$

then we get the following estimation:

$$|D^\alpha h| \leqq C \sum_{\mu+\nu=\alpha} \delta^{-|\nu|} |D^\mu f|.$$

As f is m-flat at 0, $D^\mu f$ is $(m - |\mu|)$-flat at the origin. Thus, using Landau's symbol $o(\cdot)$,

$$\sup_{\|p\| \leqq \delta} |D^\mu f(p)| = o\left(\delta^{m-|\mu|}\right),$$

therefore

$$\sup_{\|p\| \leqq \delta} |D^\alpha h(p)| = o\left(\sum_{\mu+\nu=\alpha} \delta^{m-|\mu|-|\nu|}\right) = o\left(\delta^{m-|\alpha|}\right),$$

and hence

$$\sup_{p\in\mathbb{R}^n} |D^\alpha h(p) - D^\alpha f(p)| \leqq o\left(\delta^{m-|\alpha|}\right) + \sup_{\|p\|\leqq\delta} |D^\alpha f(p)|.$$

In view of ($**$), this implies the desired relation ($*$). □

Lemma 2.20 *Let $\mathcal{U} \subset \mathbb{R}^n$ be a (nonempty) open set. Suppose $D : C^\infty(\mathcal{U}) \to C^\infty(\mathcal{U})$, $f \mapsto Df$ is a linear mapping which decreases supports, that is*

$$\operatorname{supp}(Df) \subset \operatorname{supp}(f)$$

for all functions $f \in C^\infty(\mathcal{U})$. Then for each point $p \in \mathcal{U}$ there exists a relatively compact neighbourhood $\mathcal{V}$ of p $\left(\bar{\mathcal{V}} \subset \mathcal{U}\right)$, a positive integer m and a positive real number C such that

$$\|Du\|_0 \leqq C\|u\|_m$$

holds for any function $u \in C_c^\infty(\mathcal{V} \smallsetminus \{p\})$ ($\| \ \|_k$ $(k \in \mathbb{N})$ is the seminorm introduced in 2.18).

Note Recall that any function $f \in C_c^\infty(\mathcal{V} \smallsetminus \{p\})$ is identified with a function $\tilde{f} \in C^\infty(\mathcal{U})$ according to 2.16. Since D is support-decreasing, we may apply it to f by the formula $Df := D\tilde{f}$.

Proof of the lemma. Suppose the contrary: there is a point $p \in \mathcal{U}$ such that for any relatively compact nbd $\mathcal{V} \subset \mathcal{U}$ of p, any positive integer m and positive real number C, there is a function $u \in C_c^\infty(\mathcal{V} \smallsetminus \{p\})$ such that

$$\|Du\|_0 > C\|u\|_m.$$

1st step Choose a relatively compact set $\mathcal{U}_0 \subset \mathcal{U}$ in $\mathcal{N}(p)$, and let $m := 1$, $C := 2^2$. By our assumption, there is a function $u_1 \in C_c^\infty(\mathcal{U}_0 \smallsetminus \{p\})$ such that

$$\|Du_1\|_0 > 2^2\|u_1\|_1.$$

Let $\mathcal{U}_1 := \{q \in \mathcal{U}_0 | u_1(q) \neq 0\}$. Then $\mathcal{U}_0 \smallsetminus \overline{\mathcal{U}_1} \in \mathcal{N}(p)$, so, using our assumption again, there is a function $u_2 \in C_c^\infty\left(\mathcal{U}_0 \smallsetminus \overline{\mathcal{U}_1} \smallsetminus \{p\}\right)$ such that

$$\|Du_2\|_0 > 2^4\|u_2\|_2.$$

Now we define $\mathcal{U}_2 := \left\{q \in \mathcal{U}_0 \smallsetminus \overline{\mathcal{U}_1} \middle| u_2(q) \neq 0\right\}$, and we repeat our argument. Thus, by induction, we obtain a sequence $(\mathcal{U}_k)_{k\in\mathbb{N}^*}$ of open sets and $(u_k)_{k\in\mathbb{N}^*}$ of functions such that

$$\begin{aligned}
&\overline{\mathcal{U}_k} \subset \mathcal{U}_0 \smallsetminus \{p\} \quad (k \in \mathbb{N}^*); \quad \overline{\mathcal{U}_k} \cap \overline{\mathcal{U}_\ell} = \emptyset \text{ if } k \neq \ell;\\
&u_k \in C_c^\infty\left(\mathcal{U}_0 \smallsetminus \overline{\mathcal{U}_1} \smallsetminus \ldots \smallsetminus \overline{\mathcal{U}_{k-1}} \smallsetminus \{p\}\right) \subset C_c^\infty(\mathcal{U});\\
&\mathcal{U}_k = \left\{q \in \mathcal{U}_0 \smallsetminus \overline{\mathcal{U}_1} \smallsetminus \ldots \smallsetminus \overline{\mathcal{U}_{k-1}} \middle| u_k(q) \neq 0\right\} \quad (k \in \mathbb{N}^*),
\end{aligned}$$

and, finally,

$$\|Du_k\|_0 > 2^{2k}\|u_k\|_k. \tag{$*$}$$

2nd step Consider the function

$$u := \sum_{k \in \mathbb{N}^*} \frac{2^{-k}}{\|u_k\|_k} u_k.$$

Since at each point of $\mathcal{U}_0$, at most one member of the right-hand side differs from zero, u is well-defined. We claim that u *is smooth on its domain*. To prove this, we have to show that each point $p \in \mathcal{U}_0$ has a neighbourhood on which u is smooth. We divide the points of $\mathcal{U}_0$ into three disjoint classes in the following way. First, let

$$\mathcal{V}_1 := \bigcup_{k=1}^{\infty} \mathcal{U}_k.$$

Then any point $p \in \mathcal{V}_1$ is contained in some $\mathcal{U}_k$, on which u is obviously smooth. Next, let $\mathcal{V}_2$ be the set of points in $\mathcal{U}_0$ which have a neighbourhood $\mathcal{V}$ intersecting with at most one of the $\mathcal{U}_k$'s. Then u is smooth on $\mathcal{V}$ as well. Finally, let $\mathcal{V}_3$ be the set of points in $\mathcal{U}_0$ whose every neighbourhood intersects with infinitely many of the $\mathcal{U}_k$'s. The only difficulty is to verify the smoothness of u in a neighbourhood of such points. By the Bolzano – Weierstraß theorem, $\mathcal{V}_3$ cannot be empty.

Let $p \in \mathcal{V}_3$. First we show that u is continuous at p. Since $p \notin \mathcal{U}_k$ ($k \in \mathbb{N}^*$), $u(p) = 0$. Let $\varepsilon > 0$ be arbitrary, and

$$k := \left\lceil \log_2 \frac{1}{\varepsilon} \right\rceil$$

(where the symbol $\lceil\ \rceil$ denotes upper integer part). Now we define

$$\delta := \min_{1 \leqq i \leqq k} d\left(p, \overline{\mathcal{U}_i}\right).$$

Then $\delta > 0$, otherwise $p \in \mathcal{V}_2$ would follow. If $q \in B_\delta(p)$, we have either $u(q) = 0$ or $q \in \mathcal{U}_\ell$ for some $\ell > k$, thus

$$|u(q)| = \left| \frac{2^{-\ell}}{\|u_\ell\|_\ell} u_\ell(q) \right| \leqq \left| \frac{2^{-\ell}}{\|u_\ell\|_0} u_\ell(q) \right| \leqq 2^{-\ell} < 2^{-k} \leqq 2^{-\log_2(1/\varepsilon)} = \varepsilon,$$

therefore u is continuous at p.

Now we proceed by induction. Let m be a fixed positive integer, and suppose that all partial derivatives of u up to order $m-1$ exist and are continuous, and they all vanish at p. Let $\alpha \in \mathbb{N}^n$ be a multi-index such that $|\alpha| = m-1$, and $i \in \{1, \dots, n\}$. To show that $D_i D^\alpha u(p)$ exists, we have to consider the following limit:

$$\lim_{t \to 0} \frac{D^\alpha u(p + te_i) - D^\alpha u(p)}{t} = \lim_{t \to 0} \frac{D^\alpha u(p + te_i)}{t}.$$

Let $\varepsilon > 0$ be arbitrary. Now we define k *almost* in the same way as in the previous part of the proof:

$$k := \max\left\{ m, \left\lceil \log_2 \frac{1}{\varepsilon} \right\rceil \right\}.$$

Let $t \in]0, \delta[$, where δ is defined *exactly* in the same way as above. Then either $D^\alpha u(p + te_i) = 0$, or $p + te_i \in \mathcal{U}_\ell$ for some $\ell > k$. Let

$$t_0 := \sup\{s \in]0, t[\,|p + se_i \notin \mathcal{U}_\ell\}.$$

Then the function $s \mapsto D^\alpha u(p + se_i)$ is continuous on $[t_0, t]$ and differentiable on $]t_0, t[$, so, by Lagrange's mean value theorem, there is some $\xi \in]t_0, t[$ such that

$$\begin{aligned}(s \mapsto D^\alpha u(p + se_i))'(\xi) &= D_i D^\alpha u(p + \xi e_i)\\ &= \frac{D^\alpha u(p + te_i) - D^\alpha u(p + t_0 e_i)}{t - t_0} = \frac{D^\alpha u(p + te_i)}{t - t_0}.\end{aligned}$$

Thus we have the following estimation:

$$\begin{aligned}&\left|\frac{D^\alpha u(p + te_i)}{t}\right| \leqq \left|\frac{D^\alpha u(p + te_i)}{t - t_0}\right| = |D_i D^\alpha u(p + \xi e_i)|\\ &= \left|\frac{2^{-\ell}}{\|u_\ell\|_\ell} D_i D^\alpha u_\ell(p + \xi e_i)\right| \leqq \left|\frac{2^{-\ell}}{\|D_i D^\alpha u_\ell\|_0} D_i D^\alpha u_\ell(q)\right|\\ &\qquad\qquad\qquad \leqq 2^{-\ell} < 2^{-k} \leqq 2^{-\log_2(1/\varepsilon)} = \varepsilon.\end{aligned}$$

If $t \in]-\delta, 0[$, then we proceed in the same way to obtain

$$D_i D^\alpha u(p) = \lim_{t \to 0} \frac{D^\alpha u(p + te_i) - D^\alpha u(p)}{t} = 0.$$

Therefore $D_i D^\alpha u$ exists at every point of $\mathcal{U}_0$. Its continuity is shown in the same way as that of u.

3rd step We obviously have

$$u \restriction \mathcal{U}_k = 2^{-k}(\|u_k\|_k)^{-1} u_k. \tag{$**$}$$

Since D is linear and support-decreasing, it follows that

$$(Du) \restriction \mathcal{U}_k = 2^{-k}(\|u_k\|_k)^{-1}(Du_k) \restriction u_k.$$

Thus, taking into account relations $(*)$ and $(**)$, we conclude that there is a point $p_k \in \mathcal{U}_k$ such that

$$|Du(p_k)| = 2^{-k}(\|u_k\|_k)^{-1}|Du_k(p_k)| > 2^{-k}(\|u_k\|_k)^{-1} \cdot 2^{2k}\|u_k\|_k = 2^k.$$

On the other hand, the function Du is continuous, and its support is compact, hence it is bounded. This contradicts the above assertion. $\square$

Lemma 2.21 *Suppose $\mathcal{U} \subset \mathbb{R}^n$ is a nonempty open set, and let $\mathcal{V}$ be a relatively compact open set contained in $\mathcal{U}$. Consider a support-decreasing linear mapping $D : C^\infty(\mathcal{U}) \to C^\infty(\mathcal{U})$. Assume there is a positive integer m and a positive real number C such that*

$$\|Du\|_0 \leqq C\|u\|_m \tag{$*$}$$

for all $u \in C_c^\infty(\mathcal{V})$. Then *(i)* $Du(p) = 0$ *whenever* u *is* m*-flat at* p,

(ii) *there exist smooth functions* a_α $(\alpha \in \mathbb{N}^n,\ |\alpha| \leqq m)$ *in* $C^\infty(\mathcal{V})$ *such that for each* $u \in C_c^\infty(\mathcal{V})$, $p \in \mathcal{V}$ *we have*

$$Du(p) = \sum_{|\alpha| \leqq m} a_\alpha(p)(D^\alpha u)(p).$$

Proof. (i) By Lemma 2.19, there is a sequence $(u_n)_{n\in\mathbb{N}}$ of functions in $C_c^\infty(\mathcal{V})$ such that u_n vanishes in a neighbourhood of p for each $n \in \mathbb{N}$, and

$$\lim_{n\to\infty} \|u_n - u\|_m = 0.$$

Taking into account our condition $(*)$, this implies that $(Du_n)_{n\in\mathbb{N}}$ converges uniformly to Du on $\mathcal{V}$. Since $\operatorname{supp}(Du_n) \subset \operatorname{supp}(u_n)$, and u_n vanishes near p, we have $Du_n(p) = 0$ for each $n \in \mathbb{N}$. Hence

$$Du(p) = \lim_{n\to\infty} (Du_n)(p) = 0,$$

as we claimed.

(ii) To prove the second statement, consider for an arbitrarily chosen point $p \in \mathcal{U}$ and multi-index $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ the polynomial

$$\mu_{\alpha,p}(\xi) := \left(\xi^1 - p^1\right)^{\alpha_1} \cdot \ldots \cdot (\xi^n - p^n)^{\alpha_n}$$

($\xi = \left(\xi^1, \dots, \xi^n\right)$ is a symbol). Then $\mu_{\alpha,p}$ can be viewed as a smooth function in $C^\infty(\mathcal{U})$, and the functions

$$\eta_\alpha : p \in \mathcal{U} \mapsto \eta_\alpha(p) := (D\mu_{\alpha,p})(p) \quad (\alpha \in \mathbb{N}^n)$$

also belong to $C^\infty(\mathcal{U})$. Now let $u \in C_c^\infty(\mathcal{V})$, and define

$$f := u - \sum_{|\alpha| \leqq m} \frac{1}{\alpha!}(D^\alpha u)(p)\mu_{\alpha,p}.$$

Then for each $\beta \in \mathbb{N}^n$, $|\beta| \leqq m$ we have

$$D^\beta f = D^\beta u - \sum_{|\alpha| \leqq m} \frac{1}{\alpha!}(D^\alpha u)(p)\frac{\alpha!}{(\alpha-\beta)!}\mu_{\alpha-\beta,p}.$$

Since

$$\mu_{\alpha-\beta,p}(p) = \begin{cases} 0 & \text{if } \beta \neq \alpha, \\ 1 & \text{if } \beta = \alpha, \end{cases}$$

it follows that

$$D^\beta f(p) = 0, \quad |\beta| \leqq m;$$

i.e., the function f is m-flat at p. Thus, by part (i), $Df(p) = 0$, therefore

$$Du(p) = \sum_{|\alpha| \leqq m} \frac{1}{\alpha!}(D^\alpha u)(p)\eta_\alpha(p);$$

so with the help of the smooth functions $a_\alpha := \frac{1}{\alpha!}\eta_\alpha$ ($\alpha \in \mathbb{N}^n$, $|\alpha| \leqq m$) Du can be represented in the desired form.

□

Now we introduce a more sophisticated version of the concept of a linear differential operator mentioned in 2.14.

Definition 2.22 Let $\mathcal{U} \subset \mathbb{R}^n$ be a nonempty open subset. A *differential operator* on $\mathcal{U}$ is a linear mapping $D : C^\infty(\mathcal{U}) \to C^\infty(\mathcal{U})$ with the following property: for each relatively compact open set $\mathcal{V}$ whose closure in contained in $\mathcal{U}$ there exists a finite family of functions $a_\alpha \in C^\infty(\mathcal{V})$ ($\alpha \in \mathbb{N}^n$) such that for each $u \in C^\infty(\mathcal{V})$,

$$Du = \sum_\alpha a_\alpha(D^\alpha u).$$

Note Differential operators in this more general sense also have a well-defined *order* locally, according to 2.14.

Having this concept, we are ready to formulate and prove the main result of this subsection.

Theorem 2.23 (local Peetre theorem) *Let $\mathcal{U} \subset \mathbb{R}^n$ be a nonempty open subset. If $D : C^\infty(\mathcal{U}) \to C^\infty(\mathcal{U})$ is a support-decreasing linear mapping, then D is a differential operator on $\mathcal{U}$. Conversely, any differential operator is support decreasing.*

Proof. The converse statement is clearly true. To prove the direct statement, consider a linear mapping $D : C^\infty(\mathcal{U}) \to C^\infty(\mathcal{U})$ with the property

$$\operatorname{supp}(Du) \subset \operatorname{supp}(u), \quad u \in C^\infty(\mathcal{U}).$$

Let $\mathcal{V} \subset \mathcal{U}$ be a relatively compact open set such that $\bar{\mathcal{V}} \subset \mathcal{U}$. According to 2.20, for each point $p \in \mathcal{V}$ there is a neighbourhood $\mathcal{U}_p \subset \mathcal{U}$ of p, a positive integer m_p and a positive real number C_p such that

$$\|Du\|_0 \leqq C_p\|u\|_{m_p}$$

holds for any function $u \in C_c^\infty(\mathcal{U}_p \setminus \{p\})$. The family $(\mathcal{U}_p)_{p\in\mathcal{V}}$ is an open covering of $\bar{\mathcal{V}}$, so by its compactness, there are finitely many points $p_1, \ldots, p_k$ in $\mathcal{V}$ such that the corresponding open sets $\mathcal{U}_1, \ldots, \mathcal{U}_k$ still cover $\bar{\mathcal{V}}$. Let

$$m := \max\{m_{p_i} | i \in \{1, \ldots, k\}\} \quad \text{and} \quad C := \max\{C_{p_i} | i \in \{1, \ldots, k\}\}.$$

Then for each $i \in \{1, \ldots, k\}$ and $u \in C_c^\infty(\mathcal{U}_i \setminus \{p_i\})$ we have

$$\|Du\|_0 \leqq C\|u\|_m. \tag{$*$}$$

We show that there is also a positive constant $\tilde{C}$ such that

$$\|Du\|_0 \leqq \tilde{C}\|u\|_m$$

holds for any function $u \in C_c^\infty(\mathcal{V} \smallsetminus \{p_1, \dots, p_k\})$. Using Lemma 2.12, let $(f_i)_{i=1}^{k+1}$ be a partition of unity subordinate to the open covering $\big(\mathcal{U}_1, \dots, \mathcal{U}_k, \mathcal{U} \smallsetminus \bar{\mathcal{V}}\big)$ of $\mathcal{U}$. Then each function $u \in C_c^\infty(\mathcal{V} \smallsetminus \{p_1, \dots, p_n\})$ can be written in the form

$$u = \sum_{i=1}^{k+1} f_i u = \sum_{i=1}^{k} f_i u,$$

and $(*)$ holds for every member of the sum, therefore

$$\begin{aligned}\|Du\|_0 &= \left\|D\left(\textstyle\sum_{i=1}^k f_i u\right)\right\|_0 = \left\|\textstyle\sum_{i=1}^k D(f_i u)\right\|_0 \\ &\leqq \sum_{i=1}^{k} \|D(f_i u)\|_0 \leqq \sum_{i=1}^{k} C\|f_i u\|_m = C\sum_{i=1}^{k} \|f_i u\|_m.\end{aligned}$$

Thus it is enough to show that there are positive constants C_i such that

$$\|f_i u\|_m \leqq C_i \|u\|_m \quad (i = 1, \dots, k)$$

which are independent of u (and may depend on the partition of unity chosen, however). By the generalized Leibniz rule (2.15), we have

$$\begin{aligned}\|f_i u\|_m &= \sum_{|\alpha\| \leqq m} \sup_{p \in \mathcal{V}} |D^\alpha(f_i u)(p)| \\ &= \sum_{|\alpha\| \leqq m} \sup_{p \in \mathcal{V}} \left|\textstyle\sum_{\mu+\nu=\alpha} \binom{\alpha}{\nu}(D^\nu f_i)(p)(D^\mu u)(p)\right| \\ &\leqq \sum_{|\alpha\| \leqq m} \sup_{p \in \mathcal{V}} \sum_{\mu+\nu=\alpha} \binom{\alpha}{\nu} |D^\nu f_i(p)||D^\mu u(p)| \\ &\leqq \sum_{|\alpha\| \leqq m} \sum_{\mu+\nu=\alpha} \binom{\alpha}{\nu} \sup_{p \in \mathcal{V}} |D^\nu f_i(p)||D^\mu u(p)| \\ &\leqq \sum_{|\alpha\| \leqq m} \sum_{\mu+\nu=\alpha} \binom{\alpha}{\nu} \left(\sup\nolimits_{p \in \mathcal{V}} |D^\nu f_i(p)|\right) \left(\sup\nolimits_{p \in \mathcal{V}} |D^\mu u(p)|\right) \\ &= \sum_{|\mu| \leqq m} \left[\textstyle\sum_{|\mu+\nu| \leqq m} \binom{\mu+\nu}{\nu} \sup_{p \in \mathcal{V}} |D^\nu f_i(p)|\right] \sup_{p \in \mathcal{V}} |D^\mu u(p)|.\end{aligned}$$

Letting

$$C_i := \sum_{|\mu+\nu| \leqq m} \binom{\mu+\nu}{\nu} \sup_{p \in \mathcal{V}} |D^\nu f_i(p)|,$$

we obtain

$$\|f_i u\|_m \leqq C_i \sum_{|\mu|\leqq m} \sup_{p\in\mathcal{V}} |D^\mu u(p)| = C_i \|u\|_m,$$

thus the desired estimation is valid for each term. Finally, if

$$\tilde{C} := C \sum_{i=1}^{k} C_i,$$

then we conclude

$$\|Du\|_0 \leqq C \sum_{i=1}^{k} \|f_i u\|_m \leqq C \sum_{i=1}^{k} C_i \|u\|_m = \tilde{C}\|u\|_m.$$

Lemma 2.21 then implies that for each $p \in \mathcal{V} \setminus \{p_1, \dots, p_n\}$,

$$Du(p) = \sum_{|\alpha|\leqq m} a_\alpha(p)(D^\alpha u)(p), \quad a_\alpha \in C^\infty(\mathcal{V}).$$

However, both functions Du and $\sum_{|\alpha|\leqq m} a_\alpha(D^\alpha u)$ are continuous, so the desired relation is valid at every point of $\mathcal{V}$. ☐

Transition from local to global

1st step: manifolds and bundles Since we have a chain rule for C^1-mappings between locally convex TVSs, the concept of a manifold modeled on such a vector space can be introduced without any difficulty: we may follow the well paved path of the finite-dimensional theory.

Let M be a Hausdorff space and V a locally convex TVS. By a *V-chart* on M we mean a pair $(\mathcal{U}, x)$, where $\mathcal{U} \subset M$ is an open subset of M, and x is a homeomorphism of $\mathcal{U}$ onto an open subset of V. Two charts, $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$ are said to be *smoothly compatible* if the transition mapping

$$y \circ x^{-1} : x(\mathcal{U} \cap \mathcal{V}) \to y(\mathcal{U} \cap \mathcal{V})$$

is a smooth mapping between open subsets of V in Michel – Bastiani's sense, or $\mathcal{U} \cap \mathcal{V} = \emptyset$. A *$V$-atlas* on M is a family $\mathcal{A} = (\mathcal{U}_i, x_i)_{i\in I}$ of pairwise compatible V-charts of M such that the sets $\mathcal{U}_i$ form an open covering of M. A *smooth V-structure on M* is a maximal V-atlas, and a *smooth V-manifold* is a Hausdorff space endowed with a smooth V-structure. A smooth V-manifold is said to be a *Fréchet manifold*, a *Banach manifold* and an *n-manifold*, resp., if the model space V is a Fréchet space, a Banach space and an n-dimensional (real) vector space, resp. In the latter case, without loss of generality, the Euclidean n-space $\mathbb{R}^n$ can be chosen as a model space.

The concept of smoothness of a mapping between a V-manifold M and a W-manifold N can formally be defined in the same way as in the finite-dimensional case: a mapping $\varphi : M \to N$ is said to be *smooth* if it is continuous and, for every chart $(\mathcal{U}, x)$ on M and $(\mathcal{V}, y)$ on N, the mapping

$$y \circ \varphi \circ x^{-1} : x\left(\varphi^{-1}(\mathcal{V}) \cap \mathcal{U}\right) \subset V \to W$$

is Michel – Bastiani smooth. If, in addition, φ has a smooth inverse, then it is called a *diffeomorphism*. The space of all smooth mappings from M to N will be denoted by $C^\infty(M,N)$. In particular, $C^\infty(M)$ stands for the algebra of smooth functions on M, and

$$C_c^\infty(M) := \{f \in C^\infty(M) | \operatorname{supp}(f) \text{ is compact}\}.$$

Let $\pi : E \to M$ be a smooth mapping between smooth manifolds modeled on some locally convex TVSs, and let V be a fixed locally convex space. $\pi : E \to M$ is said to be a *vector bundle* with *typical fibre* V if the following conditions are satisfied:

(i) every *fibre* $E_p := \pi^{-1}(p)$, $p \in M$, is a locally convex TVS;

(ii) for each point $p \in M$ there is a neighbourhood $\mathcal{U}$ of p and a diffeomorphism $\Phi : \mathcal{U} \times V \to \pi^{-1}(\mathcal{U})$ such that $\pi \circ \Phi = \mathrm{pr}_1$, where $\mathrm{pr}_1 : \mathcal{U} \times V \to \mathcal{U}$ is the projection to the first factor, and the mappings

$$\Phi_q : V \to E_q, \quad v \mapsto \Phi_q(v) := \Phi(q,v), \quad q \in \mathcal{U}$$

are toplinear isomorphisms.

Further terminology: M is the *base manifold*, E is the *total manifold*, and π is the *projection* of the bundle. Φ is a *trivialization* of $\pi^{-1}(\mathcal{U})$ (or a *local trivialization* of E). For a vector bundle $\pi : E \to M$, we shall use both the abbreviation π (which is unambiguous) and the shorthand E (which may be ambiguous), depending on what we wish to emphasize. If the base manifold is finite dimensional, and the typical fibre is a k-dimensional vector space, then we shall speak of a vector bundle of *rank* k.

Let $\pi_1 : E_1 \to M_1$ and $\pi_2 : E_2 \to M_2$ be vector bundles. A smooth mapping $\varphi : E_1 \to E_2$ is called *fibre preserving* if $\pi_1(z_1) = \pi_1(z_2)$ implies $\pi_2(\varphi((z_1)) = \pi_2(\varphi(z_2))$ for all $z_1, z_2 \in E$. Then φ induces a smooth mapping $\underline{\varphi} : M_1 \to M_2$ such that $\pi_2 \circ \varphi = \underline{\varphi} \circ \pi_1$. A fibre preserving mapping $\varphi : E_1 \to E_2$ is said to be a *bundle map* if it restricts to continuous linear mappings $\varphi_1 : (E_1)_p \to (E_2)_{\underline{\varphi}(p)}$, $p \in M$. If, in addition, $M_1 = M_2 =: M$, and $\underline{\varphi} = 1_M$, then φ is called a *strong bundle map*.

Suppose $\pi : E \to M$ is a vector bundle with typical fibre V and $f : N \to M$ is a smooth mapping. For each $q \in N$, let $(f^*E)_q := \{q\} \times E_{f(q)}$ be endowed with the vector space structure inherited from $E_{f(q)}$:

$$(q, z_1) + (q, z_2) := (q, z_1 + z_2), \quad \lambda(q,z) := (q, \lambda z)$$

$\big(z_1, z_2, z \in E_{f(q)},\ \lambda \in \mathbb{R}\big)$. If

$$f^*E := \bigcup_{q \in N} (f^*E)_q = \{(q,z) \in N \times E | f(q) = \pi(z)\} =: N \times_M E,$$

then f^*E carries a unique smooth structure which makes it a vector bundle with base manifold N, projection $\pi_1 := \mathrm{pr}_1 \restriction N \times_M E$ and typical fibre V. If $\Phi : \mathcal{U} \times V \to \pi^{-1}(\mathcal{U})$ is a local trivialization of E, then the mapping

$$f^*\Phi : f^{-1}(\mathcal{U}) \times V \to \pi_1^{-1}(f^{-1}(\mathcal{U})), \quad (q,v) \mapsto (q, \Phi(f(q), v))$$

is a local trivialization for f^*E. The vector bundle $\pi_1 : f^*E \to N$ so obtained is said to be the *pull-back* of π over f, and it is also denoted by $f^*\pi$. Note that the mapping

$\pi_2 := \mathrm{pr}_2 \restriction N \times_M E$ is a bundle map from f^*E to E which induces the given mapping f between the base manifolds.

For our next remarks, let us fix a vector bundle $\pi : E \to M$. A *section* of π is a smooth mapping $\sigma : M \to E$ with $\pi \circ \sigma = 1_M$; thus $\sigma(p) \in E_p$ for all $p \in M$. Similarly, a section of π over an open subset $\mathcal{U} \subset M$ is a smooth mapping $\sigma : \mathcal{U} \to E$ with $\pi \circ \sigma = 1_{\mathcal{U}}$. The *support* of a section $\sigma : \mathcal{U} \to E$ is $\mathrm{supp}(\sigma) := \overline{\{p \in \mathcal{U} | \sigma(p) \neq 0\}}$ (the closure is meant in $\mathcal{U}$). We denote by $\Gamma(\pi)$ (or $\Gamma(E)$) the set of sections of π. $\Gamma(\mathcal{U}, E)$ stands for the set of sections of E over $\mathcal{U}$; $\Gamma_c(\mathcal{U}, E) = \{\sigma \in \Gamma(\mathcal{U}, E) | \, \mathrm{supp}(\sigma) \text{ is compact}\}$. $\Gamma(\mathcal{U}, E)$, in particular, $\Gamma(\pi)$ is surely nonempty: we have the *zero section* $p \in \mathcal{U} \mapsto o(p) := 0_p :=$ *the zero vector of* E_p. $\overset{\circ}{E} := \cup_{p \in M}(E_p \setminus \{0_p\})$ will denote the *deleted bundle* for E, and $\overset{\circ}{\pi} := \pi \restriction \overset{\circ}{E}$. Sections (local sections with common domain) can be added to each other and multiplied by smooth functions on M using the standard pointwise definitions. These two operations make $\Gamma(\pi)$ and $\Gamma(\mathcal{U}, E)$ a $C^\infty(M)$-module and a $C^\infty(\mathcal{U})$-module, resp. In particular, $\Gamma(\pi)$ and $\Gamma(\mathcal{U}, E)$ (as well as $\Gamma_c(\mathcal{U}, E)$) are real vector spaces. Under some assumptions on the topological and the smooth structure of the base manifold and the TVS structure of the typical fibre, the spaces of sections can be endowed with 'similarly nice' TVS structure. For more information on this subtle problem see Kriegl – Michor's monograph [18, section 30]. In the next subsection we shall briefly discuss the finite dimensional case.

Now suppose that a smooth mapping $f : N \to M$ is also given, and consider the pullback bundle $\pi_1 : N \times_M E \to N$. A smooth mapping $S : N \to N \times_M E$ is a section of π_1 if and only if there is a smooth mapping $\underline{S} : N \to E$ such that $\pi \circ \underline{S} = f$ and $S(q) = (q, \underline{S}(q))$ for all $q \in N$; $\underline{S}$ is mentioned as the *principal part* of the section S. A smooth mapping $\underline{S} : N \to E$ satisfying $\pi \circ \underline{S} = f$ is called a *section of* E *along* f. Identifying a section of f^*E with its principal part, we get a canonical module isomorphism between $\Gamma(f^*E)$ and the module $\Gamma_f(E) = \Gamma_f(\pi)$ of sections of E along f.

2nd step: tangent bundle The crucial step in transporting of calculus from the model space V to a smooth V-manifold M is the construction of tangent vectors. Choose a point p of M and consider all triples of the form $(\mathcal{U}, x, a)$, where $(\mathcal{U}, x)$ is a chart around p and $a \in V$. The relation $\sim$ defined by

$$(\mathcal{U}, x, a) \sim (\mathcal{V}, y, b) :\Longleftrightarrow \left(y \circ x^{-1}\right)'(x(p))(a) = b$$

is an equivalence relation. The $\sim$-equivalence class $[(\mathcal{U}, x, a)]$ of a triple $(\mathcal{U}, x, a)$ is said to be a *tangent vector to* M *at* p. The set of tangent vectors to M at p is the *tangent space* to M at p; it will be denoted by T_pM. A chart $(\mathcal{U}, x)$ around p determines a bijection $\vartheta_p : T_pM \to V$ by the rule

$$v \in T_pM \mapsto \vartheta_p(v) := a \in V \quad \text{if } (\mathcal{U}, x, a) \in v.$$

There is a unique locally convex TVS structure on T_pM which makes the bijection ϑ_p a toplinear isomorphism. To be explicit, the linear structure of T_pM is given by

$$\lambda v + \mu w := \vartheta_p^{-1}(\lambda\vartheta_p(v) + \mu\vartheta_p(w)); \quad v, w \in T_pM; \quad \lambda, \mu \in \mathbb{R}.$$

The TVS structure so obtained on T_pM does not depend on the choice of $(\mathcal{U}, x)$, since if $(\mathcal{V}, y)$ is another chart around p, and $\eta_p : T_pM \to V$ is associated with $(\mathcal{V}, y)$, then $\eta_p \circ \vartheta_p^{-1} = \left(y \circ x^{-1}\right)'(x(p))$, which is a toplinear isomorphism.

Let $TM := \cup_{p\in M} T_pM$ (disjoint union), and define the mapping $\tau : TM \to M$ by $\tau(v) := p$ if $v \in T_pM$. *There is a unique smooth structure on TM which makes $\tau : TM \to M$ into a vector bundle with fibres $(TM)_p = T_pM$ and typical fibre V.* To indicate the construction of this smooth structure, let $(\mathcal{U}, x)$ be a chart on M, and assign to each $v \in \tau^{-1}(\mathcal{U}) \subset TM$ the pair $\big(x(\tau(v)), \vartheta_{\tau(v)}(v)\big) \in V \times V$. Thus we get a bijective mapping

$$\tau_{\mathcal{U}} := \big(x \circ \tau, \vartheta_{\tau(\cdot)}\big) : \tau^{-1}(\mathcal{U}) \to x(\mathcal{U}) \times V \subset V \times V.$$

As $(\mathcal{U}, x)$ runs over all charts of an atlas of M, the pairs $\big(\tau^{-1}(\mathcal{U}), \tau_{\mathcal{U}}\big)$ form an atlas for TM and make it into a smooth manifold modeled on $V \times V$ such that the topology of TM is the finest topology for which each mapping $\tau_{\mathcal{U}}$ is a homeomorphism. (For details see [8, 2.3].) The vector bundle so obtained is called the *tangent bundle* of M, and it is denoted by τ, TM or τ_M. If, in particular, $\mathcal{U}$ is an open subset of the model space V, then $\mathcal{U}$ (as well as V) can be regarded as a smooth manifold modeled on V. In this case there is a canonical identification $T\mathcal{U} \cong \mathcal{U} \times V$, which will be frequently used, without further mention.

Now take two manifolds M and N, modeled on V and W, respectively. Let $f : M \to N$ be a smooth mapping and $p \in M$. Choose charts $(\mathcal{U}, x)$ around p and $(\mathcal{V}, y)$ around $f(p)$ such that $f(\mathcal{U}) \subset \mathcal{V}$. Let $\vartheta_p : T_pM \to V$ and $\eta_{f(p)} : T_{f(p)}N \to W$ be the isomorphisms associated with $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$. Then

$$(f_*)_p := \eta_{f(p)}^{-1} \circ \big(y \circ f \circ x^{-1}\big)' (x(p)) \circ \vartheta_p : T_pM \to T_{f(p)}N$$

is a well-defined linear mapping, called the *tangent map* of f *at* p. ('Well-defined' means that $(f_*)_p$ does not depend on the choice of $(\mathcal{U}, x)$ and $(\mathcal{V}, y)$.) Having the fibrewise tangent maps, we associate to $f : M \to N$ the bundle map

$$f_* : TM \to TN, \quad f_* \restriction T_pM := (f_*)_p.$$

Let, in particular, $f : M \to W$ be a smooth mapping. Then the mapping

$$df := \mathrm{pr}_2 \circ f_* : TM \to W \times W \to W$$

is called the *differential* of f. By restriction, it leads to a continuous linear mapping

$$df(p) = d_pf := df \restriction T_pM \to W,$$

for each $p \in M$.

The sections of the tangent bundle $\tau : TM \to M$ are said to be *vector fields* on M. We denote their $C^\infty(M)$-module by $\mathfrak{X}(M)$ rather than $\Gamma(\tau)$ or $\Gamma(TM)$. Any vector field X on M induces a derivation ϑ_X of the real algebra $C^\infty(M)$ by the rule

$$f \in C^\infty(M) \mapsto \vartheta_X(f) = X.f := df \circ X.$$

It may be shown that if $X, Y \in \mathfrak{X}(M)$, then there exists a unique vector field $[X, Y]$ on M such that on each open subset $\mathcal{U}$ of M we have

$$\vartheta_{[X,Y]}(f) = \vartheta_X(\vartheta_Y(f)) - \vartheta_Y(\vartheta_X(f))$$

for all $f \in C^\infty(\mathcal{U})$, and the mapping $[\cdot,\cdot] : \mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ makes $\mathfrak{X}(M)$ a (real) Lie algebra. For an accurate proof of these claims we refer to [8]. Lang's monograph [20] treats vector fields on Banach manifolds, while Klingenberg's book [17] deals with the Hilbertian case. Notice that if $\dim M = \infty$, then not all derivations of $C^\infty(M)$ can be described as above by vector fields.

Back to the finite dimension

In the subsequent concluding part of this section we shall consider *vector bundles of finite rank* over a common base manifold M. Then, by our convention, M is also finite dimensional; let $n := \dim M$, $n \in \mathbb{N}^*$. We assume, furthermore, that *the topology of M is second countable*. Then, as it is well-known, *M admits a (smooth) partition of unity*; in particular, for each neighbourhood of every point of M there is a bump function supported in the given neighbourhood. Under these conditions we have the following useful technical result:

Lemma 2.24 *Let $\pi : E \to M$ be a vector bundle of rank k over the n-dimensional second countable base manifold M, and let p be a point of M. Then for each $z \in E_p$ there is a section $\sigma \in \Gamma(\pi)$ such that $\sigma(p) = z$. If $\mathcal{U} \subset M$ is an open set containing p, and $\sigma : \mathcal{U} \to E$ is a section of π over $\mathcal{U}$, then there is a section $\tilde{\sigma} \in \Gamma(\pi)$ which coincides with σ in a neighbourhood of p.*

Proof. Choose a local trivialization $\Phi : \mathcal{V} \times \mathbb{R}^k \to \pi^{-1}(\mathcal{V})$ of E with $p \in \mathcal{V}$ and a bump function $f \in C^\infty(M)$ supported in $\mathcal{V}$ such that $f(p) = 1$. Given $z \in E_p$, there is a unique vector $v \in \mathbb{R}^k$ for which $\Phi(p, v) = z$. Define a mapping $\sigma : M \to E$ by

$$\sigma(q) := \begin{cases} \Phi(q, f(q)v) & \text{if } q \in \mathcal{V}, \\ 0 & \text{if } q \notin \mathcal{V}. \end{cases}$$

Then σ is clearly a (smooth) section of π with the desired property $\sigma(p) = z$, so our first claim is true. The second statement can be verified similarly. □

By a *frame* of $\pi : E \to M$ over an open subset $\mathcal{U}$ of M we mean a sequence $(\sigma_i)_{i=1}^k$ of sections of E over $\mathcal{U}$ such that $(\sigma_i(p))_{i=1}^k$ is a basis of E_p for all $p \in \mathcal{U}$. If the domain $\mathcal{U}$ of the sections σ_i is not specified, we shall speak of a *local frame*. *Local frames are actually the same objects as local trivializations.* Indeed, if $(\sigma_1, \dots, \sigma_k)$ is a frame of E over $\mathcal{U}$, then the mapping

$$\Phi : \mathcal{U} \times \mathbb{R}^k \to E, \quad \left(p, \textstyle\sum_{i=1}^k \nu^i e_i\right) \mapsto \sum_{i=1}^k \nu^i \sigma_i(p)$$

is a trivialization of $\pi^{-1}(\mathcal{U})$. Conversely, if $\Phi : \mathcal{U} \times \mathbb{R}^k \to E$ is a local trivialization of E, then the mappings

$$\sigma_i : p \in \mathcal{U} \mapsto \sigma_i(p) := \Phi(p, e_i) \quad (1 \leqq i \leqq k)$$

form a local frame of E.

Now consider two vector bundles, $\pi_1 : E_1 \to M$ and $\pi_2 : E_2 \to M$ of finite rank. A mapping $\Gamma(E_1) \to \Gamma(E_2)$ will be called *tensorial* if it is $C^\infty(M)$-linear. The following simple result is of basic importance and frequently (in general, tacitly) used.

Lemma 2.25 (the fundamental lemma of strong bundle maps) *A mapping $\mathcal{F} : \Gamma(\pi_1) \to \Gamma(\pi_2)$ is tensorial if and only if there is a strong bundle map $F : E_1 \to E_2$ such that $\mathcal{F}(\sigma) = F \circ \sigma$ for all $\sigma \in \Gamma(\pi_1)$.*

For a proof see e.g. John M. Lee's text [21]. In the following we shall usually identify a tensorial mapping $\mathcal{F} : \Gamma(\pi_1) \to \Gamma(\pi_2)$ with the corresponding bundle map F, and write $F\sigma$ rather than $\mathcal{F}(\sigma)$ or $F \circ \sigma$.

The concept of a linear differential operator introduced in 2.14 can immediately be generalized to the context of vector bundles. Our initial definition is strongly motivated by the property formulated in the local Peetre theorem 2.23.

Definition 2.26 An $\mathbb{R}$-linear mapping $D : \Gamma(\pi_1) \to \Gamma(\pi_2)$, $\sigma \mapsto D(\sigma) =: D\sigma$ is said to be a *linear differential operator* if it is support-decreasing, i.e., $\operatorname{supp}(D\sigma) \subset \operatorname{supp}(\sigma)$ for any section $\sigma \in \Gamma(\pi_1)$.

Linear differential operators are natural with respect to restrictions: if $\mathcal{U}$ is an open subset of M, and $D : \Gamma(\pi_1) \to \Gamma(\pi_2)$ is a linear differential operator, then (using the abbreviation $(\pi_i)_{\mathcal{U}} := \pi_i \restriction \pi_i^{-1}(\mathcal{U})$, $i \in \{1,2\}$) there is a unique differential operator $D_{\mathcal{U}} : \Gamma((\pi_1)_{\mathcal{U}}) \to \Gamma((\pi_2)_{\mathcal{U}})$ such that $D\sigma \restriction \mathcal{U} = D_{\mathcal{U}}(\sigma \restriction \mathcal{U})$ for every section $\sigma \in \Gamma(\pi_1)$. Indeed, let $p \in \mathcal{U}$ be an arbitrary point. By Lemma 2.24, for any section $\sigma_{\mathcal{U}} \in \Gamma((\pi_1)_{\mathcal{U}})$ there is a section σ of π_1 such that $\sigma = \sigma_{\mathcal{U}}$ in a neighbourhood of p. If $(D_{\mathcal{U}}\sigma_{\mathcal{U}})(p) := (D\sigma)(p)$, then $D_{\mathcal{U}} : \Gamma((\pi_1)_{\mathcal{U}}) \to \Gamma((\pi_2)_{\mathcal{U}})$ is a well-defined linear differential operator with the desired naturality property. An equivalent formulation: D is a *local operator* in the sense that *for each open subset $\mathcal{U}$ of M and each section $\sigma \in \Gamma(\pi_1)$ such that $\sigma \restriction \mathcal{U} = 0$, we have $(D\sigma) \restriction \mathcal{U} = 0$.*

Let $(\mathcal{U}, x)$ be a chart on M. Suppose that π_1 and π_2 are trivializable over $\mathcal{U}$, i.e., there exist trivializations $\Phi_1 : \mathcal{U} \times \mathbb{R}^k \to \pi_1^{-1}(\mathcal{U})$ and $\Phi_2 : \mathcal{U} \times \mathbb{R}^\ell \to \pi_2^{-1}(\mathcal{U})$. If $\sigma \in \Gamma(\pi_1)$, then $\tilde{\sigma} := \mathrm{pr}_2 \circ \Phi_1^{-1} \circ \sigma \circ x^{-1}$ is a smooth mapping from $x(\mathcal{U}) \subset \mathbb{R}^n$ to $\mathbb{R}^k$, and the mapping $\sigma \in \Gamma(\pi_1) \mapsto \tilde{\sigma} \in C^\infty\left(x(\mathcal{U}), \mathbb{R}^k\right)$ is a linear isomorphism. Let $D : \Gamma(\pi_1) \to \Gamma(\pi_2)$ be a linear differential operator, and consider the induced operator $D_{\mathcal{U}} : \Gamma((\pi_1)_{\mathcal{U}}) \to \Gamma((\pi_2)_{\mathcal{U}})$. There exists a well-defined continuous linear mapping $\tilde{D}$ from the Fréchet space $C^\infty\left(x(\mathcal{U}), \mathbb{R}^k\right)$ into the Fréchet space $C^\infty\left(x(\mathcal{U}), \mathbb{R}^\ell\right)$ such that

$$\mathrm{pr}_2 \circ \Phi_2^{-1} \circ D_{\mathcal{U}}(\sigma \restriction \mathcal{U}) \circ x^{-1} = \tilde{D}\left(\mathrm{pr}_2 \circ \Phi_1^{-1} \circ (\sigma \restriction \mathcal{U}) \circ x^{-1}\right);$$

briefly $\widetilde{D_{\mathcal{U}}\sigma} = \tilde{D}(\tilde{\sigma})$. $\tilde{D}$ is said to be the *local expression* of D with respect to the chart $(\mathcal{U}, x)$ and the trivializations Φ_1, Φ_2.

Now we are in a position to transpose Peetre's local theorem (2.23) to the context of vector bundles.

Theorem 2.27 (Peetre's theorem) *Let π_1 and π_2 be vector bundles over M of rank k and ℓ, respectively. If $D : \Gamma(\pi_1) \to \Gamma(\pi_2)$ is a linear differential operator, then there exists a chart $(\mathcal{U}, x)$ at each point of M, such that π_1 and π_2 are trivializable over $\mathcal{U}$, and the corresponding local expression of D is of the form*

$$f \in C^\infty\left(x(\mathcal{U}), \mathbb{R}^k\right) \mapsto \sum_{|\alpha| \leqq m} A_\alpha \circ D^\alpha f,$$

where for each multi-index α such that $|\alpha| \leqq m$, A_α is a smooth mapping from $x(\mathcal{U}) \subset \mathbb{R}^n$ to $\mathcal{L}\left(\mathbb{R}^k, \mathbb{R}^\ell\right)$. □

Peetre's theorem makes it possible to define the *order* of a linear differential operator $D : \Gamma(\pi_1) \to \Gamma(\pi_2)$ *at a point* $p \in M$ as the largest $m \in \mathbb{N}$ for which there exists a multi-index α such that $|\alpha| = m$, and $A_\alpha(p) \neq 0$ in a local expression of D in a neighbourhood of p. It follows at once that if D is *of order* 0, then it may be identified with a strong bundle map $E_1 \to E_2$, therefore it acts by the rule $\sigma \in \Gamma(\pi_1) \mapsto D \circ \sigma \in \Gamma(\pi_2)$. Equivalently, in view of Lemma 2.25, D can be considered as a tensorial mapping from $\Gamma(\pi_1)$ to $\Gamma(\pi_2)$.

For the rest of this section, we let $\pi : E \to M$ be a vector bundle of rank k over an n-dimensional base manifold M. We continue to assume that M admits a partition of unity.

An important class of first order differential operators from $\Gamma(\pi)$ to $\Gamma(\pi)$ may be specified by introducing the concept of covariant differentiation. The notion is well-known, but fundamental, so we briefly recall that a mapping

$$D : \mathfrak{X}(M) \times \Gamma(\pi) \to \Gamma(\pi), \quad (X, \sigma) \mapsto D_X\sigma$$

is said to be a *covariant derivative* on E if it is tensorial in X and derivation in σ. $D_X\sigma$ is called the covariant derivative of σ in the direction of X, while for any section $\sigma \in \Gamma(\pi)$, the mapping

$$D\sigma : \mathfrak{X}(M) \to \Gamma(\pi), \quad \sigma \mapsto D_X\sigma$$

is the *covariant differential* of σ. Then $D\sigma \in \mathcal{L}_{C^\infty(M)}(\mathfrak{X}(M), \Gamma(\pi))$. For every fixed vector field X on M, the mapping

$$D_X : \Gamma(\pi) \to \Gamma(\pi), \quad \sigma \mapsto D_X\sigma$$

is obviously $\mathbb{R}$-linear. Moreover, *D_X is a local operator*. Indeed, let $\mathcal{U} \subset M$ be an open set, and choose a point $p \in M$. The existence of bump functions supported in $\mathcal{U}$ also guarantees that there is a smooth function f on M such that $f(p) = 0$ and $f = 1$ outside $\mathcal{U}$. If $\sigma \in \Gamma(\pi)$, and $\sigma \restriction \mathcal{U} = 0$, then $\sigma = f\sigma$, and

$$(D_X\sigma)(p) = (D_X(f\sigma))(p) = (Xf)(p)\sigma(p) + f(p)(D_X\sigma)(p) = 0,$$

therefore $D_X\sigma \restriction \mathcal{U} = 0$. $\mathbb{R}$-linearity and locality imply that D_X is a linear differential operator. To see that it is of first order, let $\left(\mathcal{U}, \left(u^i\right)_{i=1}^n\right)$ be a chart on M, and $(\sigma_j)_{j=1}^k$ a frame over $\mathcal{U}$. If $X \in \mathfrak{X}(M)$, $\sigma \in \Gamma(\pi)$, then $X \restriction \mathcal{U} = \sum_{i=1}^n X^i \frac{\partial}{\partial u^i}$, $\sigma \restriction \mathcal{U} = \sum_{r=1}^k f^r\sigma_r \left(X^i, f^r \in C^\infty(\mathcal{U});\ 1 \leqq i \leqq n, 1 \leqq r \leqq k\right)$, and by the local character of D_X,

$$(D_X\sigma) \restriction \mathcal{U} = D_{X \restriction \mathcal{U}}(\sigma \restriction \mathcal{U}) = \sum_{i=1}^n \sum_{r=1}^k X^i \left(\frac{\partial f^r}{\partial u^i}\sigma_r + f^r D_{\frac{\partial}{\partial u^i}}\sigma_r \right).$$

The local sections $D_{\frac{\partial}{\partial u^i}}\sigma_r$ can be combined from the frame $(\sigma_j)_{j=1}^k$ in the form

$$D_{\frac{\partial}{\partial u^i}}\sigma_r = \sum_{s=1}^k \Gamma_{ir}^s \sigma_s, \quad \Gamma_{ir}^s \in C^\infty(\mathcal{U});$$

so we get

$$(D_X\sigma) \upharpoonright \mathcal{U} = \sum_{i=1}^{n}\sum_{s=1}^{k} X^i \left(\frac{\partial f^s}{\partial u^i} + \textstyle\sum_{r=1}^{k} \Gamma_{ir}^s f^r \right) \sigma_s.$$

Since $\frac{\partial f^s}{\partial u^i} := D_i \left(f^s \circ u^{-1} \right) \circ u$, D_X is indeed of first order. Notice that the functions $\Gamma_{ir}^s \in C^\infty(\mathcal{U})$ ($1 \leqq i \leqq n$; $1 \leqq r, s \leqq k$) are called the *Christoffel symbols* of D with respect to the chart $\left(\mathcal{U}, \left(u^i\right)_{i=1}^n\right)$ and local frame $(\sigma_i)_{i=1}^k$.

3 The Chern – Rund derivative

Conventions

Throughout this section we shall work over an n-dimensional ($n \in \mathbb{N}^*$) smooth manifold M admitting a partition of unity. More precisely, the main scene of our next considerations will be the tangent bundle $\tau : TM \to M$ of M and the pull-back bundle $\tau_1 : \tau^*TM = TM \times_M TM \to TM$. We shall also need the tangent bundle $\tau_{TM} : TTM \to TM$ of TM, the deleted bundle $\overset{\circ}{\tau} : \overset{\circ}{T}M \to M$ $\left(\overset{\circ}{T}M := \{v \in TM | v \neq 0\}, \overset{\circ}{\tau} := \tau \upharpoonright \overset{\circ}{T}M\right)$ for τ and the pull-back $\overset{\circ}{\tau}{}^*TM = \overset{\circ}{T}M \times_M TM$ of TM via $\overset{\circ}{\tau}$. For the $C^\infty(TM)$-modules $\Gamma(\tau^*TM) \cong \Gamma_\tau(TM)$ and $\Gamma\left(\overset{\circ}{\tau}{}^*TM\right) \cong \Gamma_{\overset{\circ}{\tau}}(TM)$ we use the convenient notations $\mathfrak{X}(\tau)$ and $\mathfrak{X}\left(\overset{\circ}{\tau}\right)$, resp. Any vector field X on M induces a vector field $\hat{X}$ along τ and $\overset{\circ}{\tau}$ with principal parts $X \circ \tau$ and $X \circ \overset{\circ}{\tau}$, resp. $\hat{X}$ is called a *basic vector field* along τ (or $\overset{\circ}{\tau}$). Locally, the basic vector fields generate the modules $\mathfrak{X}(\tau)$ and $\mathfrak{X}\left(\overset{\circ}{\tau}\right)$. A distinguished vector field along τ is the *canonical vector field*

$$\delta : v \in TM \mapsto \delta(v) := (v, v) \in TM \times_M TM.$$

Generic vector fields along τ (or $\overset{\circ}{\tau}$) will be denoted by $\tilde{X}, \tilde{Y}, \dots$. From the module $\mathfrak{X}(\tau)$ (or $\mathfrak{X}\left(\overset{\circ}{\tau}\right)$) one can build the spaces of type (r, s) tensors along τ (or $\overset{\circ}{\tau}$). These $C^\infty(TM)$-multilinear machines can also be interpreted as 'fields'. For example, a type $(0, 2)$ tensor field $g : \mathfrak{X}\left(\overset{\circ}{\tau}\right) \times \mathfrak{X}\left(\overset{\circ}{\tau}\right) \to C^\infty\left(\overset{\circ}{T}M\right)$ can be regarded as a mapping

$$v \in \overset{\circ}{T}M \mapsto g_v \in \mathcal{L}^2\left(T_{\overset{\circ}{\tau}(v)}M, \mathbb{R}\right)$$

which has the following smoothness property: the function

$$g\left(\tilde{X}, \tilde{Y}\right) : v \in \overset{\circ}{T}M \mapsto g\left(\tilde{X}, \tilde{Y}\right)(v) := g_v\left(\tilde{X}(v), \tilde{Y}(v)\right)$$

is smooth for any two vector fields $\tilde{X}, \tilde{Y}$ along $\overset{\circ}{\tau}$.

For coordinate calculations we choose a chart $\left(\mathcal{U}, \left(u^i\right)_{i=1}^n\right)$ on M, and employ the induced chart

$$\left(\tau^{-1}(\mathcal{U}), \left(x^i, y^i\right)\right); \quad x^i := u^i \circ \tau, \quad y^i : v \in \tau^{-1}(\mathcal{U}) \mapsto y^i(v) := v\left(u^i\right)$$

($1 \leqq i \leqq n$) on TM. The coordinate vector fields

$$\frac{\partial}{\partial u^i} : f \in C^\infty(\mathcal{U}) \mapsto \frac{\partial f}{\partial u^i} := D_i\left(f \circ u^{-1}\right) \circ u \in C^\infty(\mathcal{U}) \quad (1 \leqq i \leqq n)$$

form a frame of $\tau : TM \to M$ over $\mathcal{U}$. Similarly, $\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial y^i}\right)_{i=1}^n$ is a local frame of $\tau_{TM} : TTM \to TM$. The basic vector fields

$$\widehat{\frac{\partial}{\partial u^i}} : v \in \tau^{-1}(\mathcal{U}) \mapsto \left(v, \left(\frac{\partial}{\partial u^i}\right)_{\tau(v)}\right) \quad (1 \leqq i \leqq n)$$

provide a local frame for $\tau^* TM$. Using this frame, over $\tau^{-1}(\mathcal{U})$ we have

$$\hat{X} = \sum_{i=1}^n \left(X^i \circ \tau\right) \widehat{\frac{\partial}{\partial u^i}} \quad \text{if } X \in \mathfrak{X}(M),\ X \restriction \mathcal{U} = \sum_{i=1}^n X^i \frac{\partial}{\partial u^i};$$
$$\delta = \sum_{i=1}^n y^i \widehat{\frac{\partial}{\partial u^i}}.$$

The coordinate expression of a generic section $\tilde{X}$ of $\tau^* TM$ is

$$\tilde{X} \restriction \tau^{-1}(\mathcal{U}) = \sum_{i=1}^n \tilde{X}^i \widehat{\frac{\partial}{\partial u^i}}, \quad \tilde{X}^i \in C^\infty\left(\tau^{-1}(\mathcal{U})\right) \quad (1 \leqq i \leqq n).$$

Canonical constructions on TM and $\tau^* TM$

We begin with a frequently used, simple observation. Let V be an n-dimensional real vector space, endowed with the canonical smooth structure determined by a linear isomorphism of V onto $\mathbb{R}^n$. *For any $p \in V$, V may be naturally identified with its tangent space* $T_p V$ via the linear isomorphism

$$\imath_p : v \in V \mapsto \imath_p(v) := \dot{\varrho}(0) \in T_p V, \quad \varrho(t) := p + tv \quad (t \in \mathbb{R})$$

($\dot{\varrho}(0)$ is the tangent vector of ϱ at 0 in the sense of classical manifold theory). If $(e_i)_{i=1}^n$ is a basis of V, and $\left(e^i\right)_{i=1}^n$ is its dual, then

$$\imath_p(v) = \sum_{i=1}^n e^i(v) \left(\frac{\partial}{\partial e^i}\right)_p.$$

By means of these identifications, we get an injective strong bundle map

$$\mathbf{i} : TM \times_M TM \to TTM,\ (v, w) \in \{v\} \times T_{\tau(v)} M \mapsto \mathbf{i}(v, w) := \imath_v(w) \in T_v T_{\tau(v)} M.$$

$\mathrm{Im}(\mathbf{i}) =: VTM$ is said to be the *vertical bundle* of $\tau_{TM} : TTM \to TM$. It is easy to check that $VTM = \mathrm{Ker}(\tau_*)$. By Lemma 2.25, $\mathbf{i}$ may be interpreted as a tensorial mapping from $\mathfrak{X}(\tau)$ to $\mathfrak{X}(TM)$ denoted by the same symbol. $\mathfrak{X}^v(TM) := \mathbf{i}(\mathfrak{X}(\tau))$ is the module of *vertical vector fields* on TM (in fact, $\mathfrak{X}^v(TM)$ is a Lie-subalgebra of $\mathfrak{X}(TM)$). In particular, $X^v := \mathbf{i}\hat{X}$ is said to be the *vertical lift* of $X \in \mathfrak{X}(M)$; $C := \mathbf{i}\delta$ is the *Liouville vector field* on TM. It is easy to check that

$$[X^v, Y^v] = 0, \quad [C, X^v] = -X^v; \quad X, Y \in \mathfrak{X}(M).$$

In terms of local coordinates,

$$\mathbf{i}(v,w) = \sum_{i=1}^{n} y^i(w)\left(\frac{\partial}{\partial y^i}\right)_v \quad (\tau(v) = \tau(w));$$

$$\mathbf{i}\left(\widehat{\frac{\partial}{\partial u^i}}\right)_v = \mathbf{i}\left(v, \left(\frac{\partial}{\partial u^i}\right)_{\tau(v)}\right) = \left(\frac{\partial}{\partial y^i}\right)_v,$$

$$\text{hence } \left(\frac{\partial}{\partial u^i}\right)^v = \frac{\partial}{\partial y^i} \quad (1 \leqq i \leqq n);$$

$$X^v = \sum_{i=1}^{n}\left(X^i \circ \tau\right)\frac{\partial}{\partial y^i} \quad \text{if } X \restriction \mathcal{U} = \sum_{i=1}^{n} X^i \frac{\partial}{\partial u^i}.$$

A further and surjective strong bundle map is

$$\mathbf{j} := (\tau_{TM}, \tau_*) : TTM \to TM \times_M TM, \quad z \in T_vTM \mapsto \mathbf{j}(z) := (v, \tau_*(z)),$$

which can also be regarded as a tensorial mapping from $\mathfrak{X}(TM)$ to $\mathfrak{X}(\tau)$. $\mathbf{j}$ acts on the local frame $\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial y^i}\right)_{i=1}^{n}$ by

$$\mathbf{j}\left(\frac{\partial}{\partial x^i}\right) = \widehat{\frac{\partial}{\partial u^i}}, \quad \mathbf{j}\left(\frac{\partial}{\partial y^i}\right) = 0 \quad (1 \leqq i \leqq n),$$

therefore $\mathrm{Ker}(\mathbf{j}) = \mathrm{Im}(\mathbf{i}) = \mathfrak{X}^v(TM)$. The composition $J := \mathbf{i} \circ \mathbf{j}$ is said to be the *vertical endomorphism* of $\mathfrak{X}(TM)$ (or TTM). It follows that

$$\mathrm{Im}(J) = \mathrm{Ker}(J) = \mathfrak{X}^v(TM), \quad J^2 = 0.$$

By the *complete lift* of a smooth function f on M we mean the function $f^c : v \in TM \mapsto f^c(v) := v(f) \in \mathbb{R}$. Then, obviously, $f^c \in C^\infty(TM)$. It can be shown that *for any vector field X on M there exists a unique vector field X^c on TM such that $X^c f^c = (Xf)^c$ for all $f \in C^\infty(M)$* [35]. X^c is said to be the *complete lift* of X. A great deal of calculations may be simplified by the fact that *if $(X_i)_{i=1}^n$ is a local frame of TM, then $(X_i^v, X_i^c)_{i=1}^n$ is a local frame of TTM.*

It follows immediately that $\mathbf{j}X^c = \hat{X}$, or, equivalently, $JX^c = X^v$. Concerning the vertical and the complete lifts, we have

$$[X^c, Y^c] = [X,Y]^c, \quad [X^v, Y^c] = [X,Y]^v, \quad [C, X^c] = 0; \quad X, Y \in \mathfrak{X}(M).$$

We shall need, furthermore, the following

Lemma 3.1 *If X and Z are vector fields on M, and F is a smooth function on TM, then*

$$X(F \circ Z) = (X^c F + [X,Z]^v F) \circ Z.$$

The simplest, but not too aesthetic way to prove this relation is to express everything in terms of local coordinates.

On τ^*TM there exists a canonical differential operator of first order which makes it possible to differentiate tensors along τ in vertical directions. We call this operator the *canonical v-covariant derivative*, and we denote it by ∇^v. It can explicitly be given as follows: for each section $\tilde{X}$ in $\mathfrak{X}(\tau)$,

$$\begin{aligned} \nabla^v_{\tilde{X}} F &:= \left(\mathbf{i}\tilde{X}\right) F && \text{if } F \in C^\infty(TM); \\ \nabla^v_{\tilde{X}} \tilde{Y} &:= \mathbf{j}\left[\mathbf{i}\tilde{X}, \eta\right] && \text{if } \tilde{Y} \in \mathfrak{X}(\tau) \text{ and } \eta \in \mathfrak{X}(TM) \text{ such that } \mathbf{j}\eta = \tilde{Y}. \end{aligned}$$

Using the frame $\left(\widehat{\frac{\partial}{\partial u^i}}\right)_{i=1}^n$ over $\tau^{-1}(\mathcal{U})$, if $\tilde{X} \restriction \tau^{-1}(\mathcal{U}) = \sum_{i=1}^n \tilde{X}^i \widehat{\frac{\partial}{\partial u^i}}$ and $\tilde{Y} \restriction \tau^{-1}(\mathcal{U}) = \sum_{i=1}^n \tilde{Y}^i \widehat{\frac{\partial}{\partial u^i}}$, then

$$\nabla^v_{\tilde{X}} F \restriction \tau^{-1}(\mathcal{U}) = \sum_{i=1}^n \tilde{X}^i \frac{\partial F}{\partial y^i}, \quad \nabla^v_{\tilde{X}} \tilde{Y} \restriction \tau^{-1}(\mathcal{U}) = \sum_{i,j=1}^n \tilde{X}^i \frac{\partial \tilde{Y}^j}{\partial y^i} \widehat{\frac{\partial}{\partial u^j}}.$$

From the last expression it is clear that $\nabla^v_{\tilde{X}} \tilde{Y}$ is well-defined: it does not depend on the choice of η. The mapping $\nabla^v : \left(\tilde{X}, \tilde{Y}\right) \in \mathfrak{X}(\tau) \times \mathfrak{X}(\tau) \to \nabla^v_{\tilde{X}} \tilde{Y} \in \mathfrak{X}(\tau)$ has the formal properties of a covariant derivative operator: it is tensorial in $\tilde{X}$ and satifies the derivation rule

$$\nabla^v_{\tilde{X}} F\tilde{Y} = \left(\nabla^v_{\tilde{X}} F\right) \tilde{Y} + F \nabla^v_{\tilde{X}} \tilde{Y}, \quad F \in C^\infty(TM).$$

From the very definition, or using the coordinate expression, it can easily be deduced that

$$\nabla^v_{\tilde{X}} \hat{Y} = 0, \quad \nabla^v_{\tilde{X}} \delta = \tilde{X}; \quad \tilde{X} \in \mathfrak{X}(\tau),\ Y \in \mathfrak{X}(M).$$

Using an appropriate version of Willmore's theorem on tensor derivations (see e.g. [35, 1.32]), ∇^v can uniquely be extended to a tensor derivation of the tensor algebra $\mathfrak{X}(\tau)$. For any tensor $\tilde{A}$ along τ we may also consider the (canonical) v-covariant derivatives of $\tilde{A}$ by the rule $i_{\tilde{X}} \nabla^v \tilde{A} := \nabla^v_{\tilde{X}} \tilde{A}$ ($i_{\tilde{X}}$ denotes the substitution operator associated to $\tilde{X}$). If, in particular, $F \in C^\infty(TM)$, then $\nabla^v F$ is a one-form, $\nabla^v \nabla^v F := \nabla^v(\nabla^v F)$ is a type $(0,2)$ tensor along τ. $\nabla^v \nabla^v F$ is called the (vertical) *Hessian* of F. For any two vector fields X, Y on M we have

$$\begin{aligned} \nabla^v \nabla^v F\left(\hat{X}, \hat{Y}\right) &= X^v(Y^v F) = [X^v, Y^v]F + Y^v(X^v F) \\ &= Y^v(X^v F) = \nabla^v \nabla^v F\left(\hat{Y}, \hat{X}\right), \end{aligned}$$

so the tensor $\nabla^v \nabla^v F$ is *symmetric*.

Ehresmann connections and Berwald derivatives

Definition 3.2 By an *Ehresmann connection* on TM we mean a mapping $\mathcal{H} : TM \times_M TM \to TTM$ satisfying the following conditions:

(C$_1$) $\mathcal{H} \upharpoonright \overset{\circ}{T}M \times_M TM$ is smooth;

(C$_2$) for all $v \in \overset{\circ}{T}M$, $\mathcal{H} \upharpoonright \{v\} \times T_{\tau(v)}M$ is a linear mapping to T_vTM;

(C$_3$) $\mathbf{j} \circ \mathcal{H} = 1_{TM \times_M TM}$;

(C$_4$) if $o : M \to TM$ is the zero section of TM, then $\mathcal{H}(o(p), v) = (o_*)_p(v)$ for all $p \in M$, $v \in T_pM$.

Then (C$_3$) and (C$_4$) are clearly consistent, however, *the smoothness of $\mathcal{H}$ (and the objects derived from $\mathcal{H}$) is not guaranteed on its whole domain*. (This weakening of smoothness allows more flexibility in applications.) If $HTM := \mathrm{Im}(\mathcal{H})$, then $TTM = HTM \oplus VTM$ (Whitney sum); HTM is said to be a *horizontal subbundle* of TTM.

Given an Ehresmann connection $\mathcal{H}$ on TM, there exists a unique strong bundle map $\mathcal{V} : TTM \to TM \times_M TM$, smooth in general only on $T\overset{\circ}{T}M$, such that

$$mathcalV \circ \mathbf{i} = 1_{TM \times_M TM} \quad \text{and} \quad \mathrm{Ker}(\mathcal{V}) = \mathrm{Im}(\mathcal{H}).$$

$\mathcal{V}$ is called the *vertical map* associated to $\mathcal{H}$. $\mathbf{h} := \mathcal{H} \circ \mathbf{j}$ and $\mathbf{v} := \mathbf{i} \circ \mathcal{V}$ are projectors on TTM, the *horizontal* and *vertical projector* belonging to $\mathcal{H}$, respectively. $\mathcal{H}$ (as well as $\mathcal{V}$, $\mathbf{h}$ and $\mathbf{v}$) induce tensorial mappings at the level of sections. $\mathfrak{X}^h(TM) := \mathcal{H}(\mathfrak{X}(\tau)) = \mathbf{h}(\mathfrak{X}(TM))$ is called the space of *horizontal vector fields* on TM. In particular, $X^h := \mathcal{H}(\hat{X}) = \mathbf{h}X^c$ is the *horizontal lift* of $X \in \mathfrak{X}(M)$. We have: $\mathfrak{X}(TM) = \mathfrak{X}^h(TM) \oplus \mathfrak{X}^v(TM)$ (direct sum of modules). In terms of the local frames $\left(\widehat{\frac{\partial}{\partial u^i}}\right)_{i=1}^n$ of τ^*TM and $\left(\left(\frac{\partial}{\partial x^i}\right)_{i=1}^n, \left(\frac{\partial}{\partial y^i}\right)_{i=1}^n\right)$ of TTM, we get the following coordinate expressions:

$$\mathcal{H}\left(\widehat{\frac{\partial}{\partial u^i}}\right) = \frac{\partial}{\partial x^i} - \sum_{j=1}^n N_i^j \frac{\partial}{\partial y^j}; \quad \mathcal{V}\left(\frac{\partial}{\partial x^i}\right) = \sum_{j=1}^n N_i^j \widehat{\frac{\partial}{\partial u^j}}, \quad \mathcal{V}\left(\frac{\partial}{\partial y^i}\right) = \widehat{\frac{\partial}{\partial u^i}};$$

$$\mathbf{h}\left(\frac{\partial}{\partial x^i}\right) = \frac{\partial}{\partial x^i} - \sum_{j=1}^n N_i^j \frac{\partial}{\partial y^j}, \quad \mathbf{h}\left(\frac{\partial}{\partial y^i}\right) = 0;$$

$$\mathbf{v}\left(\frac{\partial}{\partial x^i}\right) = \sum_{j=1}^n N_i^j \frac{\partial}{\partial y^j}, \quad \mathbf{v}\left(\frac{\partial}{\partial y^i}\right) = \frac{\partial}{\partial y^i} \quad (1 \leqq i \leqq n).$$

The functions N_i^j, defined on $\tau^{-1}(\mathcal{U})$ and smooth on $\overset{\circ}{\tau}^{-1}(\mathcal{U})$, are called the *Christoffel symbols* of $\mathcal{H}$ with respect to the given local frames. (The minus sign in the first formula is more or less traditional.)

Via linearization, any Ehresmann connection $\mathcal{H}$ leads to a covariant derivative $\nabla : \mathfrak{X}\left(\overset{\circ}{T}M\right) \times \mathfrak{X}(\overset{\circ}{\tau}) \to \mathfrak{X}(\overset{\circ}{\tau})$ on τ^*TM, called the *Berwald derivative* induced by $\mathcal{H}$. The explicit rules of calculation are

$$\nabla_{\mathcal{H}\tilde{X}}\tilde{Y} := \mathcal{V}\left[\mathcal{H}\tilde{X}, \mathbf{i}\tilde{Y}\right] \quad \text{and} \quad \nabla_{\mathbf{i}\tilde{X}}\tilde{Y} := \nabla^v_{\tilde{X}}\tilde{Y}; \quad \tilde{X}, \tilde{Y} \in \mathfrak{X}(\overset{\circ}{\tau}).$$

The *h-part* of ∇, given by $\nabla^h_{\tilde{X}}\tilde{Y} := \nabla_{\mathcal{H}\tilde{X}}\tilde{Y}$, has the coordinate expression

$$\nabla^h_{\widehat{\frac{\partial}{\partial u^i}}} \widehat{\frac{\partial}{\partial u^j}} = \nabla_{\left(\frac{\partial}{\partial u^i}\right)^h} \widehat{\frac{\partial}{\partial u^j}} = \sum_{k=1}^n \frac{\partial N_j^k}{\partial y^i} \widehat{\frac{\partial}{\partial u^k}} \quad (1 \leqq i, j \leqq n),$$

where the functions N_j^k are the Christoffel symbols of $\mathcal{H}$. If $\tilde{A}$ is any tensor along $\overset{\circ}{\tau}$, we may also consider the *h-Berwald differential* $\nabla^h \tilde{A}$ of $\tilde{A}$ given by $\nabla^h \tilde{A}(\tilde{X}) := \nabla^h_{\tilde{X}} \tilde{A}$. $\mathbf{t} := \nabla^h \delta$ is said to be the *tension* of $\mathcal{H}$. Then for any section $\tilde{X}$ along $\overset{\circ}{\tau}$ we have

$$\mathbf{t}(\tilde{X}) = \nabla^h \delta(\tilde{X}) = \nabla_{\mathcal{H}\tilde{X}} \delta = \mathcal{V}\left[\mathcal{H}\tilde{X}, C\right].$$

$\mathcal{H}$ is called *homogeneous* if $\mathbf{t} = 0$. If a homogeneous Ehresmann connection is of class C^1 at the zeros, then there exists a covariant derivative D on M (more precisely, on TM) such that for all $X, Y \in \mathfrak{X}(M)$,

$$(D_X Y)^v = \left[X^h, Y^v\right].$$

So *under homogeneity and C^1-differentiability Ehresmann connections lead to classical covariant derivatives on the base manifold*.

For covariant derivatives given on a generic vector bundle there is no reasonable concept of 'torsion'. However, if D is a covariant derivative on the pull-back bundle $\tau^* TM$, and an Ehresmann connection $\mathcal{H}$ is also specified on TM, we have useful generalizations of the classical torsion: the *vertical torsion* $T^v(D)$ and the *horizontal torsion* $T^h(D)$ given by

$$T^v(D)\left(\tilde{X}, \tilde{Y}\right) := D_{\mathbf{i}\tilde{X}} \tilde{Y} - D_{\mathbf{i}\tilde{Y}} \tilde{X} - \mathbf{i}^{-1}\left[\mathbf{i}\tilde{X}, \mathbf{i}\tilde{Y}\right]$$

and

$$T^h(D)\left(\tilde{X}, \tilde{Y}\right) := D_{\mathcal{H}\tilde{X}} \tilde{Y} - D_{\mathcal{H}\tilde{Y}} \tilde{X} - \mathbf{j}\left[\mathcal{H}\tilde{X}, \mathcal{H}\tilde{Y}\right],$$

respectively. (It is easy to check that both $T^v(D)$ and $T^h(D)$ are tensorial.)

The vertical torsion of any Berwald derivative ∇ vanishes. Indeed, for all $\xi, \eta \in \mathfrak{X}(TM)$ we have

$$\begin{aligned}\mathbf{i}(T^v(\nabla)(\mathbf{j}\xi, \mathbf{j}\eta)) &= \mathbf{i}\nabla_{J\xi}\mathbf{j}\eta - \mathbf{i}\nabla_{J\eta}\mathbf{j}\xi - [J\xi, J\eta] \\ &= J[J\xi, \eta] - J[J\eta, \xi] - [J\xi, J\eta] = -N_J(\xi, \eta),\end{aligned}$$

where N_J is the *Nijenhuis torsion* of J. However, as it is well-known, $N_J = 0$, therefore $T^v(\nabla) = 0$.

The horizontal torsion of the Berwald derivative induced by an Ehresmann connection $\mathcal{H}$ is said to be the *torsion of* $\mathcal{H}$. Denoting this tensor by $\mathbf{T}$, we get

$$\mathbf{i}\mathbf{T}\left(\hat{X}, \hat{Y}\right) = \left[X^h, Y^v\right] - \left[Y^h, X^v\right] - [X, Y]^v; \quad X, Y \in \mathfrak{X}(M).$$

If $\mathcal{H}$ is of class C^1 and homogeneous, then there is a covariant derivative D on M such that $(D_X Y)^v = \left[X^h, Y^v\right]$ $(X, Y \in \mathfrak{X}(M))$; hence

$$\mathbf{i}\mathbf{T}\left(\hat{X}, \hat{Y}\right) = (D_X Y - D_Y X - [X, Y])^v =: (T(D))^v,$$

so $\mathbf{T}$ reduces to the usual torsion of D.

Definition 3.3 By a *semispray* on TM we mean a mapping $S : TM \to TTM$ such that $\tau_{TM} \circ S = 1_{TM}$, $JS = C$ (or, equivalently, $\mathbf{j}S = \delta$) and $S \restriction \overset{\circ}{T}M$ is smooth. A semispray is said to be a *spray* if it is of class C^1 on TM, and has the homogeneity property $[C, S] = S$. A spray is called *affine* if it is of class C^2 (and hence smooth) on TM.

Any semispray S on TM induces an Ehresmann connection $\mathcal{H}_S$ on TM such that for all $X \in \mathfrak{X}(M)$ we have

$$\mathcal{H}_S\big(\hat{X}\big) = \frac{1}{2}(X^c + [X^v, S]).$$

The torsion of $\mathcal{H}_S$ vanishes. Conversely, *if an Ehresmann connection $\mathcal{H}$ has vanishing torsion, then there exists a semispray S on TM such that $\mathcal{H}_S = \mathcal{H}$*, i.e., '$\mathcal{H}$ is generated by a semispray'. These important results (at least in an intrinsic formulation) are due to M. Crampin and J. Grifone (independently). For details we refer to [35].

Parametric Lagrangians and Finsler manifolds

First we recall that a function $f : TM \to \mathbb{R}$ is called *positive-homogeneous of degree* r ($r \in \mathbb{R}$), briefly r^+*-homogeneous* if for each $\lambda \in \mathbb{R}^*_+$ and $v \in TM$ we have $f(\lambda v) = \lambda^r f(v)$. If f is smooth on $\overset{\circ}{T}M$, then $Cf = rf$, or, equivalently, $\nabla^v_\delta f = rf$ (Euler's relation). Conversely, this property implies that f is r^+-homogeneous on $\overset{\circ}{T}M$.

Definition 3.4 By a *parametric Lagrangian* we mean a 1^+-homogeneous function $F : TM \to \mathbb{R}$ which is smooth on $\overset{\circ}{T}M$. Then $Q := \frac{1}{2}F^2$ is called the *quadratic Lagrangian* or *energy function* associated to F. The symmetric type (0,2) tensor

$$g_F := \nabla^v\nabla^v Q = \frac{1}{2}\nabla^v\nabla^v F^2$$

along $\overset{\circ}{\tau}$ is the *metric tensor* determined by F. If, in addition,

$$F(v) > 0 \quad \text{whenever } v \in \overset{\circ}{T}M,$$

then F is called *positive definite*.

Lemma 3.5 *Let $F : TM \to \mathbb{R}$ be a parametric Lagrangian. Then*

(i) $\nabla^v\nabla^v F\left(\delta, \hat{X}\right) = 0$ *for every vector field X on M, therefore the vertical Hessian of F is degenerate.*

(ii) The quadratic Lagrangian $Q = \frac{1}{2}F^2$ is 2^+-homogeneous; C^1 on TM, smooth on $\overset{\circ}{T}M$.

(iii) The metric tensor $g_F = \nabla^v\nabla^v Q$ is 0^+-homogeneous in the sense that $\nabla^v_\delta g_F = 0$.

(iv) g and Q are related by $g_F(\delta, \delta) = 2Q$.

(v) If $\tilde{\vartheta}_F(\tilde{X}) := g\left(\tilde{X}, \delta\right)$*, then* $\tilde{\vartheta}_F$ *is a one-form along* $\overset{\circ}{\tau}$*, and* $\vartheta_F := \tilde{\vartheta}_F \circ \mathbf{j}$ *is a one-form on* $\overset{\circ}{T}M$*. We have* $\tilde{\vartheta}_F = \nabla^v Q = F\nabla^v F$.

Except the second statement of (ii), each claim can be verified by immediate calculations. For example, taking into account that $\nabla^v_{\tilde{X}}\hat{Y} = 0$ for all $\tilde{X} \in \mathfrak{X}(\tau)$, $Y \in \mathfrak{X}(M)$,

$$\begin{aligned}\nabla^v\nabla^v F\left(\delta, \hat{X}\right) &= \nabla^v_\delta(\nabla^v F)(\hat{X}) = C\left(\nabla^v_{\hat{X}}F\right) - \nabla^v F\left(\nabla^v_\delta \hat{X}\right) = C(X^v F)\\ &= [C, X^v]F + X^v(CF) = -X^v F + X^v F = 0,\end{aligned}$$

whence (i). As for the (not difficult) proof of the fact that Q is C^1 on TM, see [35, p. 1378, Observation].

Remark 3.6 Both $\tilde{\vartheta}_F$ and ϑ_F are called the *Lagrange one-form* associated to F; $\omega_F := d\vartheta_F$ is the *Lagrange two-form* (d is the classical exterior derivative on TM). ω_F and the metric tensor g_F are related by

$$\omega_F(J\xi, \eta) = g_F(\mathbf{j}\xi, \mathbf{j}\eta); \quad \xi, \eta \in \mathfrak{X}\left(\overset{\circ}{T}M\right)$$

(to check this it is enough to evaluate both sides on a pair (X^c, Y^c) with $X, Y \in \mathfrak{X}(M)$). We conclude that *the Lagrange two-form and the metric tensor associated to a parametric Lagrangian are non-degenerate at the same time*. (Non-degeneracy is meant pointwise. This implies a corresponding property at the level of vector fields, but not vice versa.)

Definition 3.7 By a *Finsler function* we mean a positive definite parametric Lagrangian whose associated metric tensor is non-degenerate (and hence is a pseudo-Riemannian metric on $\overset{\circ}{\tau}{}^*TM$). A manifold is said to be a *Finsler manifold* if its tangent bundle is endowed with a Finsler function.

Proposition 3.8 *If* (M, F) *is a Finsler manifold, then the metric tensor* g_F *is positive definite, i.e.,* g_F *is a Riemannian metric in* $\overset{\circ}{\tau}{}^*TM$.

Proof. The problem can be reduced to prove the following: if a smooth function $Q : \mathbb{R}^n \setminus \{0\} \to [0, \infty[$ is 2^+-homogeneous, and the second derivative $Q''(p) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a non-degenerate symmetric bilinear form at each point $p \in \mathbb{R}^n \setminus \{0\}$, then $Q''(p)$ is positive definite. A further reduction is provided by the fact that the index of $Q''(p)$ does not depend on the point of $p \in \mathbb{R}^n \setminus \{0\}$ (this can be seen, e.g., by an immediate continuity argument). Thus it is enough to show that $Q''(p)$ is positive definite at a suitably chosen point p.

The continuity of Q and the compactness of the Euclidean unit sphere $\partial B_1(0)$ implies the existence of a point $e \in \partial B_1(0)$ such that $Q(e) \leqq Q(a)$ for all $a \in \partial B_1(0)$. Let H be the orthogonal complement of the linear span of e. If $v \in H$, then $Q'(e)(v) = 0$, and $Q''(e)(v, v) \geqq 0$. On the other hand, applying (v) of 3.5, it follows that

$$Q''(e)(e, v) = Q'(e)(v) = 0, \quad v \in H.$$

Now let $u \in \mathbb{R}^n \setminus \{0\}$ be any vector. It can uniquely be decomposed in the form $u = \alpha e + v$; $\alpha \in \mathbb{R}$, $v \in H$. Since by Euler's relation $Q''(e)(e, e) = 2Q(e)$, we get

$$Q''(e)(u, u) = 2\alpha^2 Q(e) + Q''(e)(v, v) \geqq 0.$$

This inequality is in fact strict, because $Q''(e)$ is non-degenerate. $\square$

Note The above proof is due to P. Varjú, a student of University of Szeged (Hungary). Another argument can be found in [23].

If (M,F) is a Finsler manifold, then by 3.6 the Lagrange two-form ω_F is non-degenerate, so there exists a unique semispray S on TM such that $i_S\omega_F = -dQ$. Due to the 2^+-homogeneity of Q, S is in fact a spray, called the *canonical spray* of the Finsler manifold.

Lemma 3.9 (fundamental lemma of Finsler geometry) *Let (M,F) be a Finsler manifold. There exists a unique Ehresmann connection $\mathcal{H}$ on TM such that $\mathcal{H}$ is homogeneous, the torsion of $\mathcal{H}$ vanishes, and $dF \circ \mathcal{H} = 0$.*

Proof. We are going to show only the existense statement. As for the uniqueness, which needs more preparation and takes about one page, we refer to [35, p. 1384].

Let S be the canonical spray of (M,F), and consider the Ehresmann connection $\mathcal{H}_S$ induced by S according to the Crampin – Grifone construction. Then, as we have already pointed out, the torsion of $\mathcal{H}_S$ vanishes. Since

$$\mathcal{H}_S \text{ is homogeneous } \overset{\text{def.}}{\iff} \mathbf{t} = 0 \iff \mathbf{i} \circ \mathbf{t} = 0 \iff \left[X^h, C\right] = 0,\ X \in \mathfrak{X}(M),$$

we calculate: $\left[X^h, C\right] = \left[\frac{1}{2}(X^c + [X^v, S]), C\right] = \frac{1}{2}[[X^v,S],C] = -\frac{1}{2}([[S,C],X^v] + [[C,X^v],S]) = \frac{1}{2}([S,X^v] + [X^v,S]) = 0.$

Now we show that $dF \circ \mathcal{H}_S = 0$. Since $dQ = FdF$, and F vanishes nowhere on $\overset{\circ}{T}M$, our claim is equivalent to

$$dQ \circ \mathcal{H}_S = 0 \iff X^hQ = 0 \text{ for all } X \in \mathfrak{X}(M).$$

By the definition of S, $0 = i_S\omega_F + dQ$. We evaluate both sides on a horizontal lift X^h:

$$\begin{aligned}
0 &= \omega_F\left(S, X^h\right) + X^hQ = d\vartheta_F\left(S, X^h\right) + X^hQ = S\tilde{\vartheta}_F\left(\mathbf{j}X^h\right) - X^h\tilde{\vartheta}_F(\mathbf{j}S) \\
&\quad - \tilde{\vartheta}_F\left(\mathbf{j}\left[S, X^h\right]\right) + X^hQ \overset{3.5(\mathrm{v})}{=} S\left(\nabla^v Q\left(\mathbf{j}X^h\right)\right) - X^h(\nabla^v Q(\delta)) \\
&\quad - \nabla^v Q\left(\mathbf{j}\left[S, X^h\right]\right) + X^hQ = S(X^vQ) - 2X^hQ - J\left[S, X^h\right]Q + X^hQ \\
&= [S, X^v]Q - 2X^hQ - X^cQ + X^hQ + X^hQ,
\end{aligned}$$

taking into account that $SQ = dQ(S) = -i_S\omega_F(S) = -\omega_F(S,S) = 0$, and (see [35, 3.3, Cor. 3]) $J\left[S, X^h\right] = X^c - X^h$. Finally, by the definition of $\mathcal{H}_S$ again, $[S,X^v]Q = X^cQ - 2X^hQ$, therefore $-2X^hQ = 0$.

Thus we have proved that the canonical spray of a Finsler manifold induces an Ehresmann connection with the desired properties. □

Note In a coordinate-free form, the fundamental lemma of Finsler geometry was first stated and proved by J. Grifone [12]. We call the Ehresmann connection so described the *Barthel connection* of the Finsler manifold. (Other terms, e.g., 'nonlinear Cartan connection' are also used.) Now we briefly discuss the relation between the fundamental lemma of Finsler geometry and Riemannian geometry.

Consider the type $(0,3)$ tensor $\mathcal{C}_\flat := \nabla^v g_F = \nabla^v\nabla^v\nabla^v Q$ along $\overset{\circ}{\tau}$ and the type $(1,2)$ tensor $\mathcal{C}$ defined by the musical duality given by

$$g_F\left(\mathcal{C}\left(\tilde{X}, \tilde{Y}\right), \tilde{Z}\right) = \mathcal{C}_\flat\left(\tilde{X}, \tilde{Y}, \tilde{Z}\right); \quad \tilde{X}, \tilde{Y}, \tilde{Z} \in \mathfrak{X}(\overset{\circ}{\tau}).$$

$\mathcal{C}$, as well as $\mathcal{C}_\flat$, is called Cartan tensor of the Finsler manifold (M,F). It is easy to see that $\mathcal{C}_\flat$ vanishes if and only if g_F is the lift of a Riemannian metric $g : \mathfrak{X}(M) \times \mathfrak{X}(M) \to C^\infty(M)$ on M in the sense that $g_F = g \circ \tau$. (Then $g_F\left(\hat{X}, \hat{Y}\right) = g(X,Y) \circ \tau$ for all $X, Y \in \mathfrak{X}(M)$.) In this particular case the canonical spray of (M,F) becomes an affine spray, and, as we have already remarked, the Barthel connection induces a covariant derivative D on M such that $(D_X Y)^v = \left[X^h, Y^v\right]$ for all vector fields X, Y on M. It can be shown by an immediate calculation that D is just the Levi-Civita derivative on (M,g). Thus, roughly speaking, *if a Finsler manifold reduces to a Riemannian manifold, then its Barthel connection reduces to the Levi-Civita derivative on the base manifold*.

Covariant derivatives on a Finsler manifold

For the sake of convenient exposition, first we introduce some technical terms. Let a covariant derivative operator D on $\overset{\circ}{\tau}{}^*TM$ and an Ehresmann connection $\mathcal{H}$ on TM be given. We say that D is *strongly associated* to $\mathcal{H}$ if $D\delta = \mathcal{V}$ (= the vertical map belonging to $\mathcal{H}$). The *v-part* D^v and the *h-part* D^h of D are given by $D^v_{\tilde{X}}(\cdot) := D_{\mathbf{i}\tilde{X}}(\cdot)$ and $D^h_{\tilde{X}}(\cdot) := D_{\mathcal{H}\tilde{X}}(\cdot)$, respectively. If g is a (pseudo) Riemannian metric on $\overset{\circ}{\tau}{}^*TM$, then D is called *v-metrical*, *h-metrical* or a *metric derivative*, if $D^v g = 0$, $D^h g = 0$ and $Dg = 0$, respectively. By the *h-Cartan tensor* of a Finsler manifold (M,F) we mean the type $(0,3)$ tensor $\mathcal{C}^h_\flat := \nabla^h g_F$, or the $(1,2)$ tensor $\mathcal{C}^h$ given by

$$g_F\left(\mathcal{C}^h\left(\tilde{X}, \tilde{Y}\right), \tilde{Z}\right) = \mathcal{C}^h_\flat\left(\tilde{X}, \tilde{Y}, \tilde{Z}\right); \quad \tilde{X}, \tilde{Y}, \tilde{Z} \in \mathfrak{X}(\overset{\circ}{\tau})$$

along $\overset{\circ}{\tau}$, where ∇ is the Berwald derivative induced by the Barthel connection. (∇ will be mentioned as the *Finslerian Berwald derivative* on (M,F) in the following.)

From a ‘modern’ point of view, the covariant derivative operators introduced in Finsler geometry by L. Berwald, É. Cartan, S. S. Chern, H. Rund and later by M. Hashiguchi using classical tensor calculus, can be interpreted as covariant derivatives on $\overset{\circ}{\tau}{}^*TM$, specified by some nice properties (compatibility or ‘semi-compatibility’ with the metric tensor and the Barthel connection, vanishing of some torsion tensors). The covariant derivative constructed by É. Cartan in 1934 is an exact analogue of the Levi-Civita derivative on a Riemannian manifold, but it lives on the pull-back bundle $\overset{\circ}{\tau}{}^*TM$. To be more precise, *Cartan’s derivative* $D : \mathfrak{X}\left(\overset{\circ}{T}M\right) \times \mathfrak{X}(\overset{\circ}{\tau}) \to \mathfrak{X}(\overset{\circ}{\tau})$ *is the only covariant derivative on a Finsler manifold* (M,F) *which is metric* $(Dg_F = 0)$*, and whose vertical and horizontal torsion vanish* (the latter is taken with respect to the Barthel connection $\mathcal{H}$). Then D is strongly associated to $\mathcal{H}$, and it is related to the Finslerian Berwald derivative by

$$D^v = \nabla^v + \frac{1}{2}\mathcal{C}, \quad D^h = \nabla^h + \frac{1}{2}\mathcal{C}^h.$$

For a proof of this, as well as the next result, we refer to [35].

In 1943 S. S. Chern, using Cartan’s calculus of differential forms; later, but independently, in 1951, H. Rund by means of tensor calculus, constructed a covariant derivative on a Finsler manifold which may be described in our language as follows:

Lemma 3.10 *Let* (M,F) *be a Finsler manifold, and let* $\mathcal{H}$ *denote its Barthel connection. There exists a unique covariant derivative* D *on* $\overset{\circ}{\tau}{}^*TM$ *such that* $D^v = \nabla^v$, $D^h g = 0$

(i.e., D is h-metrical with respect to $\mathcal{H}$), and the ($\mathcal{H}$-)horizontal torsion of D vanishes. Then D is strongly associated to $\mathcal{H}$, and it is related to the Finslerian Berwald derivative by

$$D^v = \nabla^v, \quad D^h = \nabla^h + \frac{1}{2}\mathcal{C}^h.$$

□

It will be useful to summarize the classical covariant derivatives used in Finsler geometry in a tabular form.

$D = (D^v, D^h)$	$D^v g$	$D^h g$	$T^v(D)$	$T^h(D)$	$D\delta$
Berwald (∇^v, ∇^h)	$\mathcal{C}_\flat$	$\mathcal{C}_\flat^h$	0	$\mathbf{T} = 0$	$\mathbf{t} \circ \mathbf{j} + \mathcal{V} = \mathcal{V}$
Cartan	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathbf{0}$	$\mathcal{V}$
Chern – Rund $D^v := \nabla^v$	$\mathcal{C}_\flat$	$\mathbf{0}$	0	$\mathbf{0}$	$\mathcal{V}$
Hashiguchi $D^h := \nabla^h$	$\mathbf{0}$	$\mathcal{C}_\flat^h$	$\mathbf{0}$	$\mathbf{T} = 0$	$\mathcal{V}$

(Data printed in bold are prescribed.)

It is a historical curiosity that Chern's covariant derivative and Rund's covariant derivative were identified only in 1996 [1]. Another construction of the Chern – Rund derivative can be found in a quite recent paper of H.-B. Rademacher [32]: it is given locally as a *covariant derivative on the base manifold* M, 'parametrized' by a nowhere vanishing vector field on an open subset of M, satisfying some Koszul-type axioms. Now we identify Rademacher's constructions with ours presented in 3.10.

Theorem 3.11 *Let (M, F) be a Finsler manifold. Suppose $\mathcal{U} \subset M$ is an open set, and U is a nowhere vanishing vector field on $\mathcal{U}$. If g_U is defined by*

$$g_U(X, Y) := g_F\left(\hat{X}, \hat{Y}\right) \circ U; \quad X, Y \in \mathfrak{X}(\mathcal{U}),$$

then g_U is a Riemannian metric on $\mathcal{U}$, and there exists a unique covariant derivative

$$D^U : \mathfrak{X}(\mathcal{U}) \times \mathfrak{X}(\mathcal{U}) \to \mathfrak{X}(\mathcal{U}), \quad (X, Y) \mapsto D^U_X Y$$

such that D^U is torsion-free and almost metric in the sense that

$$X g_U(Y, Z) = g_U\left(D^U_X Y, Z\right) + g_U\left(Y, D^U_X Z\right) + \mathcal{C}_\flat\left(\widehat{D^U_X Y}, \hat{Y}, \hat{Z}\right) \circ U$$

for any vector fields X, Y, Z on $\mathcal{U}$. D^U is related to the Chern – Rund derivative D by

$$D^U_X Y = \left(D_{X^c}\hat{Y}\right) \circ U; \quad X, Y \in \mathfrak{X}(\mathcal{U}).$$

Proof. To show the *existence*, we define the mapping D^U by the prescription

$$\left(D^U_X Y\right)(p) := D_{X^c}\hat{Y}(U(p)); \quad X, Y \in \mathfrak{X}(\mathcal{U}),\ p \in \mathcal{U};$$

where D is the Chern – Rund derivative on (M,F). It is then obvious that D^U is additive in both of its variables. To check the $C^\infty(\mathcal{U})$-linearity in X and the Leibniz rule, we choose a smooth function f on $\mathcal{U}$, and using the properties of the Chern – Rund derivative, calculate:

$$\begin{aligned}
D^U_{fX}Y &:= \left(D_{(fX)^c}\hat{Y}\right)\circ U = \left(f^c D_{X^v}\hat{Y} + f^v D_{X^c}\hat{Y}\right)\circ U \\
&= (f^v\circ U)\left(D_{X^c}\hat{Y}\circ U\right) = fD^U_XY \quad \left(f^v := f\circ\mathring{\tau}\right); \\
D^U_X fY &:= \left(D_{X^c}f^v\hat{Y}\right)\circ U = \left((X^cF^v)\hat{Y} + f^v D_{X^c}\hat{Y}\right)\circ U = (Xf)Y + fD^U_XY.
\end{aligned}$$

Taking into account that the horizontal torsion of D vanishes, and hence $0 = D_{X^h}\hat{Y} - D_{Y^h}\hat{X} - \mathbf{j}\left[X^h, Y^h\right] = D_{X^h}\hat{Y} - D_{Y^h}\hat{X} - \widehat{[X,Y]}$, it follows that

$$\begin{aligned}
D^U_XY - D^U_YX &= \left(D_{X^c}\hat{Y} - D_{Y^c}\hat{X}\right)\circ U \\
&= \left(D_{X^h}\hat{Y} - D_{Y^h}\hat{X}\right)\circ U = \widehat{[X,Y]}\circ U = [X,Y],
\end{aligned}$$

i.e., D^U is torsion-free. It remains to show that D^U is almost metric. By Lemma 3.1 and using that D is h-metrical, we get

$$\begin{aligned}
Xg_U(Y,Z) &= X\left(g_F\left(\hat{Y},\hat{Z}\right)\circ U\right) = \left((X^c + [X,U]^v)g_F\left(\hat{Y},\hat{Z}\right)\right)\circ U \\
&= \left(\left(X^h + \mathbf{v}X^c + [X,U]^v\right)g_F\left(\hat{Y},\hat{Z}\right)\right)\circ U = g_F\left(D_{X^h}\hat{Y},\hat{Z}\right)\circ U \\
&\quad + g_F\left(\hat{Y}, D_{X^h}\hat{Z}\right)\circ U + \mathcal{C}_\flat\left(\mathcal{V}X^c + \widehat{[X,U]}, \hat{Y}, \hat{Z}\right)\circ U.
\end{aligned}$$

Finally, we identify the term $\mathcal{V}X^c + \widehat{[X,U]}$. Observe that at each point $p\in\mathcal{U}$, $U^h(U(p)) = \mathcal{H}\circ\hat{U}(U(p)) = \mathcal{H}(U(p),U(p)) = \mathcal{H}\circ\delta(U(p)) = S(U(p))$, where S is the canonical spray of (M,F). Using this, the torsion-freeness of the Barthel connection and the relation $\mathcal{C}^h(\cdot,\delta) = 0$ (as for the latter, see [35, 3.11]), we have

$$\begin{aligned}
&(\mathbf{v}X^c + [X,U]^v)\circ U = \left(\mathbf{v}X^c + \left[X^h, U^v\right] - \left[U^h, X^v\right]\right)\circ U \\
&= \left(\mathbf{v}X^c + \mathbf{i}\nabla_{X^h}\hat{U} - \mathbf{i}\nabla_{U^h}\hat{X}\right)\circ U = \left(\mathbf{v}X^c + \mathbf{i}\nabla_{X^h}\hat{U} - \mathbf{i}\nabla_S\hat{X}\right)\circ U \\
&= \left(\mathbf{v}(X^c - [S,X^v]) + \mathbf{i}\nabla_{X^h}\hat{U} - \frac{1}{2}\mathbf{i}\mathcal{C}^h\left(\hat{X},\delta\right)\right)\circ U = \mathbf{i}\left(D_{X^c}\hat{Y}\right)\circ U \\
&= \left(D^U_XY\right)^v.
\end{aligned}$$

This completes the proof of the existence statement. As for uniqueness, it can be shown (see the cited paper of Rademacher) that if D^U is an almost-metric, torsion-free covariant derivative operator on $\mathcal{U}$, then it obeys a Koszul-type formula, and hence it is uniquely determined. □

Bibliography

[1] M. Anastasiei: A historical remark on the connections of Chern and Rund *In: D. Bao, S.-S. Chern and Z. Shen (eds): Finsler Geometry* (American Mathematical Society, Providence, 1996) 171–176

[2] A. Bastiani: Applications différentiables et variétés différentiables de dimension infinie *J. Anal. Math.* **13** (1964) 1–114

[3] W. Bertram: Differential Geometry, Lie Groups and Symmetric Spaces over General Base Fields and Rings *Memoirs of the AMS* (2006) (to appear); `arXiv:math.DG/0502168`

[4] W. Bertram, H. Glöckner and K.-H. Neeb: Differential calculus over general base fields and rings *Expos. Math.* **22** (2004) 213–282; `arXiv:math.GM/0303300`

[5] Glen E. Bredon: *Topology and Geometry* (Graduate Texts in Mathematics 139, Springer Science & Business Media Inc., New York, 1993)

[6] J. Dieudonné: *Foundations of modern analysis* (Academic Press, New York, 1969)

[7] J. Dugundji: *Topology* (Allyn and Bacon, Boston, 1966)

[8] H. Glöckner and K.-H. Neeb: *Infinite-Dimensional Lie Groups, Vol. I: Basic Theory and Main Examples* (book, in preparation)

[9] H. Glöckner: Infinite-dimensional Lie groups without completeness restrictions *In: A. Strasburger, W. Wojtynski, J. Hilgert and K.-H. Neeb (eds): Geometry and Analysis on finite and infinite-dimensional Lie groups* (Banach Center Publications, 2002) 43–59

[10] H. Glöckner: Implicit functions from Topological Vector Spaces to Banach Spaces *Israel J. Math.* **155** (2006) 205–252

[11] H. Glöckner: Bundles of locally convex spaces, group actions, and hypocontinuous bilinear mappings *manuscript* (submitted) 1–22

[12] J. Grifone: Structure presque tangente et connexions I. *Ann. Inst. Fourier (Grenoble)* **22** (1972) 287–334

[13] R. Hamilton: The inverse function theorem of Nash and Moser *Bull. Amer. Math. Soc.* **7** (1982) 65–222

[14] R. S. Hamilton: Deformation of complex structures on manifolds with boundary *J. Diff. Geom.* **12** (1977) 1–45

[15] S. Helgason: *Groups and Geometric Analysis* (Academic Press, Orlando, 1984)

[16] S. Hiltunen: Implicit functions from locally convex spaces to Banach spaces *Studia Math.* **134** (1999) 235–250

[17] W. Klingenberg: *Riemannian Geometry* (Walter de Gruyter, Berlin–New York, 1982)

[18] A. Kriegl and P. W. Michor: *The Convenient Setting of Global Analysis* (Math. Surveys and Monographs 53, AMS, Providence, 1997)

[19] M. Kuranishi: Deformations of isolated singularities and $\bar{\partial}_b$ *Preprint* Columbia Univ. (1973)

[20] S. Lang: *Fundamentals of Differential Geometry* (Graduate Texts in Mathematics 191, Springer-Verlag, New York, 1999)

[21] J. M. Lee: *Introduction to Smooth Manifolds* (Graduate Texts in Mathematics 218, Springer-Verlag, New York, 2003)

[22] J. Leslie: On the group of real analytic diffeomorphisms of a compact real analytic manifold *Trans. Amer. Math. Soc.* **274** (1982) 651–669

[23] R. L. Lovas: A note on Finsler – Minkowski norms *Houston Journal of Mathematics* (to appear)

[24] T. W. Ma: Inverse mapping theorems on coordinate spaces *Bull. London Math. Soc.* **33** (2001) 473–482

[25] A. D. Michel: Differential calculus in linear topological spaces *Proc. Nat. Acad.*

Sci. **24** (1938) 340–342

[26] J. Milnor: On infinite-dimensional Lie groups *Preprint* Institute of Advanced Study, Princeton (1982)

[27] J. Milnor: Remarks on infinite-dimensional Lie groups *In: B. De Witt, R. Stora (eds): Relativité, groupes et topologie II Les Houches* (Éc. d'Été Phys. Théor., session 40, 1983, 1984) 1007-1057

[28] J. Moser: A rapidly convergent iteration method and non-linear differential equations *Ann. Scuola Norm. Sup. Pisa (3)* **20** (1966) 499–535

[29] R. Narasimhan: *Analysis on real and complex manifolds* (Masson & Cie, Paris, 1968)

[30] J. Nash: The embedding problem for Riemannian manifolds *Ann. of Math. (2)* **63** (1956) 20–63

[31] K.-H. Neeb: Nancy Lectures on Infinite-Dimensional Lie Groups *Preprint* (2002)

[32] H.-B. Rademacher: Nonreversible Finsler Metrics of Positive Flag Curvature *In: Riemann – Finsler Geometry* (MSRI Publications 50, 2004) 261-302

[33] W. Rudin: *Functional Analysis* (McGraw-Hill, New York, 1973)

[34] F. Sergeraert: Une généralisation du théorème des fonctions implicites de Nash *C. R. Acad. Sci. Paris, Ser. A* **270** (1970) 861–863

[35] J. Szilasi: A Setting for Spray and Finsler Geometry *In: P. L. Antonelli (ed): Handbook of Finsler Geometry Vol. 2* (Kluwer Academic Publishers, Dordrecht, 2003) 1183-1426

[36] A. C. Thompson: *Minkowski geometry* (Encyclopedia of mathematics and its applications 63, Cambridge University Press, Cambridge, 1996)

[37] E. Zehnder: Generalized implicit function theorems with applications to small divisor problems I, II *Comm. Pure Appl. Math* **28** (1975) 91–141; **29** (1976) 49–114

JÓZSEF SZILASI

University of Debrecen, Institute of Mathematics,
H–4010 Debrecen, P.O. Box 12, Hungary

E-mail: `szilasi@math.klte.hu`

REZSŐ L. LOVAS

University of Debrecen, Institute of Mathematics,
H–4010 Debrecen, P.O. Box 12, Hungary

E-mail: `lovasr@math.klte.hu`