Wolfgang Bertram
(Institut Elie Cartan, Nancy)

Differential Geometry over General Base Fields and Rings

Version to appear in the Memoirs of the AMS.

Last updated: november 2007

Contents

Preface

Die Regeln des Endlichen gelten im Unendlichen weiter.
(G. W. Leibniz, 1702)

Classical, real finite-dimensional differential geometry and Lie theory can be generalized in several directions. For instance, it is interesting to look for an *infinite dimensional* theory, or for a theory which works over other base fields, such as the p-adic numbers or other topological fields. The present work aims at both directions at the same time. Even more radically, we will develop the theory over certain base *rings* so that, in general, there is no dimension at all. Thus, instead of opposing "finite" and "infinite-dimensional" or "real" and "p-adic" theories, we will simply speak of "general" Lie theory or differential geometry. Certainly, this is a daring venture, and therefore I should first explain the origins and the aims of the project and say what can be done and what can *not* be done.

Preliminary versions of the present text have appeared as a series of five preprints during the years 2003 – 2005 at the Institut Elie Cartan, Nancy (still available on the web-server `http://www.iecn.u-nancy.fr/Preprint/`). The title of this work is, of course, a reference to Helgason's by now classical monograph [Hel78] from which I first learned the theory of differential geometry, Lie groups and symmetric spaces. Later I became interested in the more algebraic aspects of the theory, related to the interplay between Lie- and Jordan-theory (see [Be00]). However, rather soon I began to regret that differential geometric methods seemed to be limited to the case of the base field of real or complex numbers and to the case of "well-behaved" topological vector spaces as model spaces, such as finite-dimensional or Banach spaces (cf. the remark in [Be00, Ch.XIII] on "other base fields"). Rather surprisingly (at least for me), these limitations were entirely overcome by the joint paper [BGN04] with Helge Glöckner and Karl-Hermann Neeb where differential calculus, manifolds and Lie groups have been introduced in a very general setting, including the case of manifolds modelled on arbitrary topological vector spaces over any non-discrete topological field, and even over topological modules over certain topological base rings. In the joint paper [BeNe05] with Karl-Hermann Neeb, a good deal of the results from [Be00] could be generalized to this very general framework, leading, for example, to a rich supply of infinite-dimensional symmetric spaces.

As to differential geometry on infinite dimensional manifolds, I used to have the impression that its special flavor is due to its, sometimes rather sophisticated, functional analytic methods. On the other hand, it seemed obvious that the "purely differential" aspects of differential geometry are algebraic in nature and thus should be understandable without all the functional analytic flesh around it. In other words, one should be able to formalize the fundamental differentiation process such that its general algebraic structure becomes visible. This is indeed possible, and one possibility is opened by the general differential calculus mentioned above: for me, the most striking result of this calculus is the one presented in Chapter 6 of this work (Theorem 6.2), saying that the usual "tangent functor" T of differential geometry can, in this context, be interpreted as a functor of scalar extension from the base field or ring $\mathbb{K}$ to the *tangent ring* $T\mathbb{K} = \mathbb{K} \oplus \varepsilon\mathbb{K}$ with relation $\varepsilon^2 = 0$, sometimes also called the *dual numbers over* $\mathbb{K}$. Even though the result itself may not seem surprising (for instance, in algebraic geometry it is currently used), it seems that the proof in our context is the most natural one that one may imagine. In some sense, all the rest of the text can be seen as a try to exploit the consequences of this theorem.

I would like to add a remark on terminology. Classical "calculus" has two faces, namely *differential calculus* and *integral calculus.* In spite of its name, classical "differential geometry" also has these two faces. However, in order to avoid misunderstandings, I propose to clearly distinguish between "differential" and "integral" geometry, the former being the topic of the present work, the latter being restricted to special contexts where various kinds of integration theories can be implemented (cf. Appendix L). In a nutshell, differential geometry in this sense is the theory of k-th order Taylor expansions, for any $k \in \mathbb{N}$, of objects such as manifolds, vector bundles or Lie groups, without requiring convergence of the Taylor series. In the French literature, this approach is named by the somewhat better adapted term "développement limité" (cf., e.g., [Ca71]), which could be translated by "limited expansion". Certainly, to a mathematician used to differential geometry in real life, such a rigorous distinction between "differential" and "integral" geometry may seem pedantic and purely formal – probably, most readers would prefer to call "formal differential geometry" what remains when integration is forbidden. However, the only "formal" construction that we use is the one of the tangent bundle, and hence the choice of terminology depends on wether one considers the tangent bundle as a genuine object of differential geometry or not. Only in the very last chapter (Chapter 32) we pass the frontier separating true manifolds from formal objects by taking the projective limit $k \to \infty$ of our k-th order Taylor expansions, thus making the link with approaches by formal power series.

The present text is completely self-contained, but by no means it is an exhaustive treatise on the topics mentioned in the title. Rather, I wanted to present a first approximation of a new theory that, in my opinion, is well-adapted to capture some very general aspects of differential calculus and differential geometry and thus might be interesting for mathematicians and physicists working on topics that are related to these aspects.

Acknowledgements. I would not have dared to tackle this work if I had not had teachers who, at a very early stage, attracted my interest to foundational questions of calculus. In particular, although I do not follow the lines of thought of non-standard analysis, I am indebted to Detlef Laugwitz from whose reflections on the foundations of analysis I learned a lot. (For the adherents of non-standard analysis, I add the remark that the field ${}^*\mathbb{R}$ of non-standard numbers is of course admitted as a possible base field of our theory.)

I would also like to thank Harald Löwe for his careful proofreading of the first part of the text, and all future readers for indicating to me errors that, possibly, have remained.

Wolfgang Bertram Nancy, february 2005.

Abstract. The aim of this work is to lay the foundations of differential geometry and Lie theory over the general class of topological base fields and -rings for which a differential calculus has been developed in [BGN04], without any restriction on the dimension or on the characteristic. Two basic features distinguish our approach from the classical real (finite or infinite dimensional) theory, namely the interpretation of tangent- and jet functors as functors of *scalar extensions* and the introduction of *multilinear bundles* and *multilinear connections* which generalize the concept of vector bundles and linear connections.

Received by the editors: February 9, 2005

AMS subject classification (2000):

Primary: 22E65, 53B05, 53C35, 58A05, 58A20, 53B05, 58A32

Secondary: 22E20, 14L05, 15A69, 51K10, 58B20

Key words: Lie groups over topological fields, infinite-dimensional Lie groups, diffeomorphism groups, formal group, general and special multilinear group, symmetric space, differentiable manifold, tangent functor, linear and multilinear bundle, jet, linear connection, multilinear connection, curvature, exponential map, exponential jet, Campbell-Hausdorff formula, multilinear algebra, scalar extension, dual numbers, synthetic differential geometry, Weil functor

Author's address:

Wolfgang Bertram
Université Henri Poincaré (Nancy I)
Institut Elie Cartan
B.P. 239
54506 Vandœuvre-lès-Nancy Cedex
France
e-mail: `bertram@iecn.u-nancy.fr`
http: `http://www.iecn.u-nancy.fr/~bertram`

0. Introduction

The classical setting of differential geometry is the framework of finite-dimensional real manifolds. Later, parts of the theory have been generalized to various kinds of infinite dimensional manifolds (cf. [La99]), to analytic manifolds over ultrametric fields (finite-dimensional or Banach, cf. [Bou67] or [Se65]) or to more formal concepts going beyond the notion of manifolds such as "synthetic differential geometry" and its models, the so-called "smooth toposes" (cf. [Ko81] or [MR91]). The purpose of the present work is to unify and generalize some aspects of differential geometry in the context of manifolds over very general base fields and even -rings, featuring fundamental structures which presumably will persist even in further generalizations beyond the manifold context. We have divided the text into seven main parts of which we now give a more detailed description.

I. Basic Notions

0.1. *Differential calculus.* Differential geometry may be seen as the "invariant theory of differential calculus", and *differential calculus* deals, in one way or another, with the "infinitesimally small", either in the language of limits or in the language of infinitesimals – in the words of G.W. Leibniz: "Die Regeln des Endlichen gelten im Unendlichen weiter" [1] (quoted from [Lau86, p. 88]). In order to formalize this idea, take a topological commutative unital ring $\mathbb{K}$ (such as $\mathbb{K} = \mathbb{R}$ or any other model of the continuum that you prefer) and let "das Endliche" ("the finite") be represented by the invertible scalars $t \in \mathbb{K}^\times$ and "das unendlich (Kleine)" ("the infinitely (small)") by the non-invertible scalars. Then "Leibniz principle" may be interpreted by the assumption that $\mathbb{K}^\times$ is dense in $\mathbb{K}$: the behaviour of a continuous mapping at the infinitely small is determined by its behavior on the "finite universe". For instance, if $f : V \supset U \to W$ is a map defined on an open subset of a topological $\mathbb{K}$-module, then for all invertible scalars t in some neighborhood of 0, the "slope" (at $x \in U$ in direction $v \in V$ with parameter t)

$$f^{[1]}(x, v, t) := \frac{f(x + tv) - f(x)}{t} \tag{0.1}$$

is defined. We say that f is *of class C^1 (on U)* if the slope extends to a *continuous* map $f^{[1]} : U^{[1]} \to W$, where $U^{[1]} = \{(x, v, t) | \, x + tv \in U\}$. By Leibniz' principle, this extension is unique if it exists. Put another way, the fundamental relation

$$f(x + tv) - f(x) = t \cdot f^{[1]}(x, v, t) \tag{0.2}$$

continues to hold for all $(x, v, t) \in U^{[1]}$. The *derivative of f at x (in direction v)* is the continuation of the slope to the most singular of all infinitesimal elements, namely to $t = 0$:

$$df(x)v := f^{[1]}(x, v, 0). \tag{0.3}$$

[1] "The rules of the finite continue to hold in the infinite."

Now all the basic rules of differential calculus are easily proved, from the Chain Rule via Schwarz' Lemma up to Taylor's formula (see Chapter 1 and [BGN04]) – the reader who compares the proofs with the "classical" ones will find that our approach at the same time simplifies and generalizes this part of calculus.

0.2. *Differential geometry versus integral geometry.* The other half of Newton's and Leibniz' "calculus" is *integral* calculus. For the time being, the simple equation (0.2) has no integral analog – in other words, we cannot reverse the process of differentiation over general topological fields or rings (not even over the real numbers if we go too far beyond the Banach-space set-up). Thus in general we will not even be able to determine the space of all solutions of a "trivial" differential equation such as $df = 0$, and therefore no integration will appear in our approach to differential geometry: there will be no existence theorems for flows, for geodesics, or for parallel transports, no exponential map, no Poincaré lemma, no Frobenius theorem... what remains is really *differential* geometry, in contrast to the mixture of integration and differentiation that we are used to from the finite-dimensional real theory. At a first glance, the remaining structure may look very poor; but we will find that it is in fact still tremendously complicated and rich. In fact, we will see that many classical notions and results that usually are formulated on the global or local level or at least on the level of germs, continue to make sense on the level of *jets (of arbitrary order)*, and although we have in general no notion of convergence, we can take some sort of a limit of these objects, namely a *projective limit.*

II. Interpretation of tangent objects via scalar extensions

0.3. *Justification of infinitesimals.* We define *manifolds* and *tangent bundles* in the classical way using charts and atlasses (Ch. 2 and 3); then, intuitively, we may think of the tangent space T_pM of M at p as a "(first order) infinitesimal neighborhood" of p. This idea may be formalized by writing, with respect to some fixed chart of M, a tangent vector at the point p in the form $p + \varepsilon v$, where ε is a "very small" quantity (say, Planck's constant), thus expressing that the tangent vector is "infinitesimally small" compared to elements of M (in a chart). The property of being "very small" can mathematically be expressed by requiring that ε^2 is zero, compared with all "space quantities". This suggests that, if M is a manifold modelled on the $\mathbb{K}$-module V, then the tangent bundle should be modelled on the space $V \oplus \varepsilon V := V \times V$, which is considered as a module over the ring $\mathbb{K}[\varepsilon] = \mathbb{K} \oplus \varepsilon\mathbb{K}$ of *dual numbers over* $\mathbb{K}$ (it is constructed from $\mathbb{K}$ in a similar way as the complex numbers are constructed from $\mathbb{R}$, replacing the condition $i^2 = -1$ by $\varepsilon^2 = 0$). All this would be really meaningful if TM were a manifold not only over $\mathbb{K}$, but also over the extended ring $\mathbb{K}[\varepsilon]$. In Chapter 6 we prove that this is indeed true: the "tangent functor" T can be seen as a functor of scalar extension by dual numbers (Theorem 6.2). Hence one may use the "dual number unit" ε when dealing with tangent bundles with the same right as the imaginary unit i when dealing with complex manifolds. The proof of Theorem 6.2 is conceptual and allows to understand why dual numbers naturally appear in this context: the basic idea is that one can "differentiate the definition of differentiability" – we write (0.2) as a commutative diagram, apply the tangent functor to this diagram and get a diagram of the same form, but taken over the tangent ring $T\mathbb{K}$ of $\mathbb{K}$. The upshot is now that $T\mathbb{K}$ is nothing but $\mathbb{K}[\varepsilon]$.

Theorem 6.2 may be seen as a bridge from differential geometry to algebraic geometry where the use of dual numbers for modelizing tangent objects is a standard technique (cf. [GD71], p. 11), and also to classical geometry (see article of Veldkamp in [Bue95]). It is all the more surprising that in most textbooks on differential geometry no trace of the dual number unit ε can be found – although it clearly leaves a trace which is well visible already in the usual real framework: in the same way as a complex structure induces an almost complex structure on the underlying real manifold, a dual number structure induces a tensor field of endomorphisms having the property that its square is zero; let us call this an "almost dual structure". Now, there is a *canonical* almost dual structure on every tangent bundle (the almost trivial proof of this is given in Section 4.6). This canonical almost dual structure is "integrable" in the sense that its kernel is a distribution of subspaces that admits integral submanifolds – namely the fibers of the bundle TM. Therefore Theorem 6.2 may be seen as an analog of the well-known theorem of Newlander and Nirenberg: our integrable almost dual structure induces on TM the structure of a manifold over the ring $\mathbb{K}[\varepsilon]$.

Some readers may wish to avoid the use of rings which are not fields, or even to stay in the context of the real base field. In principle, all our results that do not directly involve dual numbers can be proved in a "purely real" way, i.e., by interpreting ε just as a formal (and very useful !) label which helps to distinguish two copies of V which play different rôles (fiber and base). But in the end, just as the "imaginary unit" i got its well-deserved place in mathematics, so will the "infinitesimal unit" ε.

0.4. *Further scalar extensions.* The suggestion to use dual numbers in differential geometry is not new – one of the earliest steps in this direction was by A. Weil ([W53]); one of the most recent is [Gio03]. However, most of the proposed constructions are so complicated that one is discouraged to iterate them. But this is exactly what makes the dual number formalism so useful: for instance, if TM is the scalar extension of M by $\mathbb{K}[\varepsilon_1]$, then the double tangent bundle $T^2M := T(TM)$ is simply the scalar extension of M by the ring

$$TTK := T(T\mathbb{K}) = \mathbb{K}[\varepsilon_1][\varepsilon_2] \cong \mathbb{K} \oplus \varepsilon_1\mathbb{K} \oplus \varepsilon_2\mathbb{K} \oplus \varepsilon_1\varepsilon_2\mathbb{K}, \tag{0.4}$$

and so on for all higher tangent bundles T^kM. As a matter of fact, most of the important notions of differential geometry deal, in one way or another, with the second order tangent bundle T^2M (e.g. Lie bracket, exterior derivative, connections) or with T^3M (e.g. curvature) or even higher order tangent bundles (e.g. covariant derivative of curvature). Therefore second and third order differential geometry really is the central part of all differential geometry, and finding a good notation concerning second and third order tangent bundles becomes a necessity. Most textbooks, if at all TTM is considered, use a component notation in order to describe objects related to this bundle. In this situation, the use of *different* symbols $\varepsilon_1, \varepsilon_2, \ldots$ for the infinitesimal units of the various scalar extensions is a great notational progress, combining algebraic rigour and transparency. It becomes clear that many structural features of T^kM are simple consequences of corresponding structural features of the rings $T^k\mathbb{K} = \mathbb{K}[\varepsilon_1, \ldots, \varepsilon_k]$ (cf. Chapter 7). For instance, on both objects there is a canonical action of the permutation group Σ_k; the subring $J^k\mathbb{K}$ of $T^k\mathbb{K}$ fixed under this action is called a *jet ring* (Chapter 8) – in the case of characteristic zero this is a truncated polynomial ring $\mathbb{K}[x]/(x^{k+1})$. Therefore the subbundle J^kM of T^kM fixed under the action of Σ_k can be seen as the scalar extension of M by the jet ring $J^k\mathbb{K}$; it can be interpreted as the *bundle of k-jets of (curves in) M*. The bundles T^kM and J^kM are the stage on which higher order differential geometry is played.

III. Second order differential geometry

0.5. *Bilinear bundles and connections.* The theory of *linear connections* and their *curvature* is a chief object of general differential geometry in our sense. When studying higher tangent bundles $T^2M, T^3M, \ldots$ or tangent bundles $TF, T^2F, \ldots$ of linear (i.e. vector-) bundles $p: F \to M$, one immediately encounters the problem that these bundles are *not linear bundles over M*. One may ask, naively: "if it is not a linear bundle, so what is it then ?" The answer can be found, implicitly, in all textbooks where the structure of TTM (or of TF) is discussed in some detail (e.g. [Bes78], [La99], [MR91]), but it seems that so far no attempt has been made to separate clearly linear algebra from differential calculus. Doing this, one is lead to a purely algebraic concept which we call *bilinear space* (cf. Appendix Multilinear Geometry). A *bilinear bundle* is simply a bundle whose fibers carry a structure of bilinear space which is preserved under change of bundle charts (Chapter 9). For instance, TTM and TF are bilinear bundles. Basically, a bilinear space E is given by a whole collection of linear (i.e., $\mathbb{K}$-module-) structures, parametrized by some other space E', satisfying some purely algebraic axioms. We then say that these linear structures are *bilinearly related*. A *linear connection* is simply given by singling out, in a fiberwise smooth way, in each fiber *one* linear structure among all the bilinearly related linear structures.

Of course, in the classical case our definition of a linear connection coincides with the usual ones – in fact, in the literature one finds many different definitions of a connection, and, being aware of the danger that we would be exposed to[1], it was not our aim to add a new item to this long list. Rather, we hope to have found the central item around which one can organize the spider's web of related notions such as *sprays, connectors, connection one-forms* and *covariant derivatives* (Chapters 10, 11, 12). We resisted the temptation to start with general Ehresmann connections on general fiber bundles. In our approach, principal bundles play a much less important rôle than in the usual framework – a simple reason for this is that the general (continuous) linear automorphism group $\mathrm{Gl}_{\mathbb{K}}(V)$ of a topological $\mathbb{K}$-module is in general no longer a Lie group, and hence the prime example of a principal bundle, the frame bundle, is not at our disposition.

---

[1] [Sp79, p. 605]: "I personally feel that the next person to propose a new definition of a connection should be summarily executed."

IV. Third and higher order differential geometry

0.6. *Multilinear bundles and multilinear connections.* The step from second order to third order geometry is important and conceptually not easy since it is at this stage that *non-commutativity* enters the picture: just as TTM is a bilinear bundle, the higher order tangent bundles $T^k M$ carry the structure of a *multilinear bundle* over M. The fibers are *multilinear spaces* (Chapter MA), which again are defined to be a set E with a whole collection of linear structures (called *multilinearly related*), parametrized by some space E', and a *multilinear connection* is given by fixing *one* of these multilinearly related structures (Chapters 15 and 16). In case $k = 2$, E' is an *abelian* (in fact, vector) group, and therefore the space of linear connections is an affine space over $\mathbb{K}$. For $k > 2$, the space E' is a *non-abelian* group, called the *special multilinear group* $\mathrm{Gm}^{1,k}(E)$.

0.7. *Curvature.* Our strategy of analyzing the higher order tangent bundle $T^k M$ is by trying to define, in a most canonical way, a sequence of multilinear connections on $TTM, T^3M, \ldots$, starting with a linear connection in the sense explained above. In other words, we want to define vector bundle structures over M on $TTM, T^3M, \ldots$, in such a way that they define, by restriction, also vector bundle structures over M on the jet bundles $J^1M, J^2M, \ldots$ which are compatible with taking the projective limit $J^\infty M$. Thus on the filtered, non-linear bundles $T^k M$ and $J^k M$ we wish to define a sequence of graded and linear structures.

It turns out that there is a rather canonical procedure how to define such a sequence of linear structures: we start with a linear connection L and define suitable derivatives $DL, D^2L, \ldots$ (Chapter 17). Unfortunately, since second covariant derivatives do in general not commute, the multilinear connection $D^k L$ on $T^{k+1}M$ is in general not invariant under the canonical action of the permutation group Σ_{k+1}. Thus, in the worst case, we get $(k+1)!$ different multilinear connections in this way. Remarkably enough, they all can be restricted to the jet bundle $J^{k+1}M$ (Theorem 19.2), but in general define still $k!$ different linear structures there. Thus, for $k = 2$, we get 2 linear structures on TTM: their difference is the *torsion* of L. For $k = 3$, we get 6 linear structures on T^3M, but their number reduces to 2 if L was already chosen to be torsionfree. In this case, their "difference" is a tensor field of type $(2, 1)$, agreeing with the classical *curvature tensor* of L (Theorem 18.3). The induced linear structure on J^3M is then unique, which is reflected by *Bianchi's identity* (Chapter 18 and Section SA.10). For general $k \geq 3$, we get a whole bunch of *curvature operators*, whose combinatorial structure seems to be an interesting topic for further work.

If L is flat (i.e, it has vanishing torsion and curvature tensors), then the sequence $DL, D^2L, \ldots$ is indeed invariant under permutations and under the so-called *shift operators* (Chapter 20), which are necessary conditions for integrating them to a "canonical chart" associated to a connection (Theorem 20.7). For general connections, the problem of defining a permutation and shift-invariant sequence of multilinear connections is far more difficult (see Section 0.16, below).

0.8. *Differential operators, duality and "representation theory."* A complete theory of differential geometry should not exclude the cotangent-"functor" T^*, which in some sense plays a more important rôle in most modern theories than the tangent functor T. In fact, the duality between T and T^* corresponds to a duality between *curves* $\gamma : \mathbb{K} \to M$ and *(scalar) functions* $f : M \to \mathbb{K}$. If the theory is rather based on curves, then the tangent functor will play a dominant role since $\gamma'(t)$ is a tangent vector, and if the theory is rather based on functions, then T^* will play a dominant role since $df(x)$ is a cotangent vector. Under the influence of contemporary algebraic geometry, most modern theories adopt the second point of view. By choosing the first one, we do not want to take a decision on "correctness" of one point of view or the other, but, being interested in the "covariant theory", we just simplify our life and avoid technical problems such as double dualization and topologies on dual spaces which inevitably arise if one tries to develop a theory of the tangent functor in an approach based on scalar functions. This allows to drop almost all restrictions on the model space in infinite dimension and opens the theory for the use of base fields of positive characteristic and even of base rings (giving the whole strength to the point of view of scalar extensions).

However, although we do not define, e.g., tangent vectors as differential operators acting on functions, they may be *represented* by such operators. Similarly, connections may be represented by covariant derivatives of sections, and so on. In other words, one can develop a *representation theory*, representing geometric objects from the "covariant setting", in a contravariant way, by operators on functions and sections – some sketchy remarks on the outline of such a theory are given in Chapter 20, the theory itself remains to be worked out. In Chapter 21, we define a particularly important differential operator, namely the *exterior derivative* of forms.

V. Lie Theory

0.9. *The Lie algebra of a Lie group.* By definition, *Lie groups* are manifolds with a smooth group structure (Chapter 5). Then, as in the classical real case, the Lie algebra of a Lie group can be defined in various ways:

(a) one may define the Lie bracket via the bracket of (left or right) invariant vector fields (as is done in most textbooks), or

(b) one may express it by a "Taylor expansion" of the group commutator $[g,h] = ghg^{-1}h^{-1}$ (as is done in [Bou72] and [Se65]; the definition by deriving the adjoint representation $\mathrm{Ad} : G \to \mathrm{Gl}(\mathfrak{g})$ is just another version of this).

For convenience of the reader, we start with the first definition (Chapter 5), and show (Theorem 23.2) that it agrees with the following version of the second definition: as in the classical theory, the tangent bundle TG of a Lie group G carries again a canonical Lie group structure, and by induction, all higher order tangent bundles T^kG are then Lie groups. The projection $T^kG \to G$ is a Lie group homomorphism, giving rise to an exact sequence of Lie groups

$$1 \quad \to \quad (T^kG)_e \quad \to \quad T^kG \quad \to \quad G \quad \to 1\,. \tag{0.5}$$

For $k=1$, the group $\mathfrak{g} := T_eG$ is simply the tangent space at the origin with vector addition. For $k=2$, the group $(T^2G)_e$ is no longer abelian, but it is two-step nilpotent, looking like a Heisenberg group: there are three abelian subgroups

$\varepsilon_1\mathfrak{g}, \varepsilon_2\mathfrak{g}, \varepsilon_1\varepsilon_2\mathfrak{g}$, all canonically isomorphic to $(\mathfrak{g}, +)$, such that the commutator of elements from the first two of these subgroups can be expressed by the Lie bracket in $\mathfrak{g}$ via the relation

$$\varepsilon_1 X \cdot \varepsilon_2 Y \cdot (\varepsilon_1 X)^{-1} \cdot (\varepsilon_2 Y)^{-1} = \varepsilon_1\varepsilon_2[X, Y]. \tag{0.6}$$

0.10. *The Lie bracket of vector fields.* Definition (a) of the Lie bracket in $\mathfrak{g}$ requires a foregoing definition of the Lie bracket of vector fields. In our general framework, this is more complicated than in the classical context because we do not have a faithful representation of the space $\mathfrak{X}(M)$ of vector fields on M by differential operators acting on functions (cf. Section 0.8). It turns out (Chapter 14) that a conceptual definition of the Lie bracket of vector fields is completely analogous to Equation (0.6), where $\mathfrak{g}$ is replaced by $\mathfrak{X}(M)$ and $(T^2G)_e$ by the space $\mathfrak{X}^2(M)$ of smooth sections of the projection $TTM \to M$. The space $\mathfrak{X}^2(M)$ carries a canonical group structure, again looking like a Heisenberg group (Theorem 14.3). Strictly speaking, for our approach to Lie groups, Definition (a) of the Lie bracket is not necessary. The main use of left- or right-invariant vector fields in the classical theory is to be integrated to a flow which gives the exponential map. In our approach, this is not possible, and we have to use rather different strategies.

0.11. *Left and right trivializations and analog of the Campbell-Hausdorff formula.* Following our general philosophy of multilinear connections, we try to linearize $(T^kG)_e$ by introducing a suitable "canonical chart" and to describe its group structure by a multilinear formula. This is possible in two canonical ways: via the zero section, the exact sequences (0.5) split, and hence we may write T^kG as a semidirect product by letting act G from the right or from the left. By induction, one sees in this way that $(T^kG)_e$ can be written via left- or right trivialization as an iterated semidirect product involving $2^k - 1$ copies of $\mathfrak{g}$. Relation (0.6) generalizes to a general "commutation relation" for this group, so that we may give an explicit formula for the group structure on the space $(T^k\mathfrak{g})_0$ which is just a certain direct sum of $2^k - 1$ copies of $\mathfrak{g}$. The result (Theorem 24.7) may be considered as a rather primitive version of the Campbell-Hausdorff formula. It has the advantage that its combinatorial structure is fairly transparent and that it works in arbitrary characteristic, and it has the drawback that inversion is more complicated than just taking the negative (as in the Campbell-Hausdorff group chunk). It should be interesting to study the combinatorial aspects of this formula in more detail, especially for the theory of Lie groups in positive characteristic.

0.12. *The exponential map.* As already mentioned, one cannot construct an exponential mapping $\exp : \mathfrak{g} \to G$ in our general context. However, the groups $(T^kG)_e$ have a nilpotent Lie algebra $(T^k\mathfrak{g})_0$, and thus one would expect that a polynomial exponential map $\exp_k : (T^k\mathfrak{g})_0 \to (T^kG)_e$ (with polynomial inverse $\log_k$) should exist. This is indeed true, provided that $\mathbb{K}$ is a field of characteristic zero (Theorem 25.2). Our proof of this fact is purely algebraic (Appendix PG, Theorem PG.6) and uses only the fact that the above mentioned analog of the Campbell-Hausdorff formula defines on $(T^kG)_e$ the structure of a *polynomial group* (see Section 0.18 below). By restriction to the jet bundle J^kG, we get an exponential map $\exp_k$ for the nilpotent groups $(J^kG)_e$, which form of projective system so that, as projective limit, we get an exponential map $\exp_\infty$ of $(J^\infty G)_e$ (Chapter 32). If there is a notion of convergence such that $\exp_\infty$ and $\log_\infty$ define convergent series, then they can indeed be used to define a "canonical chart" and thus a "canonical atlas" of G.

0.13. *Symmetric spaces.* The definition and basic theory of *symmetric spaces* follows the one given by O. Loos in [Lo69] (see also [BeNe05], where examples are constructed): a symmetric space is a manifold M equipped with a family of symmetries $\sigma_x : M \to M$, $x \in M$, such that the map $(x,y) \mapsto \mu(x,y) := \sigma_x(y)$ is smooth and satisfies certain natural axioms (Chapter 5). In [Lo69], the higher order theory of symmetric spaces is developed by using the *linear* bundles $T^{(k)}M$ whose sections are k-th order differential operators. Again following our general philosophy, we replace these bundles by the iterated tangent bundles T^kM which are not linear, but have the advantage to be themselves symmetric spaces in a natural way. In particular, we construct the *canonical connection* of M in terms of the symmetric space TTM (Chapter 26) and express its curvature by the symmetric space structure on T^3M (Chapter 27). The construction of a family $\exp_k : (T^k(T_oM))_0 \to (T^kM)_o$ of exponential mappings (whose projective limit $\exp_\infty$ can again be seen as a formal power series version of the missing exponential map $\mathrm{Exp} : T_oM \to M$) uses some general results to be given in Part VI (below) and is achieved in Chapter 30.

VI. Groups of diffeomorphisms and the exponential jet

0.14. *The group* $\mathrm{Diff}(M)$. If M is a manifold, then the group $G := \mathrm{Diff}(M)$ of all diffeomorphisms of M is in general not a Lie group. However, quite often it is useful to think of G as a Lie group with Lie algebra $\mathfrak{X}(M)$, the Lie algebra of vector fields on M, and exponential map associating to a vector field X the flow $\mathrm{Fl}_t^X \in G$ at time $t = 1$ (whenever possible). Then, according to our approach to Lie theory, we would like to analyze the group G by studying the higher order tangent group T^kG and its fiber $(T^kG)_e$ over the origin $e \in G$. It turns out that this approach makes perfectly sense for any manifold M over a general base field or -ring $\mathbb{K}$: we prove that the group $\mathrm{Diff}_{T^k\mathbb{K}}(T^kM)$ of all diffeomorphisms of T^kM which are smooth over the iterated tangent ring $T^k\mathbb{K}$, behaves in all respects like the k-th order tangent group of G. More precisely, there is an exact sequence of groups

$$1 \;\to\; \mathfrak{X}^k(M) \;\to\; \mathrm{Diff}_{T^k\mathbb{K}}(T^kM) \;\to\; \mathrm{Diff}(M) \;\to\; 1 \qquad (0.7)$$

which is the precise analog of (0.5). In particular, $\mathfrak{X}^k(M)$, the space of smooth sections of the projection $T^kM \to M$, carries a canonical group structure which turns it into a polynomial group (Theorem 28.3). For $\mathbb{K} = \mathbb{R}$ and M finite-dimensional, it is known that $\mathfrak{X}^k(M)$ carries a natural group structure (this result is due to Kainz and Michor, [KM87, Theorem 4.6], see also [KMS93, Theorem 37.7]), but the interpretation via the sequence (0.7) seems to be entirely new. A splitting of the exact sequence (0.7) is simply given by $\mathrm{Diff}(M) \to \mathrm{Diff}_{T^k\mathbb{K}}(T^kM)$, $g \mapsto T^kg$, and thus $\mathrm{Diff}_{T^k\mathbb{K}}(T^kM)$ is a semidirect product of $\mathrm{Diff}(M)$ and $\mathfrak{X}^k(M)$. As for Lie groups, we can now define left- and right-trivializations and have an analog of the Campbell-Hausdorff formula for the group structure on $\mathfrak{X}^k(M)$ (Theorem 29.2).

0.15. *Flow of a vector field.* Even in the classical situation, not every vector field $X \in \mathfrak{X}(M)$ can be integrated to a global flow $\mathrm{Fl}_t^X : M \to M$, $t \in \mathbb{R}$. In our general set-up, this cannot be done even locally. But to some extent, the following construction can be seen as an analog of the flow of a vector field: as remarked above, $\mathfrak{X}^k(M)$ carries the structure of a polynomial group, and thus by Theorem PG.6, if $\mathbb{K}$ is a field of characteristic zero, there is a bijective exponential map $\exp_k$ associated to this group. Since $\mathfrak{X}^k(M)$ plays the rôle of $(T^k G)_e$ for $G = \mathrm{Diff}(M)$, the "Lie algebra" of $\mathfrak{X}^k(M)$ should be $(T^k \mathfrak{g})_0$, for $\mathfrak{g} = \mathfrak{X}(M)$. As a $\mathbb{K}$-module, this is simply the space of section of the *axes-bundle* $A^k M$ which, by definition, is a certain direct sum over M of $2^k - 1$ copies of TM. Thus $\exp_k$ will be a bijection from the space of sections of $A^k M$ to the space of sections of $T^k M$. But using a canonical diagonal imbedding $TM \to A^k M$, every vector field X gives rise to a section δX of $A^k M$, then via the exponential map to a section $\exp_k(\delta X)$ of $T^k M$ which finally corresponds to a diffeomorphism $\Phi : T^k M \to T^k M$. On the other hand, if X admits a flow $\mathrm{Fl}_t^X : M \to M$ in the classical sense (e.g., if X is a complete vector field on a finite-dimensional real manifold), then $T^k(\mathrm{Fl}_1^X) : T^k M \to T^k M$ also defines a diffeomorphism of $T^k M$. These two diffeomorphisms are closely related to each other (see Chapter 29). Summing up, we have constructed an algebraic substitute of the "missing exponential map" $\mathfrak{X}(M) \to \mathrm{Diff}(M)$.

0.16. *The exponential jet of a connection.* In this final part we come back to the problem left open in Part IV, namely to the definition of a canonical sequence of permutation- and shift-invariant multilinear connections (Section 0.7).

In order to motivate our approach, assume M is a real finite-dimensional (or Banach) manifold equipped with an affine connection on the tangent bundle TM. Then, for every $x \in M$, the exponential map $\mathrm{Exp}_x : U_x \to M$ of the connection is a local diffeomorphism from an open neighborhood of the origin in $T_x M$ onto its image in M. Applying the k-th order tangent functor, we get $T^k \mathrm{Exp}_x : T^k(U_x) \to T^k M$, inducing bijections of fibers over 0_x and x,

$$(T^k \mathrm{Exp}_x)_{0_x} : \quad (A^k M)_x := (T^k(T_x M))_{0_x} \to (T^k M)_x. \tag{0.8}$$

As above, $(A^k M)_x$ is a certain direct sum of copies of $T_x M$, the fiber of the (k-th order) axes-bundle of M. The bijections of fibers fit together to a smooth bundle isomorphism $T^k \mathrm{Exp} : A^k M \to T^k M$, which we call the *exponential jet of the connection.* In fact, this isomorphism is a multilinear connection in the sense explained in Section 0.6, and has the desired properties of *shift invariance* and *total symmetry under the permutation group.* Therefore it is of basic interest to construct the exponential jet $\mathrm{Exp}^k : A^k M \to T^k M$ in a "synthetic" way, starting with a connection on TM and without ever using the map Exp_x itself, which doesn't exist in our general framework. In Chapters 30 and 31, we describe such a construction for Lie groups and symmetric spaces, and then outline its definition for a general (torsionfree) connection. Finally, in Chapter 32 we explain the problem of integrating jets to *germs*: given an *infinity jet* (a projective limit of k-jets, such as $\mathrm{Exp}^\infty := \lim_\leftarrow \mathrm{Exp}^k$), one would like to define a corresponding germ of a map, such as, for example, the germ of a "true" exponential map. But this is of course a problem belonging to integral geometry and not to differential geometry in our sense.

VII. Multilinear geometry

The appendix on "Multilinear geometry" gathers all purely algebraic results that are used in the main text. Two topics are treated:

(1) We introduce the concept of a *multilinear space* (over a commutative base ring $\mathbb{K}$); this is the algebraic model for the fibers of higher order tangent bundles $T^k M$ over the base M or, slightly more general, for the fibers of $T^{k-1}F$, where F is a vector bundle over M.

(2) We introduce *polynomial groups*; these are essentially formal groups (in the power-series approach, cf. [Haz78]) whose power series are finite, i.e. polynomial.

The link between (1) and (2) comes from the fact that the transition functions of bundles like $T^k F$ over M are not linear but *multilinear* (Chapter 15), and they belong to a group which can be defined in a purely algebraic way, called the *general multilinear group*. Now, it turns out that this group is a polynomial group, and hence we can apply Lie theory in a purely algebraic set-up. Conversely, much of the structure of a multilinear space is encoded in the general multilinear group. We now describe the concepts (1) and (2) in some more detail.

0.17. *Multilinear spaces.* The ingredients of the purely algebraic setting are: a family $(V_\alpha)_{\alpha \in I_k}$ of $\mathbb{K}$-modules (called a *cube of* $\mathbb{K}$*-modules*), where I_k is the lattice of all subsets of the finite set $\mathbb{N}_k = \{1, \ldots, k\}$ $(k \in \mathbb{N})$, the *total space* $E := \oplus_{\alpha \in I_k \setminus \emptyset} V_\alpha$ on which the *special multilinear group* $\mathrm{Gm}^{1,k}(E)$ acts in a non-linear way (essentially, by multilinear maps in the components V_α of E). We say that a structure on E is *intrinsic* if it is invariant under the action of the general multilinear group. There is no intrinsic linear structure on E, but the collection of *all* linear structures obtained by pushing around the initial linear structure on $\oplus_\alpha V_\alpha$ is an intrinsic object. Thus a multilinear space is not a $\mathbb{K}$-module but rather a collection of many different $\mathbb{K}$-module structures on a given set – this point of view has already turned out to be very useful in relation with Jordan theory and geometry ([Be02] and [Be04], which somewhat was at the origin of the approach developed here). In case $k = 2$, the theory is particularly simple, and therefore we first give an independent treatment of *bilinear spaces* (Chapter BA): in this case, one can easily give explicit formulae for the various $\mathbb{K}$-module structures (Section BA.1), thanks to the fact that in this case (and only in this case) the special linear group is abelian. The algebraically trained reader may start directly by reading Chapter MA. In the case of general $k \geq 2$, one should think of $\mathrm{Gm}^{1,k}(E)$ as a nilpotent Lie group (cf. Chapter PG).

We say that a multilinear space is *of tangent type* if all $\mathbb{K}$-modules V_α are canonically isomorphic to some fixed space V and hence are canonically isomorphic among each other – the terminology is motivated by the fact that the fibers of higher order tangent bundle $T^k M$ have this property. Then the symmetric group Σ_k acts on the total space E, where the action is induced by the canonical action of permutations on the set $\mathbb{N}_k$. This action is "horizontal" in the sense that it preserves the cardinality of subsets of $\mathbb{N}_k$. There is also another class of symmetry operators acting rather "diagonally" on the lattice of subsets of $\mathbb{N}_k$ and called *shift operators*. Invariance properties under these two kinds of operators and their interplay are investigated in Chapter SA.

Let us add some words on the relation of our concepts with *tensor products*. By their universal property, tensor products transform multilinear algebra into linear algebra. In a way, our concept has the same purpose, but, whereas the tensor product produces a new underlying set in order to reduce everything to *one* linear structure, we work with *one* underlying set and produce a new supply of linear structures. Thus, in principle, it is possible to reduce our set-up into the linear algebra of one linear structure by introducing some big tensor product space in a functorial way such that the action of the general multilinear group becomes linear – we made some effort to provide the necessary formulas by introducing a suitable "matrix calculus" for $\mathrm{Gm}^{1,k}(E)$. However, the definition of the "big" tensor space is combinatorially rather complicated and destroys the "geometric flavor" of the non-linear picture; instead of simplifying the theory, it causes unnecessary problems. This becomes clearly visible if the reader compares our non-linear picture with the concept of an *osculating bundle of a vector bundle* introduced by F.W. Pohl in [P62]: with great technical effort, he constructs a *linear* bundle which corresponds to $T^k F$ in the same way as the big tensor space corresponds to a multilinear space, and finally admits (loc. cit. p. 180): "In spite of several attempts, we have not been able to reduce even the [construction of the first osculating bundle] to something familiar."

Another advantage of our concept is that it behaves nicely with respect to topology: since we do not change the underlying set E, we have no problem in defining a suitable topology on E if $\mathbb{K}$ is a topological ring and the V_α are topological $\mathbb{K}$-modules. For tensor products, the situation is much worse: in general, topological tensor products of topological $\mathbb{K}$-modules have very bad properties which make them practically useless (general topological tensor products are not even associative, even over $\mathbb{K} = \mathbb{R}$, see [Gl04b]).

0.18. *Polynomial groups.* The general multilinear group $\mathrm{Gm}^{1,k}(E)$ is a special instance of a *polynomial group*: it is at the same a group and a $\mathbb{K}$-module such that the group multiplication is polynomial, and there is a common bound on the degree of all iterated product maps. One should think of a polynomial group as a nilpotent Lie group together with some global chart in which the group operations are polynomial. To such groups we may apply all results on *formal (power series) groups* (see, e.g., [D73], [Haz78] or [Se65]), but since we are in a polynomial setting, everything converges globally, and thus our results have a direct interpretation in terms of mappings. In particular, we can associate a Lie algebra to a polynomial group and define (of $\mathbb{K}$ is a field of characteristic zero) an exponential map and a logarithm (Theorem PG.6). Our proof of this result is much more elementary and closer to "usual" analysis on Lie groups than the proofs given for general formal groups in [Bou72] or [Se65] – we work with *one-parameter subgroups* essentially as in the differentiable case and do not need to invoke the Campbell-Hausdorff group chunk nor the universal enveloping algebra. Viewing formal power series groups as projective limits of polynomial groups, our result also implies corresponding results for general formal groups.

Limitations, Generalizations

0.19. *Limitations.* The purpose of the short Appendix L is to illustrate our distinction between "differential" and "integral" geometry by a list of results and notions that can *not* be treated in a purely differential context. On the other hand, a good deal of these items can be generalized from the finite-dimensional real set-up to surprisingly general situations (i.e., for rather big classes of topological base fields and of topological vector spaces) – many of these very recent results are due to H. Glöckner (see references).

0.20. *Generalizations.* Our methods apply to more general situations than just to manifolds over topological fields and rings. Some of these possibilities of generalization are discussed in Appendix G.

Related literature

It is impossible to take account of all relevant literature on such a vast area as Differential Geometry and Lie Theory, and I apologize for possibly omitting to give reference where it would have been due.

Real, finite and infinite dimensional differential geometry. For geometry on manifolds modelled on Banach-spaces, [La99] is my basic reference (the treatment of the second order tangent bundle and the Dombrowski Splitting Theorem 10.5 are motivated by the corresponding section of [La99]). From work of K.-H. Neeb (cf. [Ne02b] and [Ne06]) I learned that many notions of infinite-dimensional differential geometry from [La99] can be generalized to manifolds modelled on general locally convex vector spaces which need not even be complete. Finally, although it treats only the finite-dimensional case, the monograph "Natural Operations in Differential Geometry" by I. Kolář, P. Michor and J. Slovák [KMS93] turned out to be most closely related to the spirit of the present work. The reason for this is that an operation, if it is really "natural", is robust and vital enough to survive in most general contexts. In particular, Chapter VIII on "Product Preserving Functors" of [KMS93] had an important influence on my work. That chapter, in turn, goes back to the very influential paper "Théorie des points proches sur les variétés différentiables" by A. Weil [W53].

"Synthetic differential geometry" and "smooth toposes". These theories spring, in one way or another, from Weil's above mentioned paper. They propose a unification of differential geometry and algebraic geometry in very general categories which contain objects like $\mathrm{Diff}(M)$ and permit to use dual numbers as a means to modelize "tangent" objects (see [Ko81], [Lav87] or [MR91]). A similar approach is developed by the russian school of A. M. Vinogradov (cf. [ALV91], [Nes03]). It would be interesting to investigate the relations between those approaches and the one presented here. It seems that these theories are "dual" to ours in the sense explained in Section 0.8: whereas we base our approach on the covariant tangent functor T, those approaches are based on functions and thus are contravariant in

nature. In principle, it seems conceivable to represent (in the sense of Section 0.8) our geometric objects in "smooth toposes over $\mathbb{K}$" (cf. Appendix G).

Formal groups and other "formal" theories. Our version of Lie theory has much in common with the theory of *formal (power series) groups* (see [Haz78]), and our version of differential geometry is related to "formal" concepts such as *formal geometry* (see [CFT01]). As above, this may also be considered from a "contravariant" point of view, and then corresponds to the approach to formal groups via Hopf algebras (see [D73]). Again, it should be interesting to interprete the second approach as a "representation" of the first one.

Leitfaden

The chapters on Lie theory can be read without having to go through all the differential geometric parts, and conversely. The reader who is mainly interested in Lie groups may read Chapters 5 – 9 – 23 – 15 – 24 – PG – 25, and the reader interested in symmetric spaces could read Chapters 5 – 9 – 26 – 15 – 27 – 28 – 29 – 30. For such a reading, the interpretation of the tangent functor as a functor of scalar extension is not indispensable, provided one gives a proof of Theorem 7.5 not making use of dual numbers (which is possible).

Notation

Throughout the text, $\mathbb{K}$ denotes a commutative ring with unit 1. In certain contexts (e.g., symmetric spaces) we will assume that 2 is invertible in $\mathbb{K}$ or that $\mathbb{K}$ is a field (Section PG) or even that $\mathbb{K}$ a field of characteristic zero (when talking about exponential mappings). Some more specific notation concerning the following topics is explained in the corresponding subsections:

– differential calculus: **1.3**
– dual numbers, iterated dual numbers, jet rings: **6.1**, **7.3**, **7.4**, **8.1**
– partitions (set theory): **7.4**, **MA.1**, **MA.3**, **MA.4**
– bilinear maps, bilinear groups: **BA.1**, **BA.2**
– multilinear maps, multilinear groups: **MA.5**

Remark (added in november 2007). During the last two years, I have continued to think about some of the topics mentioned above, and concerning two points I slightly changed my mind: first, *principal bundles* and general Ehresmann connections are certainly an important topic that has to be studied in the present context [work in progress], and second, in case of *positive characteristic*, the present approach to the jet and tangent functors still leads to too much loss of information. This loss can be avoided by better exploiting the methods of proof of the general Taylor theorem (*cf.* Theorem 1.11), which means that one has to use more seriously the higher order slopes $f^{[k]}$, and not only its partial maps $T^k f$. This really is a new topic – see [Be08] for some tentative remarks.

I. Basic notions

This part contains the "conventional" aspects of the theory: having recalled the basic facts on differential calculus, we define manifolds and bundles in the usual way by charts and atlasses and we define the Lie bracket of vector fields in the old-fashioned way via its chart representation (the reader who prefers a more sophisticated definition is referred to Chapter 14). This is used in order to define, in Chapter 5, the *Lie functor* assigning a Lie algebra to a Lie group and a Lie triple system to a symmetric space.

1. Differential calculus

1.1. *Topological rings and modules.* By *topological ring* we mean a ring $\mathbb{K}$ with unit 1 together with a topology (assumed to be Hausdorff) such that the ring operations $+ : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ and $\cdot : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$ are continuous and such that the set $\mathbb{K}^\times$ of invertible elements is open in $\mathbb{K}$ and the inversion map $i : \mathbb{K}^\times \to \mathbb{K}$ is continuous. If $\mathbb{K}$ is an addition a field, it will be called a *topological field*. In the sequel we will always assume that $\mathbb{K}$ is a topological ring such that the set $\mathbb{K}^\times$ of invertible elements is dense in $\mathbb{K}$. (If $\mathbb{K}$ is a field, this means that $\mathbb{K}$ is not discrete.)

A module V over a topological ring $\mathbb{K}$ is called a *topological* $\mathbb{K}$*-module* if the structure maps $V \times V \to V$ and $\mathbb{K} \times V \to V$ are continuous. We will assume that all our topological modules are Hausdorff.

1.2. *The class* C^0 *and the "Determination Principle".* For topological $\mathbb{K}$-modules V, W and an open set $U \subset V$, we denote by $C^0(U,W) := C(U,W)$ the space of continuous maps $f : U \to W$. Since $\mathbb{K}^\times$ is assumed to be dense in $\mathbb{K}$, the value at zero of a continuous map $f : I \to W$, defined on an open neighborhood I of zero in $\mathbb{K}$, is already determined by the values of f at invertible elements, i.e., on $I \cap \mathbb{K}^\times$. We call this the *determination principle*.

1.3. *The class* C^1 *and the differential.* We say that $f : U \to W$ *is* $C^1(U,W)$ or just *of class* C^1 if there exists a C^0-map

$$f^{[1]} : U \times V \times \mathbb{K} \supset U^{[1]} := \{(x,v,t) |\, x \in U, x + tv \in U\} \to W,$$

such that

$$f(x + tv) - f(x) = t \cdot f^{[1]}(x,v,t) \tag{1.1}$$

whenever $(x,v,t) \in U^{[1]}$. (It is an easy exercise in calculus that, for $\mathbb{K} = \mathbb{R}$, $V = \mathbb{R}^n$ and $W = \mathbb{R}^m$ this coincides with the usual definition of the class C^1; see [Be07] for a very elementary discussion of these topics, and [BGN04] for more general results, including the infinite-dimensional situation.) The *differential of* f *at* x is defined by

$$df(x) : V \to W, \quad v \mapsto df(x)v := f^{[1]}(x,v,0).$$

Note that, by the Determination Principle, the map $f^{[1]}$ is uniquely determined by f and hence $df(x)$ is well-defined. The differential

$$df : U \times V \to W, \quad (x,v) \mapsto df(x)v = f^{[1]}(x,v,0);$$

is of class C^0 since so is $f^{[1]}$, and for all $v \in V$, the *directional derivative in direction* v

$$\partial_v f : U \to W, \quad x \mapsto \partial_v f(x) := df(x)v \tag{1.2}$$

is also of class C^0. We define a C^0-map Tf, called the *tangent map*, by

$$Tf : TU = U \times V \to W \times W, \quad (x,v) \mapsto (f(x), df(x)v) \tag{1.3}$$

and a C^0-map, called the *extended tangent map*, by

$$\widehat{T}f : U^{[1]} \to W \times W \times \mathbb{K}, \quad (x,v,t) \mapsto (f(x), f^{[1]}(x,v,t), t). \tag{1.4}$$

The following first order differentiation rules are easily proved.

Lemma 1.4. *For all $x \in U$, $df(x) : V \to W$ is a $\mathbb{K}$-linear C^0-map.*

Proof. We have already remarked that df and hence also $df(x)$ are C^0. Let us prove additivity of $df(x)$. For $s \in \mathbb{K}$ sufficiently close to 0, we have

$$\begin{aligned} sf^{[1]}(x, v+w, s) &= f(x + s(v+w)) - f(x) \\ &= f(x+sv+sw) - f(x+sv) + f(x+sv) - f(x) \\ &= sf^{[1]}(x+sv, w, s) + sf^{[1]}(x, v, s). \end{aligned}$$

For $s \in \mathbb{K}^\times$, we divide by s and get

$$f^{[1]}(x, v+w, s) = f^{[1]}(x+sv, w, s) + f^{[1]}(x, v, s).$$

By the Determination Principle, this equality holds for all s in a neighborhood of 0, hence in particular for $s = 0$, whence $df(x)(v+w) = df(x)w + df(x)v$. Homogeneity of $df(x)$ is proved in the same way. ■

Lemma 1.5. *If f and g are composable and of class C^1, then $g \circ f$ is C^1, and $T(g \circ f) = Tg \circ Tf$.*

Proof. Let $(x,y,t) \in U^{[1]}$. Then $f(x+ty) = f(x) + tf^{[1]}(x,y,t)$, hence

$$g(f(x+ty)) - g(f(x)) = g(f(x) + tf^{[1]}(x,y,t)) - g(f(x)) = t \cdot g^{[1]}(f(x), f^{[1]}(x,y,t), t),$$

where $(x,y,t) \to g^{[1]}(f(x), f^{[1]}(x,y,t), t)$ is C^0 since it is a composition of C^0-maps. Thus $g \circ f$ is C^1, with $(g \circ f)^{[1]}$ given by

$$(g \circ f)^{[1]}(x,y,t) = g^{[1]}(f(x), f^{[1]}(x,y,t), t)\,.$$

In particular, $d(g \circ f)(x) = dg(f(x)) \circ df(x)$ for all $x \in U$, implying our claim. ■

Lemma 1.6.

(i) *Multilinear maps of class* C^0 *are* C^1 *and are differentiated as usual. In particular, if* $f, g : U \to \mathbb{K}$ *are* C^1*, then the product* $f \cdot g$ *is* C^1*, and* $\partial_v(fg) = (\partial_v f)g + f\partial_v g$*. Polynomial maps* $\mathbb{K}^n \to \mathbb{K}^m$ *are always* C^1 *and are differentiated as usual.*

(ii) *Inversion* $i : \mathbb{K}^\times \to \mathbb{K}$ *is* C^1*, and* $di(x)v = -x^{-2}v$*. It follows that rational maps* $\mathbb{K}^n \supset U \to \mathbb{K}^m$ *are always* C^1 *and are differentiated as usual.*

(iii) *The direct product of two* C^1*-maps is* C^1*, and*

$$T(f \times g) = Tf \times Tg,$$

where we use the convention to identify, for a map $f : V \times W \to Z$*, the tangent map* $Tf : V \times W \times V \times W \to Z \times Z$ *with the corresponding map* $V \times V \times W \times W \to Z \times Z$ *(and similarly for maps defined on open subsets of* $V \times W$*). Using this convention, the tangent map of a diagonal map* $\delta_V : V \to V \times V$*,* $x \mapsto (x, x)$ *is the diagonal map* $\delta_{V \times V}$*.*

(iv) *If* $f : V_1 \times V_2 \supset U \to W$ *is* C^1*, then the rule on partial derivatives holds:*

$$df(x_1, x_2)(v_1, v_2) = d_1 f(x_1, x_2)v_1 + d_2 f(x_1, x_2)v_2.$$

Proof. One uses the same arguments as in usual differential calculus (cf. [BGN04, Section 2]). ∎

The rule on partial derivatives may also be written, by introducing for $(x_1, x_2) \in U$, the "left" and "right translations",

$$l_{x_1}(x_2) := r_{x_2}(x_1) := f(x_1, x_2).$$

Then

$$df(x_1, x_2)(v_1, v_2) = d(r_{x_2})(x_1)v_1 + d(l_{x_1})(x_2)v_2.$$

The converse of the rule on partial derivatives holds if $\mathbb{K}$ is a field (but we won't need it): if f is C^1 withe respect to the first and with respect to the second variable (in a suitable sense), then f is of class C^1 (see [BGN04, Lemma 3.9]).

1.7. *The classes* C^k *and* C^∞. Let $f : V \supset U \to F$ be of class C^1. We say that f *is* $C^2(U, F)$ or *of class* C^2 if $f^{[1]}$ is C^1, in which case we define $f^{[2]} := (f^{[1]})^{[1]} : U^{[2]} \to F$, where $U^{[2]} := (U^{[1]})^{[1]}$. Inductively, we say that f is $C^{k+1}(U, F)$ or *of class* C^{k+1} if f is of class C^k and $f^{[k]} : U^{[k]} \to F$ is of class C^1, in which case we define $f^{[k+1]} := (f^{[k]})^{[1]} : U^{[k+1]} \to F$ with $U^{[k+1]} := (U^{[k]})^{[1]}$. The map f is called *smooth* or *of class* C^∞ if it is of class C^k for each $k \in \mathbb{N}_0$. Note that $U^{[k+1]} = (U^{[1]})^{[k]}$ for each $k \in \mathbb{N}_0$, and that f is of class C^{k+1} if and only if f is of class C^1 and $f^{[1]}$ is of class C^k; in this case, $f^{[k+1]} = (f^{[1]})^{[k]}$. Next we prove the basic higher order differentiation rules.

Lemma 1.8. *If* f *and* g *are composable and of class* C^k*, then* $g \circ f$ *is of class* C^k*, and the generalized chain rule*

$$\widehat{T}^k(g \circ f) = \widehat{T}^k g \circ \widehat{T}^k f$$

holds.

Proof. For $k = 1$, this is Lemma 1.5, and the general case is proved in the same spirit by induction on k (cf. [BGN04, Prop. 4.5]). ∎

Lemma 1.9. *If f is of class C^2, then for all $x \in U$, $v, w \in V$,*

$$\partial_v \partial_w f(x) = \partial_w \partial_v f(x).$$

Proof. First of all, note that

$$\begin{aligned}\partial_w \partial_v f(x) &= \lim_{t \to 0} \frac{\partial_v f(x+tw) - \partial_v f(x)}{t} \\ &= \lim_{t \to 0} \Big(\lim_{s \to 0} \frac{f(x+tw+sv) - f(x+tw) - f(x+sv) + f(x)}{ts} \Big)\end{aligned}$$

where, for $t, s \in \mathbb{K}^\times$,

$$\begin{aligned}&\frac{f(x+sv+tw) - f(x+tw) - f(x+sv) + f(x)}{ts} \\ &\quad = \frac{1}{t}\Big(f^{[1]}(x+tw, v, s) - f^{[1]}(x, v, s) \Big) = f^{[2]}\big((x, v, s), (w, 0, 0), t\big),\end{aligned}$$

whence, by continuity of $f^{[2]}$, $\partial_w \partial_v f(x) = f^{[2]}\big((x, v, 0), (w, 0, 0), 0\big)$. Of course, we have also, by the same calculation as above, for $t, s \in \mathbb{K}^\times$,

$$\begin{aligned}&\frac{f(x+sv+tw) - f(x+sv) - f(x+tw) + f(x)}{ts} \\ &\quad = \frac{1}{s}\Big(f^{[1]}(x+sv, w, t) - f^{[1]}(x, w, t) \Big) = f^{[2]}\big((x, w, t), (v, 0, 0), s\big),\end{aligned}$$

and hence

$$f^{[2]}\big((x, v, s), (w, 0, 0), t\big) = f^{[2]}\big((x, w, t), (v, 0, 0), s\big).$$

By continuity of $f^{[2]}$, this equality holds also for $t = s = 0$, and this means that $\partial_w \partial_v f(x) = \partial_v \partial_w f(x)$. ■

Corollary 1.10. *If f is of class C^k and $x \in U$, then the map*

$$d^k f(x) : V^k \to W, \quad (v_1, \ldots, v_k) \mapsto \partial_{v_1} \ldots \partial_{v_k} f(x)$$

is a symmetric multilinear C^0-map.

Proof. The maps d^k are partial maps of $f^{[k]}$ and hence are C^0; for instance, we have seen above that $d^2 f(x)(v, w) = f^{[2]}((x, w, 0), (v, 0, 0), 0)$. Symmetry follows from Lemma 1.9, and multilinearity now follows from Lemma 1.4. ■

Theorem 1.11. (The second order Taylor expansion.) *Assume $f : U \to W$ is of class C^2. Then for all $(x, h, t) \in U^{[1]}$, f has an expansion*

$$f(x+th) = f(x) + t\, df(x)h + t^2\, a_2(x, h) + t^2\, R_2(x, h, t), \tag{1.5}$$

where the remainder term $R_2(x, h, t)$ is $O(t)$, i.e., it is of class C^0 and takes the value 0 for $t = 0$. The coefficient $a_2(x, h)$ is of class C^0 jointly in both variables and, in the variable h, is a vector-valued homogeneous form of degree 2, i.e., $a_2(x, \cdot)$

is homogeneous of degree 2*, and the map* $(v,w)\mapsto a_2(x,v+w)-a_2(x,v)-a_2(x,w)$ *is bilinear. More precisely, we have*

$$a_2(x,v+w)-a_2(x,v)-a_2(x,w)=d^2f(x)(v,w). \tag{1.6}$$

In particular $2\,a_2(x,h)=d^2f(x)(h,h)$*. Therefore, if* 2 *is invertible in* $\mathbb{K}$*, the second order Taylor expansion* (1.5) *may also be written*

$$f(x+th)=f(x)+t\,df(x)h+\frac{t^2}{2}\,d^2f(x)(h,h)+t^2\,R_2(x,h,t). \tag{1.7}$$

Proof. Since f is of class C^2, applying the fundamental relation (1.1) first to f and then to $f^{[1]}$, we get

$$\begin{aligned} f(x+th) &= f(x)+tf^{[1]}(x,h,t)\\ &= f(x)+tdf(x)h+t(f^{[1]}(x,h,t)-f^{[1]}(x,h,0))\\ &= f(x)+tdf(x)h+t^2f^{[2]}((x,h,0),(0,0,1),t)\\ &= f(x)+tdf(x)h+t^2f^{[2]}((x,h,0),(0,0,1),0)\\ &\qquad +t^2\left(f^{[2]}((x,h,0),(0,0,1),t)-f^{[2]}((x,h,0),(0,0,1),0)\right). \end{aligned}$$

We let $a_2(x,h):=f^{[2]}((x,h,0),(0,0,1),0)$ and

$$R_2(x,h,t):=f^{[2]}((x,h,0),(0,0,1),t)-f^{[2]}((x,h,0),(0,0,1),0).$$

These maps are all C^0 since so is $f^{[2]}$, and $R_2(x,h,t)$ takes the value 0 at $t=0$. One shows, as in usual differential calculus, that an expansion with these properties ("développement limité" in French) is unique (cf. [BGN04, Lemma 5.2]), and then deduces that $a_2(x,\cdot)$ is homogeneous of degree 2. Let us prove (1.6). From (1.5), together with linearity of $df(x)$, we get

$$\begin{aligned} &f(x+t(h_1+h_2))-f(x+th_1)-f(x+th_2)+f(x)\\ &\quad= t^2(a_2(x,h_1+h_2)-a_2(x,h_1)-a_2(x,h_2))+t^2O(t). \end{aligned}$$

Thus, for $t\in\mathbb{K}^\times$,

$$\begin{aligned} &\frac{f(x+t(h_1+h_2))-f(x+th_1)-f(x+th_2)+f(x)}{t^2}\\ &\qquad= a_2(x,h_1+h_2)-a_2(x,h_1)-a_2(x,h_2)+O(t). \end{aligned}$$

For $t=0$, the C^0-extension of the left hand side equals $d^2f(x)(h_1,h_2)$ (cf. proof of Lemma 1.9), and this implies (1.6). Finally, since we already know that $a_2(x,h)$ is homogeneous quadratic in h, we get for $h_1=h_2=h$ the relation $2\,a_2(x,h)=d^2f(x)(h,h)$, which implies (1.7). ∎

Corollary 1.12. *If U is an open neighborhood of the origin in V and $f : U \to W$ is of class C^2 and homogeneous of degree 2 (i.e., $f(rx) = r^2 f(x)$ whenever this makes sense), then f is a W-valued quadratic form.*

Proof. Clearly, $f(0) = 0$. Using the expansion (1.5) at $x = 0$, we get

$$t^2 f(h) = f(th) = tdf(0)h + t^2(a_2(0,h) + R_2(0,h,t)).$$

Dividing, for $t \in \mathbb{K}^\times$, by t, we get an identity whose both sides are C^0-functions of t. Letting $t = 0$, we deduce that $df(0)h = 0$. Dividing again, for $t \in \mathbb{K}^\times$, by t^2, we get $f(h) = a_2(0,h) + R_2(x,h,t)$. Both sides are C^0-functions of t, and hence for $t = 0$ we get $f(h) = a_2(0,h)$. According to (1.6), f is then a vector-valued quadratic form. ∎

It is fairly obvious that the arguments given in the proof of Theorem 1.11 can be iterated in order to prove the general expansions, for f of class C^k,

$$\begin{aligned} f(x+th) &= f(x) + \sum_{j=1}^{k} t^j a_j(x,h) + t^k R_k(x,h,t), \\ &= f(x) + \sum_{j=1}^{k} \frac{t^j}{j!} d^j f(x)(h,\ldots,h) + t^k R_k(x,h,t) \end{aligned}$$

the latter provided the integers are invertible in $\mathbb{K}$ – see [BGN04, Chapter 5] for the details. In this work we will use only the second order Taylor expansion. Thus the present summary of basic results on differential calculus is sufficient for our purposes. For more information, we refer to [BGN04]; see also Appendix L, Section L.3, for examples of some important categories of topological fields and vector spaces and their main properties, and Appendix G, Section G.1, for the generalization of the preceding set-up to general "C^0-concepts", as introduced in [BGN04].

2. Manifolds

2.1. *Manifolds and manifolds with atlas.* We fix a topological $\mathbb{K}$-module V as "model space" of our manifold. A C^k*-manifold with atlas* is a topological space M together with a V*-atlas* $\mathcal{A} = (\varphi_i, U_i)_{i\in I}$. This means that U_i, $i \in I$ is a covering of M by open sets, and $\varphi_i : M \supset U_i \to \varphi_i(U_i) \subset V$ is a *chart*, i.e. a homeomorphisms of the open set $U_i \subset M$ onto an open set $\varphi_i(U_i) \subset V$, and any two charts $(\varphi_i, U_i), (\varphi_j, U_j)$ are C^k*-compatible* in the sense that

$$\varphi_{ij} := \varphi_i \circ \varphi_j^{-1}|_{\varphi_j(U_i \cap U_j)} : \varphi_j(U_i \cap U_j) \to \varphi_i(U_i \cap U_j)$$

and its inverse φ_{ji} are of class C^k. If M is C^k for all k, we say that M is a smooth or C^∞-manifold. We see no reason to assume that the topology of M is Hausdorff (compare [La99, p. 23] for this issue). Also, we see no reason to assume that the atlas $\mathcal{A}$ is *maximal*. Of course, any atlas can be completed to a maximal one, if one wishes to do so. (See, however, Section 2.4 below, for a word of warning concerning maximal atlasses.) A *manifold over* $\mathbb{K}$ is a manifold together with a maximal atlas.

Sometimes it may be useful to consider also the disjoint union of two manifolds (possibly modelled on non-ismorphic topological modules) as a manifold; then one should call V*-manifolds* the class of manifolds defined before.

Let M, N be C^k-manifolds with atlas. A map $f : M \to N$ is *of class* C^k if, for all choices of charts (φ, U) of M and (ψ, W) of N,

$$\psi \circ f \circ \varphi^{-1} : \varphi(U \cap f^{-1}(W)) \to \psi(W)$$

is of class C^k. From the chain rule (Lemma 1.8) it follows that C^k-manifolds with atlas form a category. We denote by $\mathrm{Diff}(M)$ or $\mathrm{Diff}_{\mathbb{K}}(M)$ the automorphism group of M in this category, also called the group of *diffeomorphisms*.

In particular, all topological $\mathbb{K}$-modules are C^∞-manifolds and we can define *smooth functions on* M to be smooth maps $f : M \to \mathbb{K}$. The space $C^\infty(M)$ of smooth functions on M may be reduced to the constants, and it may also happen that $C^\infty(U_i)$ is reduced to the constants for all chart domains U_i (e.g. case of topological vector spaces that admit no non-zero continuous linear forms, cf. [BGN04, Example 8.2]). Therefore it is no longer possible to define differential geometric objects via their action on smooth functions.

2.2. *Direct products.* In the category of smooth manifolds over $\mathbb{K}$ one can form direct products: given two C^k-manifolds with atlas $(M, \mathcal{A}), (N, \mathcal{B})$, endow $M \times N$ with the product topology, and the charts are given by the maps $\varphi_i \times \psi_j$. These charts are again C^k-compatible and define an atlas $\mathcal{A} \times \mathcal{B}$.

2.3. *Submanifolds.* For our purposes, the following strong definition of submanifolds will be convenient (of course, one may consider also weaker conditions): if E is a submodule of a topological $\mathbb{K}$-module V, then we say that E is *admissible* if there exists a complementary submodule F such that the bijection

$$E \times F \to V, \quad (e, f) \mapsto e + f$$

is a homeomorphism. Thus E and F are closed and complemented submodules of V. Now a *submanifold* is defined as usual to be a subset $N \subset M$ such that there exists a subatlas of $\mathcal{A}$ having the property that $U_i \cap N$ is either empty or corresponds to an admissible submodule of V.

2.4. *Locality.* For classical real manifolds, any topological refinement of the open chart domain cover $(U_i)_{i \in I}$ of M defines a new and equivalent atlas of M. In particular, a maximal atlas of M contains, for any point $x \in M$, a system of open neighborhoods of x as possible chart domains. This fact remains true if $\mathbb{K}$ is a field, because of the following Locality Lemma ([BGN04, Lemma 4.9]): *If $\mathbb{K}$ is a field and if f is C^k on each open set U_j of an open cover $(U_j)_{j \in J}$ of the open set $U \subset V$, f is also of class C^k on U.* For general base rings, this may become false. In this case, one should rather think of a maximal atlas as something like a "maximal bundle atlas" of a fiber bundle, with fibers corresponding to ideals of non-invertible elements of $\mathbb{K}$. As long as one minds this remark, the theory of manifolds over rings is not really different from the theory of manifolds over fields.

3. Tangent bundle and general fiber bundles

3.1. *An equivalence relation describing M.* Assume M is a manifold with atlas modelled on V. A point $p \in M$ is described in the form $p = \varphi_i^{-1}(x)$ with $x \in \varphi_i(U_i)$ and $i \in I$. In a different chart it is given by $p = \varphi_j^{-1}(y)$. In other words, M is the set of equivalence classes $S/\sim$, where

$$S := \{(i,x) | \, x \in \varphi_i(U_i)\} \subset I \times V,$$

and $(i,x) \sim (j,y)$ iff $\varphi_i^{-1}(x) = \varphi_j^{-1}(y)$ iff $\varphi_{ji}(x) = y$. We write $p = [i,x] \in M = S/\sim$.

3.2. *An equivalence relation describing TM.* Next we define an equivalence relation on the set

$$TS := S \times V \subset I \times V \times V$$

by:

$$\begin{aligned} (i,x,v) \sim (j,y,w) \quad &:\Leftrightarrow \quad \varphi_j \circ \varphi_i^{-1}(x) = y, \; d(\varphi_j \circ \varphi_i^{-1})(x)v = w \\ &\Leftrightarrow \quad \varphi_{ji}(x) = y, \; d\varphi_{ji}(x)v = w. \end{aligned}$$

Since $\varphi_{ii} = \mathrm{id}$, $\varphi_{ij}\varphi_{jk} = \varphi_{ik}$, the chain rule implies that this is an equivalence relation. Again we denote equivalence classes by $[i,x,v]$, and we let

$$TM := TS/\sim .$$

If $[i,x,v] = [j,y,w]$, then $[i,x] = [j,y]$, and hence the map

$$\pi : TM \to M, \quad [i,x,v] \mapsto [i,x]$$

is well-defined. For $p = [i,x] \in M$, we let

$$T_pM := \pi^{-1}(p) = \{[i,x,v] \in TM | \, v \in V\}.$$

The map

$$T_x\varphi_i^{-1} : V \to T_pM, \quad v \mapsto [i,x,v]$$

is a bijection (it is surjective by definition and injective since the differentials are bijections), and we can use it to define the structure of a topological $\mathbb{K}$-module on T_pM which actually does not depend on (i,x) because $T_y\varphi_j^{-1} = T_x\varphi_i^{-1} \circ d\varphi_{ij}(y)$. The $\mathbb{K}$-module T_pM is called the *tangent space of M at x.*

We define an atlas $T\mathcal{A} := (T\varphi_i)_{i \in I}$ on TM by:

$$TU_i := \pi^{-1}(U_i), \qquad T\varphi_i : TU_i \to V \times V, \; [i,x,v] \mapsto (x,v)$$

Change of charts is now given by

$$T\varphi_{ij} : (x,v) \mapsto (\varphi_{ij}(x), d\varphi_{ij}(x)v)$$

which is C^{k-1} if φ_{ij} is C^k. Thus $(TM, T\mathcal{A})$ is a manifold with atlas.

If $f : M \to N$ is C^k we define its *tangent map* by

$$Tf : TM \to TN, \quad [i, x, v] \mapsto [j, f_{ij}(x), df_{ij}(x)v]$$

where $f_{ij} = \psi_j \circ f \circ \varphi_i^{-1}$ (supposed to be defined on a non-empty open set). In other words,

$$Tf = (T\psi_j)^{-1} \circ (f_{ij}, df_{ij}) \circ T\varphi_i = (T\psi_j)^{-1} \circ Tf_{ij} \circ T\varphi_i$$

with Tf_{ij} as defined in (1.3). This is well-defined, linear in fibers and C^{k-1}. Clearly the functorial rules hold, i.e. we have defined a covariant functor T from the category of smooth manifolds over $\mathbb{K}$ into itself.

Using the product atlas from 2.2 on the direct product $M \times N$, we see that there is a natural isomorphism

$$T(M \times N) \cong TM \times TN$$

which is directly constructed by using the equivalence relation defining TM.

If $f : M \to \mathbb{K}$ is a smooth function, then $T_x f : T_x M \to T_{f(x)}\mathbb{K} = \mathbb{K}$ gives rise to a function $TM \to \mathbb{K}$, linear in fibers, which we denote by df. The product rule (Lemma 1.6 (i)) implies that $d(fg) = f\,dg + g\,df$.

3.3. *General fiber bundles.* General fiber bundles over M are defined following the same pattern as above for the tangent bundle: assume M is modelled on V and let N some manifold modelled on a topological $\mathbb{K}$-module W. Assume that, for all triples $(i, j, x) \in I \times I \times V$ such that $\varphi_i^{-1}(x) \in \varphi_j^{-1}(U_j)$, an element

$$g_{ji}(x) \in \mathrm{Diff}(N)$$

of the diffeomorphism group of N is given such that the cocycle relations

$$g_{ik}(x) = g_{ij}(\varphi_{jk}(x)) \circ g_{jk}(x), \quad g_{ii}(x) = \mathrm{id}_N$$

are satisfied and such that

$$(x, w) \mapsto g_{ij}(x, w) := g_{ij}(x)w$$

is smooth wherever defined. Then we define an equivalence relation on $S \times N$ by

$$(i, x, v) \sim (j, y, w) \quad :\Leftrightarrow \quad \varphi_{ji}(x) = y,\ g_{ji}(x)v = w.$$

By the cocycle relations, this is indeed an equivalence relation, and by the smoothness assumption, $F := S \times N/\sim$ can be turned into a manifold modelled on $V \times W$ and locally isomorphic to $U_i \times N$ and such that the projection $p : F \to M$, $[i, x, w] \mapsto [i, x]$ is a well-defined smooth map whose fibers are all diffeomorphic to N. When we speak of a "chart" of a bundle, we always mean a bundle chart, and we never consider maximal atlasses on bundles.

Homomorphisms or *bundle maps* are pairs $\widetilde{f} : F \to F'$, $f : M \to M'$ of smooth maps such that $p' \circ \widetilde{f} = f \circ p$. Using charts, it is seen that fibers of bundles are submanifolds, and bundle maps induce by restriction smooth maps on fibers.

3.4. *Vector bundles.* If N carries an additional structure ($\mathbb{K}$-module, affine space, projective space...) and the $g_{ij}(x)$ respect this structure, then each fiber also carries this structure. In particular, if N is a $\mathbb{K}$-module and the transition maps respect this structure, then F is called a *vector bundle.* In this case, the bundle atlas is of the form $\widetilde{A} = (p^{-1}(U_i), \widetilde{\varphi}_i)_{i\in I}$ with

$$\widetilde{\varphi}_i : p^{-1}(U_i) \to V \times W, \quad [i,x,w] \mapsto (x,w).$$

Change of charts is given by the transition functions

$$\widetilde{g}_{ij}(x,w) = (\varphi_{ij}(x), g_{ij}(x)w).$$

Homomorphisms are required to respect the extra structure on the fibers. Thus homomorphisms of vector bundles are such that restriction to fibers induces $\mathbb{K}$-linear maps.

If F is a vector bundle over M, then there is a well defined map

$$z : M \to F, \quad [i,x] \mapsto [i,x,0],$$

called the *zero section*, and if $(\widetilde{f}, f)$ is a homomorphism of vector bundles, then f can be recovered from $\widetilde{f}$ via $f = p \circ \widetilde{f} \circ z$.

3.5. *Direct sum of vector bundles.* If F and E are fiber bundles over M, then, by definition, $F \times_M E$ is the fiber bundle over M whose fiber over x is $F_x \times E_x$. The transition functions are given by $g_{ij}(x) \times f_{ij}(x)$. If E and F are in addition vector bundles, then $E \times_M F$ is also a vector bundle, denoted by $E \oplus_M F$, with transition functions, in matrix form,

$$g_{ij}(x) \oplus f_{ij}(x) = \begin{pmatrix} g_{ij}(x) & 0 \\ 0 & f_{ij}(x) \end{pmatrix}.$$

We do not define tensor products, dual or hom-bundles of vector bundles in our general context – see Appendix L, Sections L1 and L2 for explanatory comments.

4. The Lie bracket of vector fields

4.1. *Vector fields and derivations.* A *section* of a vector bundle F over M is a smooth map $X : M \to F$ such that $p \circ X = \mathrm{id}_M$. If F is a vector bundle, then the sections of F form a module, denoted by $\Gamma^\infty(F)$ or simply by $\Gamma(F)$, over the ring $C^\infty(M)$. Sections of TM are also called *vector fields*, and we also use the classical notation $\mathfrak{X}(M)$ for $\Gamma(TM)$. In a chart (U_i, φ_i), vector fields can be identified with smooth maps $X_i : V \supset \varphi_i(U_i) \to V$, given by

$$X_i := \mathrm{pr}_2 \circ T\varphi_i \circ X \circ \varphi_i^{-1} : \varphi_i^{-1}(U_i) \to U_i \to TU_i \cong U_i \times V \to V.$$

Similarly, sections of an arbitrary vector bundle are locally represented by smooth maps

$$X_i : V \supset \varphi_i^{-1}(U_i) \to W.$$

If the chart (U_i, φ_i) is fixed, for brevity of notation we will often suppress the index i and write the chart representation of X in the form

$$U \to U \times W, \quad x \mapsto x + X(x) \quad \text{or} \quad (x, X(x)).$$

For a vector field $X : M \to TM$ and a smooth function $f : M \to \mathbb{K}$, recall that the differential $df : TM \to \mathbb{K}$ is smooth and hence we can define a smooth function $L_X f$ by $L_X f := df \circ X$. Then we have the Leibniz rule:

$$L_X(fg) = d(fg) \circ X = (fdg + gdf) \circ X = gL_X f + fL_X g.$$

Thus the map $X \mapsto L_X$ is a $\mathbb{K}$-linear map into the space of derivations of $C^\infty(M)$. However, it will in general neither be injective nor surjective, not even when restricted to suitable open subsets (cf. [BGN04, Examples 8.2 and 8.3]). Therefore it cannot be used to define the Lie algebra structure on $\mathfrak{X}(M)$. In the following theorem, we define the Lie bracket by using its expression in a chart; an intrinsic definition of the Lie bracket needs a closer look at the second tangent bundle TTM and is postponed to Chapter 14.

Theorem 4.2. *There is a unique structure of a Lie algebra over $\mathbb{K}$ on $\mathfrak{X}(M)$ such that for all $X, Y \in \mathfrak{X}(M)$ and $(i, x) \in S$,*

$$[X, Y]_i(x) = dX_i(x)Y_i(x) - dY_i(x)X_i(x).$$

Proof. Uniqueness is clear. Let us show that, on the intersection of two chart domains, the bracket $[X, Y]$ is independent of the choice of chart: assume $(i, x) \sim (j, y)$, i.e. $y = \varphi_j \varphi_i^{-1} x = \varphi_{ji}(x)$; then $X_j(y) = d\varphi_{ji}(x)X_i(x)$ or

$$X_j \circ \varphi_{ji} = T\varphi_{ji} \circ X_i. \tag{4.1}$$

We have to show that $[X, Y]$ has the same transformation property under changes of charts. For simplicity, we let $\varphi := \varphi_{ji}$ and calculate:

$$\begin{aligned} dY_j(\varphi(x))X_j(\varphi(x)) &= d(T\varphi \circ Y_i)(x)T\varphi \circ X_i(x) \\ &= d^2\varphi(x)(X_i(x), Y_i(x)) + d\varphi(x)dY_i(x)X_i(x). \end{aligned}$$

We exchange the rôles of X and Y and take the difference of the two equations thus obtained: we get, using Schwarz' lemma (Lemma 1.9),

$$dY_j(\varphi(x))X_j(\varphi(x)) - dX_j(\varphi(x))Y_j(\varphi(x)) = d\varphi(x)(dY_i(x)X_i(x) - dX_i(x)Y_i(x))$$

which had to be shown. Summing up, the bracket operation $\mathfrak{X}(M) \times \mathfrak{X}(M) \to \mathfrak{X}(M)$ is well-defined, and it clearly is $\mathbb{K}$-bilinear and antisymmetric in X and Y. For the case that 2 is not invertible in $\mathbb{K}$, note that we clearly also have the identity $[X, X] = 0$.

All that remains to be proved is the Jacobi identity. This is done by a direct computation which involves only the composition rule and Schwarz' lemma: define a (chart dependent) "product" of X_i and Y_i by

$$(X_i \cdot Y_i)(x) := dY_i(x)X_i(x).$$

(Later this will be interpreted as $\nabla_{X_i} Y_i$, the canonical flat connection of the chart U_i, see Chapter 10.) Then, by a direct calculation, one shows that this product is a *left symmetric* or *Vinberg algebra* (cf., e.g., [Koe69]), i.e. it satisfies the identity

$$X_i \cdot (Y_i \cdot Z_i) - (X_i \cdot Y_i) \cdot Z_i = Y_i \cdot (X_i \cdot Z_i) - (Y_i \cdot X_i) \cdot Z_i.$$

But it is immediately checked that for every left symmetric algebra, the commutator $[X_i, Y_i] = X_i \cdot Y_i - Y_i \cdot X_i$ satisfies the Jacobi identity. ■

The Lie bracket is natural in the following sense: assume $\varphi : M \to N$ is a smooth map and $X \in \mathfrak{X}(M)$, $Y \in \mathfrak{X}(N)$. We say that the pair (X, Y) is *φ-related* if

$$Y \circ \varphi = T\varphi \circ X.$$

Lemma 4.3. *If (X, Y) and (X', Y') are φ-related, then so is $([X, X'], [Y, Y'])$. In particular, the diffeomorphism group of M acts by automorphisms of the Lie algebra $\mathfrak{X}(M)$.*

Proof. This is the same calculation as the one given in the preceding proof after Eqn. (4.1). ■

Moreover, from the definitions it follows easily that

$$\mathfrak{X}(M) \to \mathrm{Der}(C^\infty(M)), \quad X \mapsto -L_X$$

is a homomorphism of Lie algebras. (The sign in the definition of the Lie bracket is a matter of convention.)

4.4. *Infinitesimal automorphisms.* If X is a vector field on M, then an *integral curve* is a smooth map $\varphi : \mathbb{K} \supset I \to M$ such that $T_t\varphi \cdot \mathbf{1} = X(\varphi(t))$. In a chart, this is equivalent to the usual differential equation $\varphi'(t) = X(\varphi(t))$. In our general set-up, we have no existence or uniqueness statements for solutions of ordinary differential equations, and hence we will never work with integral curves or flows.

However, a very rough infinitesimal version of the flow of a vector field can be defined (for more refined versions see Chapter 29): we say that a diffeomorphism

$f : TM \to TM$ is an *infinitesimal automorphism* if f preserves fibers, and in each fiber f acts by translations. Clearly, this defines a group, denoted by

$$\mathrm{InfAut}(M) := \{ f \in \mathrm{Diff}(TM) | \, \forall p \in M : \exists v_p \in T_pM : \forall v \in T_pM : \; f(v) = v + v_p \}.$$

Clearly, $X(p) := v_p = f(0_p)$ defines a map $X : M \to TM$ that can also be written $X = f \circ z$, where $z : M \to TM$ is the zero section, whence is a smooth section of TM, i.e., it is a vector field on M. Conversely, if $X : M \to TM$ is a vector field on M, we define a map

$$\widetilde{X} : TM \to TM, \quad v \mapsto v + X(\pi(v))).$$

In each fiber, $\widetilde{X}$ acts by translations, and using charts, we see that $\widetilde{X}$ is smooth. The inverse of $\widetilde{X}$ is $\widetilde{-X}$; more generally, we have $\widetilde{X+Y} = \widetilde{X} \circ \widetilde{Y}$, and hence $\widetilde{X}$ is an infinitesimal automorphism. Summing up, we have defined a group isomorphism $(\mathfrak{X}(M), +) \cong \mathrm{InfAut}(M)$. Note that the same notions could be defined by replacing TM by an arbitrary vector bundle F over M. In Chapter 28, "higher order versions" of the group $\mathrm{InfAut}(M)$ will be defined.

4.5. *Tensor fields, forms.* Let E and F be vector bundles over M. An *E-valued k-multilinear form on F* is a smooth map $A : F \times_M \ldots \times_M F \to E$ (k factors) such that $p_E \circ A = p_{\times^k_M F}$, i.e., A maps fibers over x to fibers over x, and $A_x : F_x \times \ldots \times F_x \to E_x$ is $\mathbb{K}$-linear in all k variables. If $F = TM$, then we speak also of *E-valued forms*; if $F = TM$ and $E = M \times \mathbb{K}$ is the trivial line bundle, then A is called a *tensor field of type $(k,0)$*, and if $F = E = TM$, then A is called *tensor field of type $(k,1)$*. *Differential forms of degree k* are skew-symmetric tensor fields of type $(k,0)$. (We don't define tensor fields of type (r,s) with $s > 1$ since we have not defined the tensor product bundles; cf. Appendix L2.) Note that we do not interprete forms as sections of a vector bundle over M; for instance, one-forms are defined as maps $\omega : TM \to \mathbb{K}$, but T^*M is not defined as a bundle (cf. Appendix L1).

4.6. *Almost (para-) complex and dual structures.* A tensor field J of type $(1,1)$, i.e., a smooth map $J : TM \to TM$ acting linearly on all fibers, such that $J^2 = -\mathbf{1}$ will be called an *almost complex structure* (even if the base field is not real), and if $J^2 = \mathbf{1}$, J will be called a *polarization* (or a *para-complex structure* in case 2 is invertible in $\mathbb{K}$ and both eigenspaces of J_x are isomorphic as submodules of T_xM for all $x \in M$). If $J^2 = 0$ and $\ker(J_x) = \mathrm{im}(J_x)$ for all $x \in M$, then J will be called an *almost dual structure*.

4.7. *The canonical almost dual structure on the tangent bundle.* The tangent bundle TM of a manifold M carries a natural almost dual structure. This can be seen as follows: recall that the transition maps of the bundle atlas of TM are given by $T\varphi_{ij}(x;v) = (\varphi_{ij}(x), d\varphi_{ij}(x)v)$. Therefore, using the rule on partial derivatives, the transition maps of the bundle atlas of $TTM = T(TM)$ are given by

$$\begin{aligned} TT\varphi_{ij}(x,v;x',v') &= (T\varphi_{ij}(x,v), d(T\varphi_{ij})(x,v)(x',v')) \\ &= (\varphi_{ij}(x), d\varphi_{ij}(x)v, d\varphi_{ij}(x)x', d\varphi_{ij}(x)v' + d^2\varphi_{ij}(x)(v,x')) \\ &= (\varphi_{ij}(x), d\varphi_{ij}(x)v, \begin{pmatrix} d\varphi_{ij}(x) & 0 \\ d^2\varphi_{ij}(x)(v,\cdot) & d\varphi_{ij}(x) \end{pmatrix} \begin{pmatrix} x' \\ v' \end{pmatrix}). \end{aligned}$$

The matrix in the last line describes the differential $d(T\varphi_{ij})(x,v)$. It is clear that this matrix commutes with the matrix

$$Q := \begin{pmatrix} 0 & 0 \\ \mathbf{1}_V & 0 \end{pmatrix}.$$

Therefore, if $p = [i,x] \in M$ and $u = [i,x,v] \in T_pM$, and we identify $T_u(TM)$ with $V \times V$ via the chart $TT\varphi_i$, then the endomorphism ε_u of $T_u(TM)$ defined by Q does not depend on the chart. Hence

$$\varepsilon := (\varepsilon_u)_{u \in TM}$$

defines a tensor field of type $(1,1)$ on TM such that $\varepsilon^2 = 0$ and the kernel of ε_u is exactly $T_pM \cong T_u(T_pM) \subset T_u(TM)$. This tensor field is natural in the sense that it is invariant under *all* diffeomorphisms of the kind Tf where f is a diffeomorphism of M: this is proved by the above calculation, with φ_{ij} replaced by f. (In [Bes78, p. 20/21] this tensor field, in a slightly different presentation, is called the "vertical endomorphism".)

The tensor field ε is *integrable* in the following sense: the distribution of subspaces given by $\ker \varepsilon$ admits *integral submanifolds* which are simply the fibers of the bundle TM. Comparing with the situation of an almost complex structure (replace the condition $\varepsilon^2 = 0$ by $\varepsilon^2 = -\mathbf{1}$), one might conjecture that in analogy with the theorem of Newlander and Nirenberg from the complex case, also in this case it may be true that indeed TM is already a manifold defined over the ring $\mathbb{K} \oplus \varepsilon\mathbb{K}$ with relation $\varepsilon^2 = 0$, i.e. over the dual numbers. In Chapter 6 we will show that this is indeed true.

4.8. *Remarks on differential operators of degree zero.* In the real finite-dimensional case, tensor fields can be interpreted as (multi-)differential operators of degree zero. This is in general no longer possible in our set-up. By definition, a *(multi-) differential operator of degree zero* between vector bundles E, F over M is given by a collection of maps from sections to sections, for all open sets $U \subset M$, or at least for all chart domains U,

$$\widetilde{A}_U : \Gamma(F|_U) \times \ldots \times \Gamma(F|_U) \to \Gamma(F|_U), \quad (\xi_1, \ldots, \xi_k) \mapsto \widetilde{A}_U(\xi_1, \ldots, \xi_k) \tag{4.2}$$

(k factors) such that

(1) $\widetilde{A}_U$ is $\mathcal{C}^\infty(U)$-multilinear,
(2) for $x \in U$, the value $\widetilde{A}_U(\xi_1, \ldots, \xi_k)(x)$ depends only on the values of the sections $\xi_1, \ldots, \xi_k$ at the point x.

If $A : F \times_M \ldots \times_M F \to E$ is a tensor field as in Section 4.5, then clearly we obtain a multidifferential operator of degree zero by letting

$$(\widetilde{A}_U(\xi_1, \ldots, \xi_k))(x) := A_x(\xi_1(x), \ldots, \xi_k(x)).$$

Conversely, given a collection of mappings $\widetilde{A}_U$ having Properties (1) and (2), we can define a map $A : F \times_M \ldots \times_M F \to E$, multilinear in fibers and depending smoothly on the base M, in the following way: let $x \in M$, choose a chart U around

x, extend elements $v_1, \ldots, v_k \in F_x$ to constant sections $\widetilde{v}_1, \ldots, \widetilde{v}_k$ of F over U (i.e. $(\widetilde{v}_i)(x) = v_i$) and let

$$A_x(v_1, \ldots, v_k) := (\widetilde{A}_U(\widetilde{v}_1, \ldots, \widetilde{v}_k))(x). \tag{4.3}$$

(This is well-defined: independence from U and from the extension of elements in the fiber to local sections follows from (2).) However, we cannot guarantee smooth dependence on $v_1, \ldots, v_k$; this has to be checked case by case in application situations. For instance, a one-form $A : TM \to \mathbb{K}$ corresponds to $\omega := \widetilde{A}_M : \mathfrak{X}(M) \to C^\infty(M)$, and for vector fields X, Y, we let

$$(L_X \omega)(Y) := L_X(\omega(Y)) - \omega([X, Y]). \tag{4.4}$$

Then $L_X \omega$ satisfies Properties (1) and (2); smooth dependence on $v = Y(x)$ is easily checked in a chart, and hence $L_X \omega$ defines, via (4.3), a tensor field which we denote by $L_X A$. Similarly we could define the Lie derivative of an arbitrary $(k, 0)$- or $(k, 1)$-form, and using this, the exterior derivative of a differential form could be defined in the usual way. We will, however, give a more intrinsic definition of the exterior derivative later on (Chapters 13 and 22). Finally, let us remark that in some situations, Condition (2) is automatically satisfied: it is automatic in case $\mathbb{K}$ is a field and the fibers of F are finite-dimensional (essentially, the proof of [La99, Lemma VIII.2.3] applies to this situation). Moreover, if cutoff functions in the sense of [La99] exist, then it suffices to work with $\widetilde{A}_M$ instead of the whole collection of the $\widetilde{A}_U$'s. See Chapter 21 for more general remarks on differential operators.

5. Lie groups and symmetric spaces: basic facts

5.1. *Manifolds with multiplication.* A *product* or *multiplication map* on a manifold M is a smooth binary map $m : M \times M \to M$, and *homomorphisms of manifolds with multiplication* are smooth maps that are compatible with the respective multiplication maps. *Left and right multiplication operators*, defined by $l_x(y) = m(x,y) = r_y(x)$, are partial maps of m and hence smooth self maps of M. Applying the tangent functor to this situation, we see that (TM, Tm) is again a manifold with multiplication, and tangent maps of homomorphisms are homomorphisms of the respective tangent bundles. The tangent map Tm is given by the formula

$$T_{(x,y)}m(v,w) = T_{(x,y)}m((v,0_y) + (0_x,w)) = T_x(r_y)v + T_y(l_x)w. \tag{5.1}$$

Formula (5.1) is nothing but the rule on partial derivatives (Lemma 1.5 (iv)), written in the language of manifolds. In particular, (5.1) shows that the canonical projection and the zero section,

$$\pi : TM \to M, \quad v \to \pi(v), \qquad z : M \to TM, \quad p \mapsto 0_p \tag{5.2}$$

are homomorphisms of manifolds with multiplication. We will always identify M with the subspace $z(M)$ of TM. Then (5.1) implies that the operator of left multiplication by $p = 0_p$ in TM is nothing but $T(l_p) : TM \to TM$, and similarly for right multiplications.

5.2. *Lie groups.* A *Lie group over* $\mathbb{K}$ is a smooth $\mathbb{K}$-manifold G carrying a group structure such that multiplication $m : G \times G \to G$ and inversion $i : G \to G$ are smooth. Homomorphisms of Lie groups are smooth group homomorphisms. Clearly, Lie groups and their homomorphisms form a category in which direct products exist.

Applying the tangent functor to the defining identities of the group structure (G, m, i, e), it is immediately seen that then $(TG, Tm, Ti, 0_{T_eG})$ is again a Lie group such that $\pi : TG \to G$ becomes a homomorphism of Lie groups and such that the zero section $z : G \to TG$ also is a homomorphism of Lie groups.

5.3. *The Lie algebra of a Lie group.* A vector field $X \in \mathfrak{X}(G)$ is called *left invariant* if, for all $g \in G$, $X \circ l_g = Tl_g \circ X$, and similarly we define *right invariant vector fields.* If X is right invariant, then $X(g) = X(r_g(e)) = T_e r_g \cdot X(e)$; thus X is uniquely determined by the value $X(e)$, and thus the map

$$\mathfrak{X}(G)^{r_G} \to T_eG, \quad X \mapsto X(e) \tag{5.3}$$

from the space of right invariant vector fields into T_eG is injective. It is also surjective: if $v \in T_eG$, then left multiplication with v in TG, $l_v : TG \to TG$, preserves fibers (because π is a homomorphism) and hence defines a vector field

$$v^R := l_v \circ z : G \to TG, \quad g \mapsto l_v(0_g) = v \cdot 0_g = Tm(v, 0_g) = T_e r_g(v)$$

which is right invariant since right multiplications commute with left multiplications. Now, the space $\mathfrak{X}(G)^{r_G}$ is a Lie subalgebra of $\mathfrak{X}(M)$; this follows immediately from Lemma 4.3 because X is right invariant iff the pair (X, X) is r_g-related for all $g \in G$. The space $\mathfrak{g} := T_eG$ with the Lie bracket defined by

$$[v, w] := [v^R, w^R]_e$$

is called the *Lie algebra of* G.

Theorem 5.4.

(i) *The Lie bracket* $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ *is of class* C^0.

(ii) *For every homomorphism* $f : G \to H$, *the tangent map* $\dot{f} := T_e f : \mathfrak{g} \to \mathfrak{h}$ *is a homomorphism of Lie algebras.*

Proof. (i) Pick a chart $\varphi : U \to V$ of G such that $\varphi(e) = 0$. Since $w^R(x) = Tm(w, x)$ depends smoothly on (x, w), it is represented in the chart by a smooth map (which again will be denoted by $w^R(x)$). But this implies that $[v^R, w^R](x) = d(v^R)(x)w^R(x) - d(w^R)(x)v^R(x)$ depends smoothly on v, w and x and hence $[v, w]$ depends smoothly on v, w.

(ii) First one checks that the pair of vector fields $(v^R, (\dot{f}v)^R)$ is f-related, and then one applies Lemma 4.3 in order to conclude that $\dot{f}[v, w] = [\dot{f}v, \dot{f}w]$. ■

The functor from Lie groups over $\mathbb{K}$ to C^0-Lie algebras over $\mathbb{K}$ will be called the *Lie functor (for* $\mathbb{K}$*-Lie groups and* $\mathbb{K}$*-Lie algebras).* We say that a C^0-Lie algebra over $\mathbb{K}$ is *integrable* if there is a Lie group over $\mathbb{K}$ such that $\mathfrak{g} = \mathrm{Lie}(G)$. The following is the *integrability problem for* $\mathbb{K}$*-Lie algebras*:

IP1. Which $\mathbb{K}$-Lie algebras are integrable ?

IP2. If a $\mathbb{K}$-Lie algebra is integrable, how can we describe all equivalence classes of Lie groups belonging to this Lie algebra ?

These are difficult problems which fall out of the scope of the present work. Let us just mention that all full matrix groups $\mathrm{Gl}(n, \mathbb{K})$ are Lie groups (with atlas given by the natural chart), as well as all orthogonal groups $\mathrm{O}(b, \mathbb{K}^n)$ corresponding to non-degenerate forms $b : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}$ – in these cases the manifold structure can be defined by "Cayley's rational chart", cf. [Be00] and [BeNe05]. For $\mathrm{SL}(n, \mathbb{K})$ one can define an atlas via the Bruhat decomposition; thus all "classical groups over $\mathbb{K}$" are Lie groups, and their Lie algebras are calculated in the usual way, namely, they are subalgebras of some matrix algebra $\mathfrak{gl}(V)$ with the usual Lie bracket $[X, Y] = XY - YX$, and $V \cong \mathbb{K}^m$. In fact, for $X \in \mathfrak{gl}(V)$, the value of the right invariant vector field X^R at $g \in \mathrm{Gl}(V)$ is $X^R(g) = Xg$ (product in $\mathrm{End}(V)$), and hence the Lie bracket can be calculated in the natural chart $\mathrm{End}(V)$:

$$[X, Y] = [X^R, Y^R]_e = d(X^R)(e) \cdot Y^R(e) - d(Y^R)(e) \cdot X^R(e) = XY - YX.$$

(This explains our choice of defining the Lie algebra via *right* invariant vector fields; the other choice would have given the opposite Lie algebra structure on $\mathfrak{gl}(V)$. Since we do not invoke any contravariant constructions, we can avoid the well-known "sign dilemma", cf. [Be00, p. 32/33], [Hel78, p. 122].) If V is not isomorphic to some $\mathbb{K}^m$, then in general there is no Lie group structure on $\mathrm{Gl}(V)$ – but a good substitute of general linear groups is given by unit groups of "continuous inverse algebras" (cf. Section 25.6). See [Gl04e] for these and more general constructions of Lie groups over topological fields.

5.5. *Symmetric spaces.* A *reflection space (over* $\mathbb{K}$*)* is a smooth manifold with a multiplication map $m : M \times M \to M$ such that, for all $x, y, z \in M$, and writing $\sigma_x(y) = m(x, y)$ instead of $l_x(y)$,

(M1) $m(x, x) = x$

(M2) $m(x, m(x, y)) = y$, i.e. $\sigma_x^2 = \mathrm{id}_M$, (M3) $m(x, m(y, z)) = m(m(x, y), m(x, z))$, i.e. $\sigma_x \in \mathrm{Aut}(M, m)$.

(Reflection spaces – "Spiegelungsräume" in German – have been introduced by O. Loos in [Lo67] in the finite dimensional real case.) The left multiplication operator σ_x is, by (M1)–(M3), an automorphism of order two fixing x; it is called the *symmetry around* x. The "trivial reflection space" $\sigma_x = \mathrm{id}_M$ for all x is not excluded by the axioms (M1)–(M3). We say that (M, m) is a *symmetric space (over* $\mathbb{K}$*)* if (M, m) is a reflection space such that 2 is invertible in $\mathbb{K}$ and the property

(M4) for all $x \in M$, $T_x(\sigma_x) = -\mathrm{id}_{T_x M}$

holds. The assumption that 2 is invertible in $\mathbb{K}$ guarantess that 0 is the only fixed point of the differential $T_x(\sigma_x) : T_x M \to T_x M$, and without this assumption (M4) would be useless. In the finite dimensional case over $\mathbb{K} = \mathbb{R}$, or, more generally, in any context where we have an implicit function theorem at out disposition, (M4) implies that x is an *isolated fixed point* of σ_x and hence our definition contains the one from [Lo69] as a special case (see [Ne02a, Lemma 3.2] for the case of Banach symmetric spaces). The group $G(M)$ generated by all $\sigma_x \sigma_y$ is a (normal) subgroup of $\mathrm{Aut}(M, m)$, called the *group of displacements.* A distinguished point $o \in M$ is called a *base point.* With respect to a base point, one defines the *quadratic representation*

$$Q := Q_o : M \to G(M), \quad x \mapsto Q(x) := \sigma_x \sigma_o \tag{5.4}$$

and the *powers* (for $n \in \mathbb{Z}$)

$$x^{-1} := \sigma_o(x), \quad x^{2n} := Q(x)^n.o, \quad x^{2n+1} := Q(x)^n.x. \tag{5.5}$$

By a straightforward calculation, one proves then the *fundamental formula*

$$Q(Q(x)y) = Q(x)Q(y)Q(x) \tag{5.6}$$

and the *power associativity rules* (cf. [Lo69] or [Be00, Lemma I.5.6])

$$m(x^n, x^m) = x^{2n-m}, \quad (x^m)^n = x^{mn}. \tag{5.7}$$

A symmetric space is called *abelian* or *commutative* if the group $G(M)$ is commutative. For instance, any topological $\mathbb{K}$-module with the product

$$m(u, v) = u - v + u = 2u - v \tag{5.8}$$

is a commutative symmetric space. See Section 26.1 for some more basic facts on commutative symmetric spaces.

Proposition 5.6. *Assume* (M, m) *is a symmetric space over* $\mathbb{K}$.

(i) *The tangent bundle* (TM, Tm) *of a reflection space is again a reflection space.*

(ii) *The tangent bundle* (TM, Tm) *of a symmetric space is again a symmetric space.*

Proof. (i) We express the identities (M1)–(M3) by commutative diagrams to which we apply the tangent functor T. Since T commutes with direct products, we get the same diagrams and hence the laws (M1)–(M3) for Tm (cf. [Lo69] for the explicit form of the diagrams).

(ii) We have to prove that (M4) holds for (TM, Tm). First of all, note that the fibers of $\pi : TM \to M$ (i.e. the tangent spaces) are stable under Tm because π is a homomorphism. We claim that for $v, w \in T_pM$ the explicit formula $Tm(v,w) = 2v - w$ holds (i.e. the structure induced on tangent spaces is the canonical "flat" symmetric structure (5.8) of an affine space). In fact, from (M1) for Tm we get $v = Tm(v,v) = T_p(\sigma_p)v + T_p(r_p)v = -v + T_p(r_p)v$, whence $T_p(r_p)v = 2v$ and

$$Tm(v,w) = T_p(\sigma_p)w + T_p(r_p)v = 2v - w.$$

Now fix $o \in M$ and $v \in T_oM$. We choose 0_o as base point in TM. Then $Q(v) = \sigma_v \sigma_{0_o}$ is, by (M3), an automorphism of (TM, Tm) such that $Q(v)0_o = \sigma_v(0_o) = 2v$. But multiplication by $\frac{1}{2}$,

$$\frac{1}{2} : TM \to TM, \quad u \mapsto \frac{1}{2}u$$

also is an automorphism of (TM, Tm), as shows Formula (5.1). Therefore the automorphism group of TM acts transitively on fibers, and after conjugation of σ_v with $(\frac{1}{2}Q(v))^{-1}$ we may assume that $v = 0_o$. But in this case the proof of our claim is easy: we have $\sigma_{0_o} = T\sigma_o$, and since $T_o\sigma_o = -\mathrm{id}_{T_oM}$, the canonical identification $T_{0_o}(TM) \cong T_oM \oplus T_oM$ yields $T_{0_o}(\sigma_{0_o}) = (-\mathrm{id}_{T_oM}) \times (-\mathrm{id}_{T_oM}) = -\mathrm{id}_{T_{0_o}TM}$, whence (M4). ■

The alert reader may have noticed that the preceding proof of (M4) already contains the construction of a canonical connection on TM: in fact, the argument of the proof shows that the "vertical space" $V_v := T_v(T_pM) \subset T_v(TM)$ has a canonical complement $H_v = \frac{1}{2}Q(v)H_0$, where H_0 is one of the factors of the canonical decomposition $T_{0_p}(TM) \cong T_pM \oplus T_pM$ – see Chapter 26 for the further theory concerning this.

5.7. *The algebra of derivations of* M. A vector field $X : M \to TM$ on a symmetric space M is called a *derivation* if X is also a homomorphism of symmetric spaces. This can be rephrased by saying that $(X \times X, X)$ is m-related. Lemma 4.3 therefore implies that the space $\mathfrak{g} \subset \mathfrak{X}(M)$ of derivations is stable under the Lie bracket. It is also easily checked that it is a $\mathbb{K}$-submodule of $\mathfrak{X}(M)$, and hence $\mathfrak{g} \subset \mathfrak{X}(M)$ is a Lie subalgebra.

5.8. *The Lie triple system of a symmetric space with base point.* We fix a base point $o \in M$. The map $X \mapsto T\sigma_o \circ X \circ \sigma_o$ is a Lie algebra automorphism of $\mathfrak{X}(M)$ of order 2 which stabilizes $\mathfrak{g}$. We let

$$\mathfrak{g} = \mathfrak{g}^+ \oplus \mathfrak{g}^-, \quad \mathfrak{g}^\pm = \{X \in \mathfrak{g} \,|\, T\sigma_o \circ X \circ \sigma_o = \pm X\}$$

be its associated eigenspace decomposition (recall our assumption on $\mathbb{K}$!). The space $\mathfrak{g}^+$ is a Lie subalgebra of $\mathfrak{X}(M)$, whereas $\mathfrak{g}^-$ is only closed under the triple bracket

$$(X,Y,Z) \mapsto [X,Y,Z] := [[X,Y],Z].$$

Proposition 5.9.

(i) *The space* $\mathfrak{g}^+$ *is the kernel of the evaluation map* $\mathrm{ev}_o : \mathfrak{g} \to T_oM$, $X \mapsto X(o)$.

(ii) *Restriction of* ev_o *yields a bijection* $\mathfrak{g}^- \to T_oM$, $X \mapsto X(o)$.

Proof. (i) Assume $X \in \mathfrak{g}^+$. Then $T_o\sigma X(o) = X(\sigma_o(o)) = X(o)$ implies $-X(o) = X(o)$ and hence $X(o) = 0$. On the other hand, if $X(o) = 0$, then $X(\sigma_o(p)) = X(m(o,p)) = Tm(X(o), X(p)) = Tm(0_o, X(p)) = T\sigma_o X(p)$, whence $X \in \mathfrak{g}^+$.

(ii) By (i), $\mathfrak{g}^- \cap \ker(\mathrm{ev}_o) = \mathfrak{g}^- \cap \mathfrak{g}^+ = 0$, and hence $\mathrm{ev}_o : \mathfrak{g}^- \to T_oM$ is injective. It is also surjective: let $v \in T_oM$. Consider the map

$$\widetilde{v} = \frac{1}{2}Q(v) \circ z : M \to TM, \quad p \mapsto \frac{1}{2}Q(v)0_p = \frac{1}{2}Tm(v, Tm(0_o, 0_p)). \tag{5.9}$$

It is a composition of homomorphisms and hence is itself a homomorphism from M into TM. Moreover, as seen in the proof of Proposition 5.6, $\widetilde{v}(o) = v$. Thus we will be done if can show that $\widetilde{v} \in \mathfrak{g}^-$. First of all, $\widetilde{v}$ is a vector field since, for all $p \in M$ and $w \in T_pM$, $Q(v)w \in T_{m(o,m(o,p))}M = T_pM$. Finally,

$$\begin{aligned} T\sigma_o \circ \widetilde{v} \circ \sigma_o &= \frac{1}{2}T\sigma_o \circ Q(v) \circ z \circ \sigma_o \\ &= \frac{1}{2}Q(T\sigma_o v) \circ z = \frac{1}{2}Q(-v) \circ z = -\widetilde{v}. \end{aligned}$$

(The notation $\widetilde{v}$ is taken from [Lo69]. In later chapters and in [Be00] we use also the notation $l(v) := \widetilde{v}$ and consider $l(v)$ as a "vector field extension of v".) ∎

The space $\mathfrak{m} := T_oM$ with triple bracket given by

$$[u, v, w] := [[\widetilde{u}, \widetilde{v}], \widetilde{w}](o)$$

is called the *Lie triple system (Lts) associated to* (M, o). It satisfies the identities of an abstract Lie triple system over $\mathbb{K}$ (cf. [Lo69, p. 78/79] or [Be00]): for all $u, v, w \in \mathfrak{m}$,

(LT1) $[u, u, v] = 0$,

(LT2) $[u, v, w] + [v, w, u] + [w, u, v] = 0$,

(LT3) the operator $R(u, v) := [u, v, \cdot]$ is a derivation of the trilinear product $[\cdot, \cdot, \cdot]$.

In fact, these identities are easy consequences of the defining identities of a Lie algebra: they hold in $\mathfrak{g}$ and hence also in the subspace $\mathfrak{g}^-$ of $\mathfrak{g}$ which is stable under taking triple Lie brackets.

Theorem 5.10. *Let* M *be a symmetric space over* $\mathbb{K}$ *with base point* o.

(i) *The triple Lie bracket of the Lts* $\mathfrak{m}$ *associated to* (M, o) *is of class* C^0.

(ii) *If* $\varphi : M \to M'$ *is a homomorphism of symmetric spaces such that* $\varphi(o) = o'$, *then* $\dot{\varphi} := T_o\varphi : \mathfrak{m} \to \mathfrak{m}'$ *is a homomorphism of associated Lts.*

Proof. One uses the same arguments as in the proof of Theorem 5.4. ∎

5.11. *Lie functor, group case, and the classical spaces.* The functor described by the preceding theorem will be called the *Lie functor for symmetric spaces and Lie triple systems (defined over* $\mathbb{K}$*)*, and the same integrability problem as for Lie groups arises. It generalizes the integrability problem for Lie groups because every Lie group can be turned into a symmetric space by letting

$$\mu : G \times G \to G, \quad (x, y) \mapsto xy^{-1}x;$$

then (G, μ) is a symmetric space, called *of group type* (Properties (M1)–(M3) are immediate; for (M4) one proves as usual in Lie group theory that $T_e j = -\operatorname{id}_{T_e G}$ for the inversion map $j : G \to G$); the Lie triple system of (G, μ, e) is then $\mathfrak{g}$ with the triple bracket

$$[X, Y, Z] = \frac{1}{4}[[X, Y], Z]$$

– for proving this, one can use the same arguments as in [Lo69, p. 81]. Note that, in the group case, $\operatorname{Aut}(G, \mu)$ acts transitively on the space since all left and right translations are automorphisms; however, the action of the transvection group $G(G, \mu)$ is in general not transitive – e.g., take the Lie group $\mathrm{Gl}(n, \mathbb{K})$; we have $\sigma_x \sigma_y = l_x r_x l_{y^{-1}} r_{y^{-1}}$ in terms of left and right translations, and hence the determinant of $\sigma_x \sigma_y(z)$ is congruent to the determinant of z modulo a square in $\mathbb{K}$. This shows that the orbit structure of $G = \mathrm{Gl}(n, \mathbb{K})$ under the action of $G(G, \mu)$ is at least as complicated as the structure of $\mathbb{K}^\times / (\mathbb{K}^\times)^2$.

The "classical spaces" can all be constructed in the following way: assume G is a Lie group and σ an involution of G; we assume moreover that the *space of symmetric elements*

$$M := \{g \in G \mid \sigma(g) = g^{-1}\} \tag{5.11}$$

is a submanifold of G in the sense of 2.3. (The last assumption is not automatically satisfied since we do not have an exponential map at our disposition; however, in all "classical cases" mentioned below, it is easily checked either directly or by remarking that in "Jordan coordinates" it is indeed automatic, cf. [Be00].) Then M is a subsymmetric space of (G, μ) in the obvious sense. If $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{q}$ is the decomposition of $\mathfrak{g} = \operatorname{Lie}(G)$ into $+1$- and -1-eigenspaces of $T_e\sigma$, then, as in [Lo69, p. 82], it is seen that the Lie triple system of (M, μ, e) is $\mathfrak{q}$ with $[X, Y, Z] = [[X, Y], Z]$. For an arbitrary base point $g \in M$, the Lie triple system structure on $T_g M$ is described in the same way as for $g = e$, but replacing σ by the involution $I_g \circ \sigma$ where $I_g(x) = gxg^{-1}$ is conjugation by g. In contrast to the group case, here the action of the automorphism group $\operatorname{Aut}(M)$ is in general no longer transitive on M. For instance, if $G = \mathrm{Gl}(n, \mathbb{K})$ and $\sigma = \mathrm{id}$, then the orbits in $M = \{g \in G \mid g^2 = \mathbf{1}\}$ are described by the matrices $I_{p,q} = \begin{pmatrix} \mathbf{1}_p & 0 \\ 0 & -\mathbf{1}_q \end{pmatrix}$, $p + q = n$. Similarly, taking for G a classical matrix group (in the sense explained after Theorem 5.4) and for σ a "classical involution" (see [Be00, Chapter I.6] for a fairly exhaustive list in the real case) we get analogs of all matrix series from Berger's classification of irreducible real symmetric spaces [B57], but where the number of non-isomorphic types may be considerably bigger in case of base-fields or rings different from $\mathbb{R}$ or $\mathbb{C}$. All these symmetric spaces are "Jordan symmetric spaces" in the sense of [BeNe05] where a very general construction of symmetric spaces (of classical or non-classical type) is described.

II. Interpretation of tangent objects via scalar extensions

6. Scalar extensions. I: Tangent functor and dual numbers

6.1. *The ring of dual numbers over* $\mathbb{K}$. The *ring of dual numbers over* $\mathbb{K}$ is the truncated polynomial ring $\mathbb{K}[x]/(x^2)$; it can also be defined in a similar way as complex numbers, replacing the condition $i^2 = -1$ by $\varepsilon^2 = 0$:

$$\mathbb{K}[\varepsilon] := \mathbb{K} \oplus \varepsilon\mathbb{K}, \quad (a + \varepsilon b)(a' + \varepsilon b') = aa' + \varepsilon(ab' + ba'). \tag{6.1}$$

In analogy with the terminology for complex numbers, let us call $z = x + \varepsilon y$ a *dual number*, x its *spacial part*, y its *infinitesimal part* and ε the *infinitesimal* or *dual number unit.* A dual number is invertible iff its spacial part is invertible, and inversion is given by

$$(x + \varepsilon y)^{-1} = \frac{x - \varepsilon y}{x^2}. \tag{6.2}$$

It follows that $\mathbb{K}[\varepsilon]$ is again a topological ring with dense group of invertible elements. Taking the spacial part is a ring homomorphism $\mathbb{K}[\varepsilon] \to \mathbb{K}$, and its kernel is the ideal $\varepsilon\mathbb{K}$ of "purely infinitesimal" numbers. If $\mathbb{K}$ is a field, then this is the unique maximal ideal of $\mathbb{K}[\varepsilon]$. A matrix realization of $\mathbb{K}[\varepsilon]$ in the algebra of 2×2-matrices over $\mathbb{K}$ is given by

$$\mathbb{K}[\varepsilon] \cong \mathbb{K}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \oplus \mathbb{K}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{6.3}$$

For any topological $\mathbb{K}$-module V, the scalar extension

$$V \otimes_{\mathbb{K}} \mathbb{K}[\varepsilon] = V \otimes_{\mathbb{K}} (\mathbb{K} \oplus \varepsilon\mathbb{K}) = V \oplus \varepsilon V$$

(with $\mathbb{K}[\varepsilon]$-action given by $(a + \varepsilon b)(v + \varepsilon w) = ab + \varepsilon(aw + bv)$) is a topological $\mathbb{K}[\varepsilon]$-module. In particular, it is a smooth manifold over $\mathbb{K}[\varepsilon]$. More generally, for every open set $U \subset V$, the set $TU := U \oplus \varepsilon V$ is open in $V \oplus \varepsilon V$ and hence is a manifold over $\mathbb{K}[\varepsilon]$.

Theorem 6.2. *If M is a manifold of class C^{k+1} over $\mathbb{K}$, modelled on V, then TM is, in a canonical way, a manifold of class C^k over $\mathbb{K}[\varepsilon]$, modelled on the scalar extended $\mathbb{K}[\varepsilon]$-module $V \oplus \varepsilon V$. If $f : M \to N$ is of class C^{k+1} over $\mathbb{K}$, then $Tf : TM \to TN$ is of class C^k over $\mathbb{K}[\varepsilon]$. Thus the tangent functor can be characterized as the "functor of scalar extension" from manifolds over $\mathbb{K}$ into manifolds over $\mathbb{K}[\varepsilon]$ agreeing on open submanifolds of $\mathbb{K}$-modules with the algebraic scalar extension functor.*

Proof. Change of charts on TM is described by the transition maps $T\varphi_{ij}(x, v) = (\varphi_{ij}(x),\ d\varphi_{ij}(x)v)$. Therefore the theorem follows from the following claim:

Theorem 6.3. *Assume V and W are topological $\mathbb{K}$-modules, U open in V and $f : V \supset U \to W$ is of class C^{k+1} over $\mathbb{K}$. Then $Tf : V \oplus \varepsilon V \supset U \times V \to W \times W = W \oplus \varepsilon W$ is of class C^k over $\mathbb{K}[\varepsilon]$. More precisely, the "higher order difference quotient maps" defined in* Section 1.7 *are related by*

$$(Tf)^{[k]} = T(f^{[k]}),$$

where the difference quotient map on the left hand side is taken with respect to $\mathbb{K}[\varepsilon]$.

Proof. It suffices to prove the claim for $k = 1$, i.e. that

$$(Tf)^{[1]} = T(f^{[1]}),$$

the general case then follows by a straightforward induction. Assume $f : V \supset U \to W$ is a map of class C^2. We are going to apply the tangent functor T to the defining identity

$$t \cdot f^{[1]}(x, v, t) = f(x + tv) - f(x) \qquad (!)$$

of $f^{[1]}$. Let us write (!) in the form of a commutative diagram:

$$\begin{array}{ccccc}
U^{[1]} & \overset{\beta}{\to} & U^{[1]} \times U & \overset{\gamma}{\to} & U \times U \\
\alpha \downarrow & & & & \downarrow f \times f \\
\mathbb{K} \times U^{[1]} & & & & W \times W \\
\mathrm{id}_{\mathbb{K}} \times f^{[1]} \downarrow & & & & \downarrow \varphi \\
\mathbb{K} \times W & \overset{m_W}{\to} & W & = & W
\end{array}$$

where $\alpha(x, v, t) = (t, x, v, t)$, $\beta(x, v, t) = (x, v, t, x)$, $\gamma(x, v, t, y) = (x + tv, y)$ and $\varphi(u, v) = u - v$, i.e.

$$\begin{array}{ccccc}
(x, v, t) & \mapsto & (x, v, t, x) & \mapsto & (x + tv, x) \\
\downarrow & & & & \downarrow \\
(t, x, v, t) & & & & (f(x + tv), f(x)) \\
\downarrow & & & & \downarrow \\
(t, f^{[1]}(x, v, t)) & \mapsto & tf^{[1]}(x, v, t) & = & f(x + tv) - f(x)
\end{array}$$

Applying the tangent functor T to this diagram, we get a diagram of the same type, where all maps are replaced by their tangent maps. (Note that this is the same argument as in the proof that the tangent group TG of G is a Lie group, resp. the analog for symmetric spaces, cf. Chapter 5; it only uses that T is a covariant functor commuting with direct products and diagonal imbeddings.) Let us describe the tangent map $Tm : T\mathbb{K} \times T\mathbb{K} \to T\mathbb{K}$ of $m : \mathbb{K} \times \mathbb{K} \to \mathbb{K}$: for $t \in \mathbb{K}^\times$,

$$m^{[1]}((x_1, x_2), (v_1, v_2), t) = \frac{1}{t}(m(x_1 + tv_1, x_2 + tv_2) - m(x_1, x_2)) = v_1 x_2 + x_1 v_2 + t v_1 v_2,$$

and letting $t = 0$, we get $Tm((x_1, x_2), (v_1, v_2)) = (x_1 x_2, v_1 x_2 + x_1 v_2)$. Comparing with (6.1), we see that $(T\mathbb{K}, Tm)$ is isomorphic to the ring $\mathbb{K}[\varepsilon]$ of dual numbers over $\mathbb{K}$. For this reason, $\mathbb{K}[\varepsilon]$ will also be called the *tangent ring of* $\mathbb{K}$. In a similar way, we calculate the tangent maps of the structural maps $m_V : \mathbb{K} \times V \to V$ and $a : V \times V \to V$ of the topological $\mathbb{K}$-module V are C^1 and find that

$$\begin{aligned} Tm_V((r,s),(x,v)) &= (rx, rv + sx), \\ Ta((x,v),(x',v')) &= (x + x', v + v') \end{aligned}$$

which is also obtained by writing $r+\varepsilon s$ for (r,s) and $x+\varepsilon v$ for (x,v) and calculating in the scalar extended module $V \oplus \varepsilon V$. Hence $TV \cong V \oplus \varepsilon V$ as modules over the ring $T\mathbb{K} \cong \mathbb{K}[\varepsilon]$, and we call $V \oplus \varepsilon V$ also the *tangent module of* V. Summing up, when applying the tangent functor to (!), all structural maps for V, W are replaced by the corresponding structural maps for the corresponding scalar extended modules, and f and $f^{[1]}$ are replaced by Tf and $T(f^{[1]})$. This in turn means that the identity

$$t \cdot T(f^{[1]})(x,v,t) = Tf(x + tv) - Tf(x) \tag{!!}$$

holds for all $(x,v,t) \in TU \times TV \times T\mathbb{K}$ with $x + tv \in TU$, and this means that Tf is of class C^1 over the ring $T\mathbb{K}$, and that $(Tf)^{[1]} = T(f^{[1]})$. ■

6.4. *Application to Lie groups and symmetric spaces.* For any $\mathbb{K}$-Lie algebra $\mathfrak{g}$, the scalar extended algebra $\mathfrak{g} \otimes_{\mathbb{K}} \mathbb{K}[\varepsilon]$ can be described as

$$\mathfrak{g} \oplus \varepsilon\mathfrak{g}, \quad [X + \varepsilon Y, X' + \varepsilon Y'] = [X, X'] + \varepsilon([X, Y'] + [X', Y]),$$

i.e. one simply takes the ε-bilinear bracket on $\mathfrak{g} \oplus \varepsilon\mathfrak{g}$ and takes account of the relation $\varepsilon^2 = 0$. Similarly, the scalar extended Lie triple system $\mathfrak{m} \otimes_{\mathbb{K}} \mathbb{K}[\varepsilon] \cong \mathfrak{m} \oplus \varepsilon\mathfrak{m}$ of a $\mathbb{K}$-Lie triple system $\mathfrak{m}$ can be described.

Theorem 6.5. *If G is a Lie group over $\mathbb{K}$, then TG is a Lie group over $\mathbb{K}[\varepsilon]$ with Lie algebra $\mathfrak{g} \oplus \varepsilon\mathfrak{g}$, and if M is a symmetric space over $\mathbb{K}$ with base point o and associated Lie triple system $\mathfrak{m}$, then TM is a symmetric space over $\mathbb{K}[\varepsilon]$ with base point 0_o and associated Lie triple system $\mathfrak{m} \oplus \varepsilon\mathfrak{m}$.*

Proof. If M is a manifold with product map $m : M \times M \to M$, then, by Theorem 6.2, $Tm : TM \times TM \to TM$ is smooth over $\mathbb{K}[\varepsilon]$. Thus, if G is a Lie group over $\mathbb{K}$, then TG is a Lie group, not only over $\mathbb{K}$, but also over $\mathbb{K}[\varepsilon]$. Moreover, all constructions used to define the Lie bracket of G naturally commute with the functor of scalar extension by $\mathbb{K}[\varepsilon]$, and hence the Lie algebra of TG is the scalar extended Lie algebra $\mathfrak{g} \otimes_{\mathbb{K}} \mathbb{K}[\varepsilon]$ of $\mathfrak{g}$. Similar arguments apply to symmetric spaces. ■

6.6. *Comparison of the structures of $T\mathbb{K}$ and of TM.* The fiber over 0 in the ring $T\mathbb{K}$ is the ideal $\varepsilon\mathbb{K}$; it is the kernel of the canonical projection $T\mathbb{K} \to \mathbb{K}$, which is a ring homomorphism. In other words, we have an exact sequence of non-unital rings

$$0 \to \varepsilon\mathbb{K} \to T\mathbb{K} \to \mathbb{K} \to 0. \tag{6.4}$$

On the level of manifolds, this corresponds to the inclusion of the tangent space T_oM at a base point $o \in M$, followed by the canonical projection:

$$T_oM \to TM \to M.$$

By a direct computation, one checks that for any $\lambda \in \mathbb{K}$, the map

$$l_\lambda : \mathbb{K}[\varepsilon] \to \mathbb{K}[\varepsilon], \quad r + \varepsilon s \mapsto r + \varepsilon \lambda s$$

is a ring homomorphism and that

$$l : \mathbb{K} \times T\mathbb{K} \to T\mathbb{K}, \quad (\lambda, z) \mapsto l_\lambda(z)$$

is an action of $\mathbb{K}$ on $T\mathbb{K}$. If $\lambda = -1$, we may call l_{-1} the "spatial conjugation"; if $\lambda = 0$, l_0 corresponds to the projection onto $\mathbb{K}$. On the level of manifolds, we have corresponding smooth maps

$$l_\lambda : TM \to TM, \; v \mapsto \lambda v, \qquad l : \mathbb{K} \times TM \to TM$$

where the product is taken in the $\mathbb{K}$-module $T_{\pi(v)}M$. Finally, there is also an inclusion map $\mathbb{K} \to \mathbb{K}[\varepsilon]$, which corresponds to the zero section

$$z_{TM} : M \to TM, \quad x \mapsto 0_x.$$

6.7. *The trivial scalar extension functor.* Not every manifold over $\mathbb{K}[\varepsilon]$ is (isomorphic to) a tangent bundle, not even locally. In fact, since $\varepsilon\mathbb{K}$ is an ideal in $T\mathbb{K}$, every $\mathbb{K}$-module is also a $\mathbb{K}[\varepsilon]$-module, just by letting act ε trivially (composition of the action of $\mathbb{K}$ and the ring homomorphism $\mathbb{K}[\varepsilon] \to \mathbb{K}$, $r + \varepsilon s \mapsto r$). In a similar way, every manifold over $\mathbb{K}$ is also a manifold over $\mathbb{K}[\varepsilon]$, just by letting act ε trivially everywhere; we call this *the trivial scalar extension functor*. Therefore, in order to characterize the $\mathbb{K}[\varepsilon]$-manifold structure in Theorem 6.2, it was necessary to state explicitly that, on chart domains, it coincides with the algebraic scalar extension. Occasionally, it may be useful to consider $\mathbb{K}$-manifolds M as $\mathbb{K}[\varepsilon]$-manifolds in the trivial way. For instance, the projection $\pi : TM \to M$ is then smooth over $\mathbb{K}[\varepsilon]$ – indeed, in a chart the projection $V \oplus \varepsilon V \to V$ is then $\mathbb{K}[\varepsilon]$-linear, whence smooth over $\mathbb{K}[\varepsilon]$.

Theorem 6.2 implies that $\mathrm{Diff}_{\mathbb{K}}(M) \to \mathrm{Diff}_{T\mathbb{K}}(TM)$, $g \mapsto Tg$ is a well-defined imbedding. In fact, it is not too difficult to determine now the full group of $T\mathbb{K}$-diffeomorphisms of TM:

Theorem 6.8. *Assume M and N are manifolds over $\mathbb{K}$.*

(1) *For any vector field $X : M \to TM$, the infinitesimal automorphism $\widetilde{X} : TM \to TM$, $v \mapsto v + X(\pi(v))$ (cf.* Section 4.4*) is smooth over $\mathbb{K}[\varepsilon]$.*

(2) *Assume $F : TM \to TN$ is smooth over $T\mathbb{K}$. Then F maps fibers to fibers (i.e., there is a smooth $f : M \to N$ with $\pi_N \circ F = f \circ \pi_M$) and F acts $\mathbb{K}$-affinely from fibers to fibers.*

(3) *The following is an exact and splitting sequence of groups:*

$$0 \;\to\; \mathfrak{X}(M) \;\overset{X \mapsto \widetilde{X}}{\to}\; \mathrm{Diff}_{T\mathbb{K}}(TM) \;\overset{F \mapsto f}{\to}\; \mathrm{Diff}_{\mathbb{K}}(M) \;\to 1.$$

In particular, any $T\mathbb{K}$-smooth diffeomorphism $F : TM \to TM$ may be written in a unique way as $F = \widetilde{X} \circ Tf$ for some $X \in \mathfrak{X}(M)$ and $f \in \mathrm{Diff}(M)$.

Proof. (1) We use the following simple lemma:

Lemma 6.9. *Assume* $F : V \oplus \varepsilon V \supset U \oplus \varepsilon V \to W \oplus \varepsilon W$ *is of the form* $F(a+\varepsilon b) = \varepsilon g(a)$ *with* $g : U \to W$ *of class* C^∞ *over* $\mathbb{K}$. *Then* F *is of class* C^∞ *over* $T\mathbb{K}$.

Proof. The claim can be checked by a direct computation. Here is a computation-free argument: scalar multiplication by ε,

$$l_\varepsilon : W \oplus \varepsilon W \to W \oplus \varepsilon W, \quad x + \varepsilon v \mapsto \varepsilon x,$$

is $\mathbb{K}[\varepsilon]$-linear and continuous, whence of class C^∞ over $\mathbb{K}[\varepsilon]$. Now, we may write $F(a+\varepsilon b) = \varepsilon(g(a) + \varepsilon dg(a)b) = l_\varepsilon \circ Tg(a+\varepsilon b)$, and thus F is seen to be a composition of $\mathbb{K}[\varepsilon]$-smooth maps. ■

Coming back to the proof of the theorem, in a bundle chart we write $\widetilde{X}(x+\varepsilon v) = x + \varepsilon v + \varepsilon X(x) = \mathrm{id}(x+\varepsilon v) + \varepsilon X(x)$, and hence $\widetilde{X}$ is represented by the sum of the identity map and a map having the form given in the preceding lemma. It follows that $\widetilde{X}$ is smooth over $\mathbb{K}[\varepsilon]$.

(2) We represent F with respect to charts and write again $F : TU \to TU'$ for the chart representation. Since F is smooth over $T\mathbb{K}$, we have the following second order Taylor expansion of F at the point $x = x + \varepsilon 0 \in TU$ (Theorem 1.11):

$$F(x+\varepsilon v) = F(x) + \varepsilon dF(x)v + \varepsilon^2 a_2(x,v) + \varepsilon^2 R_2(x,v,\varepsilon) = F(x) + \varepsilon dF(x)v. \quad (6.5)$$

Since $\varepsilon dF(x)v \in \varepsilon W$, this shows that F maps fibers to fibers. Next, we decompose $F(x+\varepsilon v) = F_1(x+\varepsilon v) + \varepsilon F_2(x+\varepsilon v)$ into its "spacial" and "infinitesimal part" $F_i : TU \to W$. We let also $f := F_1|_U : U \to W$, which is the chart representation of the map $f := \pi_N \circ F \circ z_M : M \to N$. Over $\mathbb{K}$, we may write $dF = d(F_1 + \varepsilon F_2) = dF_1 + \varepsilon dF_2$, and hence we get from (6.5)

$$\begin{aligned} F(x+\varepsilon v) &= F(x) + \varepsilon dF(x)v = F_1(x) + \varepsilon(dF_1(x)v + F_2(x)) \\ &= f(x) + \varepsilon df(x)v + \varepsilon F_2(x) = Tf(x)v + \varepsilon F_2(x). \end{aligned}$$

Thus F acts affinely on fibers (the linear part is $df(x)$ and the translation part is translation by $F_2(x)$).

(3) The first map is well-defined by (1); clearly, it is an injective group homomorphism. The second map is surjective since it admits the section $g \mapsto Tg$. Clearly $\mathfrak{X}(M)$ belongs to the kernel of the second map. Conversely, assume $F \in \mathrm{Diff}_{T\mathbb{K}}(TM)$ is such that $f = \mathrm{id}_M$. In a chart representation as in the proof of Part (2), this means that $F_1(x) = x$. But as seen above, then the linear part of $F|_{T_xM}$ is the identity, hence $F|_{T_xM}$ acts as a translation for all x, and thus F is an infinitesimal automorphism. This proves exactness of the given sequence.

Finally, as for any split exact sequence, the middle group $\mathrm{Diff}_{T\mathbb{K}}(TM)$ can now be written as a semidirect product of $\mathfrak{X}(M)$ and $\mathrm{Diff}(M)$, implying the last statement. ■

In fact, the proof of Part (2) of the theorem shows a bit more than anounced, namely that any $T\mathbb{K}$-smooth map $F : TM \to TN$ can be written in the form $F(u) = Tf(u) + Y(\pi_M(u))$ where $Y : M \to TN$ is a "vector field over f" (represented in a chart by $F_2|_U : U \to W$), i.e., $\pi_N \circ Y = f$. – In Chapter 28 these facts will be generalized, and we will show that the sequence from the preceding theorem is the precise analog of the sequence $0 \to \mathfrak{g} \to TG \to G \to 1$ of a Lie group G.

6.10. *Remark on the interpretation of the tangent functor as a contraction of the direct product functor.* Intuitively, the tangent bundle may be seen in the following way: take the direct product $M \times M$, realize M as the diagonal in $M \times M$, take the vertical fiber over the point $x = (x, x)$ and enlarge it more and more (with fixed center (x, x)) until you only see the "infinitesimal neighborhood of x", and finally forget the rest. What you get is the tangent bundle, seen as a "contraction of the direct product". This intuitive picture can be put on a completely rigorous footing: look at the family of rings $\mathbb{K}[x]/(x^2 - tx)$ with $t \in \mathbb{K}$; taking the classes of 1 and x as a $\mathbb{K}$-basis, this ring is described by the product

$$(a + bx)(c + dx) = ac + (ad + bc + tbd)x. \tag{6.6}$$

A matrix realization of this ring in the algebra of 2×2-matrices over $\mathbb{K}$ is given by

$$\mathbb{K}[x]/(x^2 - tx) \cong \mathbb{K}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \oplus \mathbb{K}\begin{pmatrix} t & 0 \\ 1 & 0 \end{pmatrix}.$$

If t is invertible in $\mathbb{K}$ (in particular, for $t = 1$), this ring is isomorphic to the direct product of rings $\mathbb{K} \times \mathbb{K}$: take, in the ring $\mathbb{K} \times \mathbb{K}$ with pointwise product $(x, y)(x', y') = (xx', yy')$, the new basis $e = (1, 1), f = (0, t)$; e is the unit element in $\mathbb{K} \times \mathbb{K}$, and $f^2 = tf$; hence the product in $\mathbb{K} \times \mathbb{K}$ is described by $(x_1 e + v_1 f)(x_2 e + v_2 f) = x_1 x_2 e + (x_1 v_2 + v_1 x_2 + t v_1 v_2) f$, which is the same as (6.6). If $t = 0$ we get the dual numbers over $\mathbb{K}$; they appear here as a "contraction for $t \to 0$" of the rings $\mathbb{K} \times \mathbb{K} \cong \mathbb{K}[x]/(x^2 - tx)$. In a similar way, differential calculus can be seen as a contraction of difference calculus. Namely, recall from Equation (1.4) the definition of the extended tangent map $\widehat{T} f(x, v, t) = (f(x), f^{[1]}(x, v, t), t)$. Since the functorial rule $\widehat{T}(f \circ g) = \widehat{T} f \circ \widehat{T} g$ holds, all arguments used in the proof of Theorem 6.3 can be adapted to the case of the functor $\widehat{T}(--, t)$. (Note that, for invertible t, this functor is, by a change of variables, isomorphic to the direct product functor.) In particular, applying this functor to the ring $\mathbb{K}$ itself, we get

$$\widehat{T}(m)((x_1, x_2), (v_1, v_2), t)) = (x_1 x_2, v_1 x_2 + x_1 v_2 + t v_1 v_2, t).$$

Fixing and suppressing the last variable, this leads to a product $(\mathbb{K} \times \mathbb{K}) \times (\mathbb{K} \times \mathbb{K}) \to \mathbb{K} \times \mathbb{K}$ defined by

$$(x_1, v_1) \cdot (x_2, v_2) := \widehat{T} m((x_1, x_2), (v_1, v_2), t) = (x_1 x_2, v_1 x_2 + x_1 v_2 + t v_1 v_2).$$

As in the case $t = 0$, this defines on $\mathbb{K} \times \mathbb{K}$ the structure of a topological ring; comparing with (6.5), we see that this is nothing but the ring $\mathbb{K}[x]/(x^2 - tx)$. Applying our functor $\widehat{T}(--, t)$ to (!), we thus see that for all $t \in \mathbb{K}$, $\widehat{T} f(--, t)$ is of class C^k over the ring $\mathbb{K}/(x^2 - tx)$.

For invertible t, and in particular for $t = 1$, the preceding claims can all be checked by direct and elementary calculations. Taking the "limit case $t \to 0$", we may again deduce Theorem 6.3, thus giving another proof of it. However, proving the limit case Theorem 6.3 by checking the definition of differentiability over $\mathbb{K}[\varepsilon]$ directly, leads to fairly long and involved calculations. (The reader may try this as an exercise, to get accustomed to the definitions.)

7. Scalar extensions. II: Higher order tangent functors

7.1. *Higher order tangent bundles.* For any manifold M, the *second order tangent bundle* is $T^2M := T(TM)$, and the *k-th order tangent bundle* is inductively defined by $T^kM := T(T^{k-1}M)$. A repeated application of Theorem 6.2 shows:

Theorem 7.2. *The k-th order tangent bundle T^kM is, in a canonical way, a manifold over the ring of iterated dual numbers $T^k\mathbb{K} = \mathbb{K}[\varepsilon_1]\ldots[\varepsilon_k]$, and if $f: M \to N$ is smooth over $\mathbb{K}$, then $T^kf: T^kM \to T^kN$ is smooth over $T^k\mathbb{K}$.* ∎

Thus the higher order tangent functor T^k can be seen as the functor of scalar extension by the ring $T^k\mathbb{K}$. We will add the index ε_i to the tangent functor symbol T if we want to indicate that we mean scalar extension with "new infinitesimal unit" ε_i. Psychologically, we associate ε_1 to the "first" scalar extension and ε_k to the "last" scalar extension; but we will see below that all orders are equivalent, i.e., there is an action of the permutation group Σ_k by automorphisms of the whole structure. – When working with chart representations, we may use, for bundle chart domains of TM over $U \subset M$, resp. of TTM over U, the following two equivalent notations

$$\begin{aligned} TU &= \{(x,v)|\, x \in U, v \in V\} = \{x + \varepsilon v|\, x \in U, v \in V\}, \\ TTU &= \{(x, v_1, v_2, v_{12})| x \in U, v_1, v_2, v_{12} \in V\} \\ &= \{x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12} | x \in U, v_1, v_2, v_{12} \in V\} \end{aligned}$$

but the more classical component notation will be gradually abandoned in favor of the "algebraic notation" which turns out to be more suitable (the component notation corresponds to the one used in [La99] and in most other references where TTM is treated).

7.3. *Structure of $TT\mathbb{K}$ and of TTM.* First of all, one should note that TTM is a fiber bundle over M, but it is *not* a vector bundle – the transition maps, calculated in Section 4.7, are not linear in fibers. The "algebraic structure" of TTM will be discussed in Chapter 9. Here we focus on properties that are directly related to properties of the ring

$$TT\mathbb{K} \cong (\mathbb{K}[\varepsilon_1])[\varepsilon_2] = \mathbb{K} \oplus \varepsilon_1\mathbb{K} \oplus \varepsilon_2\mathbb{K} \oplus \varepsilon_1\varepsilon_2\mathbb{K},$$

(relations: $\varepsilon_1^2 = 0 = \varepsilon_2^2$ and $\varepsilon_1\varepsilon_2 = \varepsilon_2\varepsilon_1$). We are going to discuss the following features of $TT\mathbb{K}$: projections; injections; ideal structure; canonical automorphisms and endomorphisms.

(A) Projections. The projections onto the "spacial part" corresponding to the various extensions fit together to a commutative diagram

$$\begin{array}{ccccc} & & \mathbb{K}[\varepsilon_1][\varepsilon_2] & & \\ & \swarrow & & \searrow & \\ \mathbb{K}[\varepsilon_1] & & & & \mathbb{K}[\varepsilon_2]\,. \\ & \searrow & & \swarrow & \\ & & \mathbb{K} & & \end{array} \qquad (7.1)$$

For every $\mathbb{K}$-module V, we get by scalar extension a $TT\mathbb{K}$-module TTV and a diagram of $\mathbb{K}$-modules corresponding to (7.1), which, on the level of manifolds, corresponds to the diagram of vector bundle projections

$$\begin{array}{ccccc} & & TTM & & \\ & \swarrow & & \searrow & \\ T_{\varepsilon_1}M & & & & T_{\varepsilon_2}M\,. \\ & \searrow & & \swarrow & \\ & & M & & \end{array} \tag{7.2}$$

In a bundle chart, (7.2) is represented by

$$\begin{array}{ccccc} & & x+\varepsilon_1 v_1+\varepsilon_2 v_2+\varepsilon_1\varepsilon_2 v_{12} & & \\ & \swarrow & & \searrow & \\ x+\varepsilon_1 v_1 & & & & x+\varepsilon_2 v_2\,. \\ & \searrow & & \swarrow & \\ & & x & & \end{array} \tag{7.3}$$

(B) Injections. There is a diagram of inclusions of rings

$$\begin{array}{ccccc} & & \mathbb{K}[\varepsilon_1][\varepsilon_2] & & \\ & \nearrow & & \nwarrow & \\ \mathbb{K}[\varepsilon_1] & & & & \mathbb{K}[\varepsilon_2] \\ & \nwarrow & & \nearrow & \\ & & \mathbb{K} & & \end{array} \tag{7.4}$$

which corresponds to a diagram of zero sections of vector bundles

$$\begin{array}{ccccc} & & TTM & & \\ & \nearrow & & \nwarrow & \\ T_{\varepsilon_1}M & & & & T_{\varepsilon_2}M\,. \\ & \nwarrow & & \nearrow & \\ & & M & & \end{array} \tag{7.5}$$

(C) Ideal structure. In the following, we abbreviate $R := TT\mathbb{K}$. The following are inclusions of ideals:

$$\begin{array}{ccccc} & & \varepsilon_1 R = \varepsilon_1\mathbb{K}\oplus\varepsilon_1\varepsilon_2\mathbb{K} = \varepsilon_1\mathbb{K}[\varepsilon_2] & & \\ & \nearrow & & \searrow & \\ \varepsilon_1\varepsilon_2 R = \varepsilon_1\varepsilon_2\mathbb{K} & & & & R\,. \\ & \searrow & & \nearrow & \\ & & \varepsilon_2 R = \varepsilon_2\mathbb{K}\oplus\varepsilon_1\varepsilon_2\mathbb{K} = \varepsilon_2\mathbb{K}[\varepsilon_1] & & \end{array} \tag{7.6}$$

Here, also the composed inclusion $\varepsilon_1\varepsilon_2\mathbb{K}\subset R$ is an ideal. The three ideals from (7.6) are kernels of three projections appearing in the following three exact sequences of (non-unital) rings.

$$\begin{array}{ccccccccc} 0 & \to & \varepsilon_2 R & \to & R & \xrightarrow{p_1} & \mathbb{K}[\varepsilon_1] & \to & 0, \\ 0 & \to & \varepsilon_1 R & \to & R & \xrightarrow{p_2} & \mathbb{K}[\varepsilon_2] & \to & 0, \\ 0 & \to & \varepsilon_1\varepsilon_2 R & \to & R & \xrightarrow{p_1\times p_2} & \mathbb{K}[\varepsilon_1]\times\mathbb{K}[\varepsilon_2] & \to & 0. \end{array} \tag{7.7}$$

The first two sequences correspond to the two versions of the vector bundle projection $TTM \to TM$ that appear in (7.2), where the kernel describes just the fibers. The third sequence corresponds to a sequence of smooth maps

$$TM \to TTM \to TM \times_M TM. \tag{7.8}$$

The first map in (7.8) is the injection given in a chart by $x + \varepsilon v \mapsto x + \varepsilon_1 \varepsilon_2 v$ and the second map in (7.8) is the surjection given by

$$x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1 \varepsilon_2 v_{12} \mapsto (x + \varepsilon_1 v_1, x + \varepsilon_2 v_2);$$

it is just the vector bundle direct sum of the two projections $TTM \to TM$ from (7.7). The injection from (7.8) may intrinsically be defined as $\varepsilon_2 \circ T_{\varepsilon_1} z = \varepsilon_1 \circ T_{\varepsilon_2} z$, where $z : M \to TM$ is the zero section and $\varepsilon_i : TTM \to TTM$ is the almost dual structure corresponding to ε_i (cf. Section 4.7). The image of this map is called the *vertical bundle*,

$$VM := \varepsilon_1 \varepsilon_2 TM := \varepsilon_1 T_{\varepsilon_2} z(TM) = \varepsilon_2 T_{\varepsilon_1} z(TM); \tag{7.9}$$

it is a subbundle of TTM that carries the structure of a vector bundle over M isomorphic to TM.

(D) Automorphisms, endomorphisms. An important feature of $TT\mathbb{K}$ is the "exchange" or "flip automorphism" κ exchanging ε_1 and ε_2,

$$\kappa : TT\mathbb{K} \to TT\mathbb{K},$$
$$x_0 + x_1\varepsilon_1 + x_2\varepsilon_2 + x_{12}\varepsilon_1\varepsilon_2 \mapsto x_0 + x_2\varepsilon_1 + x_1\varepsilon_2 + x_{12}\varepsilon_1\varepsilon_2.$$

It induces a $\mathbb{K}$-linear automorphism of TTV and a canonical diffeomorphism (over $\mathbb{K}$) of TTM, given in a chart by

$$\begin{aligned} &\kappa : TTU \to TTU, \\ &x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12} \mapsto x + \varepsilon_1 v_2 + \varepsilon_2 v_1 + \varepsilon_1 \varepsilon_2 v_{12}. \end{aligned} \tag{7.10}$$

This map is chart independent: the transition functions of TTM are calculated in Section 4.7; they are given by $T^2\varphi_{ij}$, and because of the symmetry of $d^2\varphi_{ij}(x)$ (Schwarz' lemma), they commute with κ. Thus $\kappa : TTM \to TTM$ does not depend on the chart.

Recall from Section 6.6 that the ring $\mathbb{K}$ acts by endomorphisms on $T\mathbb{K}$:

$$l : \mathbb{K} \times T\mathbb{K} \to T\mathbb{K}, \quad (\lambda, a + \varepsilon b) \mapsto a + \varepsilon\lambda b$$

which corresponds to the action $l : \mathbb{K} \times TM \to TM$ by scalar multiplication in tangent spaces. By taking the tangent map Tl, we get two actions $T_{\varepsilon_i} l : T\mathbb{K} \times TT\mathbb{K} \to TT\mathbb{K}$, $i = 1, 2$, corresponding to two actions $T\mathbb{K} \times TTM \to TTM$ coming from the two $T\mathbb{K}$-vector bundle structures of TTM over TM. Both actions are interchanged by the flip κ. The combined action

$$l^{(2)} : T\mathbb{K} \times TTM \to TTM, \quad (\lambda, u) \mapsto T_{\varepsilon_1} l(\lambda, T_{\varepsilon_2} l(\lambda, u)) \tag{7.11}$$

commutes with κ. In a chart, for $r, s \in \mathbb{K}$ and $u = x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12}$,

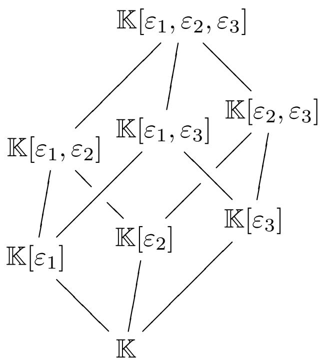
$$\begin{aligned} T_{\varepsilon_2} l(r + \varepsilon_2 s, u) &= x + \varepsilon_1 v_1 + \varepsilon_2 r v_2 + \varepsilon_1\varepsilon_2 (r v_{12} + s v_1), \\ l^{(2)}(r + \varepsilon_2, u) &= x + r\varepsilon_1 v_1 + r\varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 (r^2 v_{12} + r s v_1 + r s v_2). \end{aligned}$$

7.4. *Structure of $T^k\mathbb{K}$ and of T^kM.* The structure of the ring $T^k\mathbb{K} = \mathbb{K}[\varepsilon_1, \ldots, \varepsilon_k]$ can be analyzed in a similar way as we did above for $k = 2$. In the following, we focus on the case $k = 3$,

$$T^3\mathbb{K} = \mathbb{K} \oplus \varepsilon_1\mathbb{K} \oplus \varepsilon_2\mathbb{K} \oplus \varepsilon_3\mathbb{K} \oplus \varepsilon_1\varepsilon_2\mathbb{K} \oplus \varepsilon_1\varepsilon_3\mathbb{K} \oplus \varepsilon_2\varepsilon_3\mathbb{K} \oplus \varepsilon_1\varepsilon_2\varepsilon_3\mathbb{K},$$

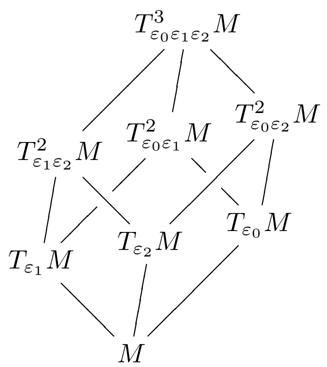
with the obvious relations.

(A) Projections. There is a commutative cube of ring projections:



(7.12)

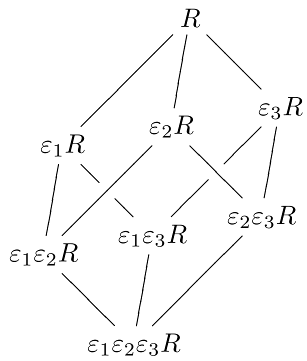
On the level of manifolds, the corresponding diagram of vector bundle projections will be written



(B) Injections. The diagrams written above can also be read, from the base to the top, as inclusions of subrings; on the level of manifolds, this corresponds to a diagram of zero sections of vector bundles. For general k, we have similar diagrams of projections and injections which could be represented by k-dimensional cubes.

(C) Ideal structure. There is also a cube of inclusions of ideals of the ring

$R = T^3\mathbb{K}$:



(7.13)

which on the level of manifolds corresponds to a cube of inclusions of various "vertical bundles"; but only the last and smallest of these vertical bundles, $\varepsilon_1\varepsilon_2\varepsilon_3 TM$, is a vector bundle over M. For general k, the first two exact sequences of (7.7) generalize, for $j = 1, \ldots, k$, to

$$0 \quad \to \quad \varepsilon_j R \quad \to \quad R \quad \stackrel{p_j}{\to} \quad \mathbb{K}[\varepsilon_1, \ldots, \widehat{\varepsilon}_j, \ldots, \varepsilon_k] \quad \to \quad 0 \tag{7.14}$$

where a "hat" means "omit this coefficient". The general form of the third sequence from (7.7) is, for any choice $1 \leq j_1 < \ldots < j_\ell \leq k$,

$$0 \quad \to \quad \bigcap_{i=1}^{\ell} \varepsilon_{j_i} R \quad \to \quad R \quad \stackrel{\prod p_{j_i}}{\longrightarrow} \prod_{i=1}^{\ell} \mathbb{K}[\varepsilon_1, \ldots, \widehat{\varepsilon}_{j_i}, \ldots, \varepsilon_k] \quad \to \quad 0. \tag{7.15}$$

In particular, for the "maximal choice" $j_i = i$, $i = 1, 2, \ldots, k$, we get

$$0 \quad \to \quad \varepsilon_1 \ldots \varepsilon_k R \quad \to \quad R \quad \to \quad \prod_{i=1}^{k} \mathbb{K}[\varepsilon_1, \ldots, \widehat{\varepsilon}_j, \ldots, \varepsilon_k] \quad \to \quad 0. \tag{7.16}$$

On the level of manifolds, (7.16) corresponds to the "most vertical" bundle

$$\varepsilon_1 \ldots \varepsilon_k TM \quad \to \quad T^k M \quad \to \quad \prod_{i=1}^{k} T^{k-1} M.$$

There are also various projections $T^k\mathbb{K} \to T^\ell\mathbb{K}$ for all $\ell = 0, \ldots, k-1$; their kernels are certain sums of the ideals considered so far. On the level of manifolds, this corresponds to an iterated fibration associated to the projections $T^k M \to T^\ell M$. In particular, for $\ell = 0$, we obtain the "augmentation ideal" of $T^k\mathbb{K}$ containing all linear combinations of the ε_i "without constant term". It corresponds to the fibers of $T^k M$ over M.

(D) Automorphisms, endomorphisms. According to Diagram (7.12), $T^3 M$ can be seen in three different ways as the second order tangent bundle of TM. Thus we have three different versions of the canonical flip κ (Section 7.3 (D)), corresponding to the canonical action of the three transpositions (12), (13) and (23) on $T^3 M$. The action of (12) may also be interpeted as $T_{\varepsilon_3}\kappa$ with $\kappa : TTM \to TTM$ as in 7.3 (D). Altogether this gives us a canonical action of the permutation group Σ_3 by $\mathbb{K}$-diffeomorphisms of $T^3 M$. In a chart, this action is given by permuting the symbols $\varepsilon_1, \varepsilon_2, \varepsilon_3$. Similarly, for general k we have an action of the permutation group Σ_k on $T^k M$.

The action $l : \mathbb{K} \times TM \to TM$ gives rise to an action $T^k l : T^k\mathbb{K} \times T^{k+1} M \to T^{k+1} M$ which is not invariant under the Σ_k-action, but as for $k = 2$ one can put these actions together to define some invariant action – see Section 20.4 for more details.

7.4. *Transition functions and chart representation of higher order tangent maps.* For a good understanding of the structure of $T^k M$ we will need to know the structure of the transition functions $T^k \varphi_{ij}$ of $T^k M$, and more generally, the structure of higher order tangent maps $T^k f$. For $k = 1$, recall that, for a smooth map $f : V \supset U \to W$ (which represents a smooth map $M \to N$ with respect to charts), we have

$$Tf(x + \varepsilon v) = f(x) + \varepsilon df(x)v. \tag{7.17}$$

Using the rule on partial derivatives (Lemma 1.6 (iv)), we may calculate $d(Tf)$ and hence TTf. One gets

$$\begin{aligned} T^2 f(x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12}) = f(x) + \varepsilon_1 df(x)v_1 + \varepsilon_2 df(x)v_2 + \\ \varepsilon_1\varepsilon_2\big(df(x)v_{12} + d^2 f(x)(v_1, v_2)\big). \end{aligned} \tag{7.18}$$

Iterating this procedure once more, we get

$$\begin{aligned} &T^3 f(x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_3 v_3 + \varepsilon_1\varepsilon_2 v_{12} + \varepsilon_3\varepsilon_1 v_{13} + \varepsilon_3\varepsilon_2 v_{23} + \varepsilon_1\varepsilon_2\varepsilon_3 v_{123}) \\ &= f(x) + \sum_{j=1,2,3} \varepsilon_j df(x)v_j + \sum_{1\le i<j\le 3} \varepsilon_i\varepsilon_j(df(x)v_{ij} + d^2 f(x)(v_i, v_j)) + \\ &\qquad \varepsilon_1\varepsilon_2\varepsilon_3 \cdot \big(df(x)v_{123} + d^2 f(x)(v_1, v_{23}) + d^2 f(x)(v_2, v_{13}) + \\ &\qquad\qquad d^2 f(x)(v_3, v_{12}) + d^3 f(x)(v_1, v_2, v_3)\big). \end{aligned}$$

In order to state the general result, we introduce a multi-index notation (cf. also Appendix MA, Sections MA.1 and MA.3). For multi-indices $\alpha, \beta \in I_k := \{0,1\}^k$ we define

$$\varepsilon^\alpha := \varepsilon_1^{\alpha_1} \cdots \varepsilon_k^{\alpha_k}, \qquad |\alpha| := \sum_i \alpha_i.$$

Moreover, we fix a total ordering on our index set I_k; for the moment, any ordering would do the job, but for later purposes, we agree to take the usual lexicographic ordering on $I_k = \{0,1\}^k$. Then a general element in the fiber over x can be written in the form $x + \sum_{\alpha>0} \varepsilon^\alpha v_\alpha$ with $v_\alpha \in V$, i.e., our chart domain is written in the form

$$T^k U = \{x + \sum_{\substack{\alpha\in I_k \\ \alpha>0}} \varepsilon^\alpha v_\alpha |\, x \in U, \forall \alpha :\, v_\alpha \in V\}.$$

A *partition* of $\alpha \in I_k$ is an ℓ-tuple $\Lambda = (\Lambda^1, \ldots, \Lambda^\ell)$ with $\Lambda^i \in I_k$ such that $0 < \Lambda^1 < \ldots < \Lambda^\ell$ and $\sum_i \Lambda^i = \alpha$; the integer $\ell = \ell(\Lambda)$ is called the *length* of the partition. The set of all partitions (of length ℓ) of α is denoted by $\mathcal{P}(\alpha)$ (resp. by $\mathcal{P}_\ell(\alpha)$).

Theorem 7.5. *The higher order tangent maps $T^k f : T^k U \to T^k W$ of a C^k-map $f : U \to W$ are expressed in terms of the usual higher differentials*

$d^j f : U \times V^j \to W$ *via the following formula:*

$$T^k f(x + \sum_{\alpha>0} \varepsilon^\alpha v_\alpha) = f(x) + \sum_{\substack{\alpha\in I_k \\ \alpha>0}} \varepsilon^\alpha \big(\sum_{\ell=1}^{|\alpha|} \sum_{\Lambda\in\mathcal{P}_\ell(\alpha)} d^\ell f(x)(v_{\Lambda^1},\ldots,v_{\Lambda^\ell})\big)$$
$$= f(x) + \sum_{j=1}^k \sum_{\substack{\alpha\in I \\ |\alpha|=j}} \varepsilon^\alpha \big(\sum_{\ell=1}^{j} \sum_{\substack{\Lambda^1+\ldots+\Lambda^\ell=\alpha \\ \Lambda^1<\ldots<\Lambda^\ell}} d^\ell f(x)(v_{\Lambda^1},\ldots,v_{\Lambda^\ell})\big).$$

Proof. The claim is proved by induction on k. For $k=1$, this is simply formula (7.17). Assume now that the formula holds for $k\in\mathbb{N}$. We know that the iterated tangent map $\widetilde{f} := T^k f : T^k U \to T^k W$ is smooth over $T^k\mathbb{K}$ and that $\widetilde{f}|_U$ coincides, via the zero sections, with f, i.e., $\widetilde{f}(x) = f(x) \in W$ for all $x\in U$. In the following proof we will use only these properties of $\widetilde{f}$. As in the proof of Theorem 6.8 (2), a second order Taylor expansion gives, for ε being one the ε_i,

$$\widetilde{f}(x+\varepsilon v) = f(x) + \varepsilon d\widetilde{f}(x)v + \varepsilon^2 a_2(x,v) + \varepsilon^2 R_2(x,v,\varepsilon) = f(x) + \varepsilon\partial_v f(x),$$

which, of course, is again (7.17). Note that we used $d\widetilde{f}(x)v = \partial_v\widetilde{f}(x) = \partial_v f(x) = df(x)v$, which holds since $x,v\in V$ and $\widetilde{f}|_U = f$. Now we are going to repeat this argument, using that, if $\widetilde{f}$ is C^3 over $TT\mathbb{K}$, then all maps $\partial_u f$, $u\in TTU$, are C^2 over $TT\mathbb{K}$ and hence also over $T\mathbb{K}$:

$$\begin{aligned}\widetilde{f}(x+\varepsilon_1 v_1+\varepsilon_2 v_2 +\varepsilon_1\varepsilon_2 v_{12}) &= \widetilde{f}(x+\varepsilon_1 v_1+\varepsilon_2(v_2+\varepsilon_1 v_{12}))\\ &= \widetilde{f}(x+\varepsilon_1 v_1) + \varepsilon_2(\partial_{v_2+\varepsilon_1 v_{12}}\widetilde{f})(x+\varepsilon_1 v_1)\\ &= f(x) + \varepsilon_1\partial_{v_1} f(x) + \varepsilon_2\partial_{v_2+\varepsilon_1 v_{12}}\widetilde{f}(x) + \varepsilon_1\varepsilon_2\partial_{v_1}\partial_{v_2+\varepsilon_1 v_{12}}\widetilde{f}(x)\\ &= f(x) + \varepsilon_1\partial_{v_1} f(x) + \varepsilon_2\partial_{v_2} f(x) + \varepsilon_1\varepsilon_2(\partial_{v_{12}} + \partial_{v_1}\partial_{v_2})f(x).\end{aligned}$$

The proves again (7.18), and again the resulting expression depends only on $f|_U$. Now, for the general induction step, we let $u = x+v = x+\sum_{\alpha\in I_{k+1}}\varepsilon^\alpha v_\alpha \in T^{k+1}U = T(T^k U)$ and decompose $u = x + v' + \varepsilon_{k+1}v''$ correspondingly. We let also $\omega = (0,\ldots,0,1)\in I_{k+1}$ and identify I_k with the subset of $\alpha\in I_{k+1}$ such that $\alpha_{k+1}=0$. Then $\varepsilon^\omega = \varepsilon_{k+1}$, and I_{k+1} is the disjoint union of I_k and $I_k+\omega$. By induction,

$$\widetilde{f}(x+v) = \widetilde{f}\big(x+\varepsilon^\omega v_\omega + \sum_{\substack{\alpha\in I_k \\ \alpha>0}} \varepsilon^\alpha(v_\alpha + \varepsilon^\omega v_{\alpha+\omega})\big)$$
$$= \widetilde{f}(x+\varepsilon^\omega v_\omega) + \sum_{\substack{\alpha\in I_k \\ \alpha>0}} \varepsilon^\alpha \sum_{\ell=1}^{|\alpha|}\sum_{\Lambda\in\mathcal{P}_\ell(\alpha)} (\partial_{v_{\Lambda^1}+\varepsilon^\omega v_{\Lambda^1+\omega}}\cdots\partial_{v_{\Lambda^\ell}+\varepsilon^\omega v_{\Lambda^\ell+\omega}}\widetilde{f})(x+\varepsilon^\omega v_\omega)$$
$$= f(x) + \varepsilon^\omega\partial_{v_\omega} f(x) + \sum_{\substack{\alpha\in I_k \\ \alpha>0}} \varepsilon^\alpha \sum_{\ell=1}^{|\alpha|}\sum_{\Lambda\in\mathcal{P}_\ell(\alpha)} (\partial_{v_{\Lambda^1}+\varepsilon^\omega v_{\Lambda^1+\omega}}\cdots\partial_{v_{\Lambda^\ell}+\varepsilon^\omega v_{\Lambda^\ell+\omega}}\widetilde{f})(x)$$
$$+ \varepsilon^\omega\big(\partial_{v_\omega}\sum_{\substack{\alpha\in I_k \\ \alpha>0}} \varepsilon^\alpha \sum_{\ell=1}^{|\alpha|}\sum_{\Lambda\in\mathcal{P}_\ell(\alpha)} (\partial_{v_{\Lambda^1}+\varepsilon^\omega v_{\Lambda^1+\omega}}\cdots\partial_{v_{\Lambda^\ell}+\varepsilon^\omega v_{\Lambda^\ell+\omega}}\widetilde{f})\big)(x).$$

Next, we use linearity of the map $y \mapsto \partial_y f$ in order to expand the whole expression. Fortunately, due to the relation $(\varepsilon^\omega)^2 = \varepsilon_{k+1}^2 = 0$, many terms vanish, and the remaining sum is a sum over partitions Ω of elements $\beta \in I_{k+1}$ of the following three types:

(a) Ω is a partition Λ of some $\alpha \in I_k$,

(b) Ω is obtained from a partition Λ of $\alpha \in I_k$ by adjoining ω as last component (then $\ell(\Omega) = \ell(\Lambda) + 1$),

(c) Ω is obtained from a partition Λ of $\alpha \in I_k$ by adding ω to some component of Λ (then $\ell(\Omega) = \ell(\Lambda)$).

But it is clear that every partition Ω of any element $\beta \in I_{k+1}$ is exactly of one of the preceding three types, and therefore $T^{k+1} f(x+v)$ is of the form given in the claim. ■

As a consequence of the theorem, one can get another proof of the existence of a canonical action of the permutation group Σ_k on $T^k M$: in a chart, Σ_k acts via its canonical action on $\{0,1\}^k$, and the chart formula shows that T^k commutes with this action because of the symmetry of the higher differentials. – Another consequence of the preceding proof is the following uniqueness result on the "scalar extension of f by $T^k f$":

Theorem 7.6. *Assume $f : M \to N$ is smooth over $\mathbb{K}$ and $\widetilde{f} : T^k M \to T^k N$ is a $T^k\mathbb{K}$-smooth extension of f in the sense that $\widetilde{f} \circ z_{T^k M} = z_{T^k N} \circ f$:*

$$\begin{array}{ccc} T^k M & \stackrel{\widetilde{f}}{\to} & T^k N \\ \uparrow & & \uparrow \\ M & \stackrel{f}{\to} & N \end{array}$$

Then $\widetilde{f} = T^k f$.

Proof. The preceding proof has shown that $\widetilde{f}$ agrees with $T^k f$ in any chart representation, whence $\widetilde{f} = T^k f$. ■

8. Scalar extensions. III: Jet functor and truncated polynomial rings

8.1. *Jet rings and truncated polynomial rings.* Recall the canonical action of the permutation group Σ_k by automorphisms of the higher order tangent ring $T^k\mathbb{K} = \mathbb{K}[\varepsilon_1, \ldots, \varepsilon_k]$. The subring fixed under this action will be called the *(k-th order) jet ring of* $\mathbb{K}$, denoted by $J^k\mathbb{K}$. For $k = 0, 1$ we have $J^0\mathbb{K} = \mathbb{K}$, $J^1\mathbb{K} = T\mathbb{K} = \mathbb{K}[\varepsilon] = \mathbb{K}[x]/(x^2)$, and for $k = 2, 3$,

$$J^2\mathbb{K} = \{a_0 + \varepsilon_1 a_1 + \varepsilon_2 a_1 + \varepsilon_1\varepsilon_2 a_2 | \, a_0, a_1, a_2 \in \mathbb{K}\} = \mathbb{K} \oplus \mathbb{K}(\varepsilon_1 + \varepsilon_2) \oplus \mathbb{K}\varepsilon_1\varepsilon_2,$$
$$J^3\mathbb{K} = \mathbb{K} \oplus \mathbb{K}(\varepsilon_1 + \varepsilon_2 + \varepsilon_3) \oplus \mathbb{K}(\varepsilon_1\varepsilon_2 + \varepsilon_2\varepsilon_3 + \varepsilon_1\varepsilon_3) \oplus \mathbb{K}\varepsilon_1\varepsilon_2\varepsilon_3.$$

For general k, we introduce the notation

$$\delta := \delta_k := \sum_{i=1}^k \varepsilon_i, \qquad \delta^{(2)} := \sum_{1 \leq i < j \leq k} \varepsilon_i \varepsilon_j, \qquad \delta^{(j)} := \delta_k^{(j)} := \sum_{1 \leq i_1 < \ldots < i_j \leq k} \varepsilon_{i_1} \cdots \varepsilon_{i_j} \tag{8.1}$$

(the j-th elementary symmetric polynomial in the variables $\varepsilon_1, \ldots, \varepsilon_k$) which defines a basis of the $\mathbb{K}$-module $J^k\mathbb{K}$:

$$J^k\mathbb{K} = \mathbb{K} \oplus \mathbb{K}\delta \oplus \mathbb{K}\delta^{(2)} \oplus \ldots \oplus \mathbb{K}\delta^{(k)}. \tag{8.2}$$

An element of $T^k\mathbb{K}$ and of $J^k\mathbb{K}$ is invertible iff its lowest coefficient is invertible; therefore $J^k\mathbb{K}$ has a dense unit group if $\mathbb{K}$ has a dense unit group. The powers of δ are

$$\delta^2 = (\sum_i \varepsilon_i)^2 = \sum_{i,j} \varepsilon_i\varepsilon_j = 2\sum_{i<j} \varepsilon_i\varepsilon_j = 2\delta^{(2)}, \ldots$$
$$\delta^j = j!\delta^{(j)}, \quad \ldots \quad , \delta^k = k!\varepsilon_1 \cdots \varepsilon_k, \qquad \delta^{k+1} = 0.$$

More generally, we have the multiplication rule

$$\delta^{(i)} \cdot \delta^{(j)} = \binom{i+j}{i} \delta^{(i+j)}.$$

If the integers $2, \ldots, k$ are invertible in $\mathbb{K}$, we can take $1, \delta, \delta^2, \ldots, \delta^k$ as a new $\mathbb{K}$-basis in $J^k\mathbb{K}$:

$$J^k\mathbb{K} = \mathbb{K} \oplus \mathbb{K}\delta \oplus \mathbb{K}\delta^2 \oplus \mathbb{K}\delta^3 \ldots \oplus \mathbb{K}\delta^k.$$

For any $\mathbb{K}$, the ring

$$\mathbb{K}[\delta_k] := \mathbb{K}[x]/(x^{k+1})$$

is called a *truncated polynomial ring (over* $\mathbb{K}$*).* A $\mathbb{K}$-basis of $\mathbb{K}[x]/(x^{k+1})$ is given by the classes of the polynomials $1, x, x^2, \ldots, x^k$. We denote these classes by $\delta^0, \delta, \delta^2, \ldots, \delta^k$; then we have the relation $\delta^{k+1} = 0$, and elements are multiplied by the rule

$$\sum_{i=0}^k a_i\delta^i \cdot \sum_{j=0}^k b_j\delta^j = \sum_{\ell=0}^k (\sum_{i+j=\ell} a_i b_j)\delta^\ell.$$

These relations show that, if $2, \ldots, k$ are invertible in $\mathbb{K}$, then the jet ring $J^k\mathbb{K}$ and the truncated polynomial ring $\mathbb{K}[\delta_k]$ are isomorphic. For general $\mathbb{K}$, the unit element of $\mathbb{K}[\delta_k]$ is $1 = \delta^0$, and, given $a = \sum_i a_i \delta^i$, the condition $ab = 1$ for $b = \sum_j b_j \delta^j$ is a triangular system of linear equations whose matrix has a diagonal of coefficients all equal to a_0; this system can be solved if and only if a_0 is invertible. In other words, the unit group of the ring $J^k\mathbb{K}$ is the set of truncated polynomials having an invertible lowest coefficient. Clearly, this group is dense in $J^k\mathbb{K}$ if the unit group of $\mathbb{K}$ is dense in $\mathbb{K}$.

Theorem 8.2. *If M is a manifold over $\mathbb{K}$, we denote by*

$$J^k M := (T^k M)^{\Sigma_k}$$

the fixed point set of the canonical action of the permutation group Σ_k on $T^k M$. Then $J^k M$ is a subbundle of the bundle $T^k M$ over M, and it is a smooth manifold over the k-th order jet ring $J^k\mathbb{K}$. If $f : M \to N$ is smooth over $\mathbb{K}$, then the restriction $J^k f : J^k M \to J^k N$ is well-defined and is smooth over $J^k\mathbb{K}$. Summing up, J^k can be seen as the functor of scalar extension from $\mathbb{K}$ to $J^k\mathbb{K}$. If the integers $2, \ldots, k$ are invertible in $\mathbb{K}$, then the corresponding statements are true with respect to the truncated polynomial ring $\mathbb{K}[x]/(x^{k+1})$ instead of $J^k\mathbb{K}$.

Proof. First of all we prove that $J^k M$ is a subbundle of $T^k M$. In a bundle chart over $U \subset M$, the fiber over x is represented by

$$(J^k U)_x = \{x + \sum_{i=1}^{k} \varepsilon_i v_1 + \sum_{i<j} \varepsilon_i \varepsilon_j v_2 + \ldots + \varepsilon_1 \cdots \varepsilon_k v_k \mid v_1, \ldots, v_k \in V\}. \quad (8.3)$$

This is a $\mathbb{K}$-submodule and corresponds to the chart representation of a submanifold provided we can find a complementary submodule. In case $k = 2$, this amounts to find a complementary submodule of the diagonal in $V \times V$: for instance, we may take $0 \times V$ (if 2 in invertible in $\mathbb{K}$, we may also take the antidiagonal); similarly for general k. Thus $J^k M$ is a submanifold of $T^k M$ in the sense of 2.3; from (8.3) it is then seen to be a subbundle. If $f : M \to M$ is a smooth map, then $T^k f : T^k M \to T^k N$ is, due to the symmetry of higher differentials, compatible with the action of the permutation group Σ_k, and hence gives rise to a well-defined and $\mathbb{K}$-smooth map

$$J^k f : J^k M \to J^k N, \quad u \mapsto T^k f(u)$$

such that we have the functorial rules $J^k(f \circ g) = J^k f \circ J^k g$, $J^k \mathrm{id}_M = \mathrm{id}_{J^k M}$. Smoothness of $J^k f$ over $J^k\mathbb{K}$ follows simply by restriction: we know that $T^k f : T^k U \to T^k W$ is smooth over $T^k\mathbb{K}$ (Theorem 7.2); we write Condition (1.1) for differentiability over $T^k\mathbb{K}$ and restrict to Σ_k-invariants and thus get Condition (1.1) for differentiability ofer $J^k\mathbb{K}$. By induction, we see that $J^k f$ is of class C^ℓ over $J^k\mathbb{K}$ for all ℓ.

Finally, the last statement follows from the isomorphism of topological rings $J^k\mathbb{K} \cong \mathbb{K}[x]/(x^{k+1})$ explained in Section 8.1. ■

The bundle $J^k M$ over M is called the *k-th order jet bundle of M* – see Section 8.8 below for a discussion with more conventional definitions of jets.

Theorem 8.3. *If G is a Lie group over $\mathbb{K}$, then, for $k \in \mathbb{N}$, $T^k G$, resp. $J^k G$ are Lie groups over the rings $T^k\mathbb{K}$, resp. over $J^k\mathbb{K}$, and the corresponding Lie algebra is the scalar extension of the Lie algebra $\mathfrak{g}$ of G by the ring $T^k\mathbb{K}$, resp. by $J^k\mathbb{K}$. Similarly, if M is a symmetric space over $\mathbb{K}$ with base point o, then, for $k \in \mathbb{N}$, $T^k M$, resp. $J^k M$, are symmetric spaces over $T^k\mathbb{K}$, resp. over $J^k\mathbb{K}$, with base point being the origin in the fiber over o. The associated Lie triple system is the scalar extension of the Lie triple system associated to (M, o) by the ring $T^k\mathbb{K}$, resp. by $J^k\mathbb{K}$. The various natural projections and inclusions are homomorphisms of Lie groups, resp. of symmetric spaces.*

Proof. This is the generalization of Theorem 6.5, and its version for $T^k\mathbb{K}$ follows by induction from that theorem. The version for $J^k\mathbb{K}$ then follows by restriction to Σ_k-invariants. ∎

8.4. *The structure of $J^k\mathbb{K}$ and of $J^k M$.* Not all properties of the ring $T^k\mathbb{K}$ are invariant under the permutation group, and hence the structure of $J^k\mathbb{K}$ is somewhat poorer than the structure of $T^k\mathbb{K}$.

(A) Projections. The various exact sequences $0 \to \varepsilon T^{k-1}\mathbb{K} \to T^k\mathbb{K} \to T^{k-1}\mathbb{K} \to 0$ which are all of the form $0 \to \varepsilon R \to R[\varepsilon] \to R \to 0$ (where ε is one of the ε_i) induce, by restriction to invariants under permutations, an exact sequence

$$0 \to \delta^{(k)}\mathbb{K} \to J^k\mathbb{K} \xrightarrow{p} J^{k-1}\mathbb{K} \to 0, \qquad p(\sum_{i=0}^{k} \delta^{(i)} v_i) = \sum_{i=0}^{k-1} \delta^{(i)} v_i. \tag{8.4}$$

More generally, for all $\ell \leq k$ we have projections

$$\pi_{\ell,k} : J^k\mathbb{K} \to J^\ell\mathbb{K}, \quad \sum_{i=0}^{k} \delta^{(i)} v_i \mapsto \sum_{i=0}^{\ell} \delta^{(i)} v_i \tag{8.5}$$

which all come from the various projections $T^k\mathbb{K} \to T^\ell\mathbb{K}$. On the space level, $\pi_{k,k+1}$ corresponds to the *jet bundle projection*

$$\begin{array}{ccc} J^{k+1}M & \subset & T^{k+1}M \\ \downarrow & & \downarrow \\ J^k M & \subset & T^k M \end{array} \tag{8.6}$$

and the projections $\pi_{\ell,k} : J^k M \to J^\ell M$ are given in a chart by

$$J^k M \to J^\ell M, \quad x + \sum_{j=1}^{k} \delta^{(j)} v_j \mapsto x + \sum_{j=1}^{\ell} \delta^{(j)} v_j. \tag{8.7}$$

(B) Non-existence of injections. The "zero-sections" $z_j : \varepsilon_j T^k M \to T^{k+1} M$ of the vector bundles $p_j : T^{k+1}M \to T^k M$ are, for $k \geq 1$, not compatible with the action of the symmetric group: for instance, in a chart we have

$$z_{k+1}(x + \varepsilon_1 v_1 + \ldots + \varepsilon_k v_k) = x + \varepsilon_1 v_1 + \ldots + \varepsilon_k v_k + 0,$$

hence the image of z_{k+1} does not belong to $J^{k+1}M$. Therefore, when restricting the vector bundle $T^{k+1}M \to T^kM$ to invariants under the symmetric group, we obtain a "vector bundle without zero section", that is, an *affine bundle* $J^{k+1}M \to J^kM$. Thus the fiber over $u \in J^kM$,

$$(V^{k+1,1}M)_u = (J^{k+1}M)_{x+\sum_{j=1}^k \delta^{(j)}v_j} = \{x + \sum_{j=1}^{k+1} \delta^{(j)}v_j \,|\, v_{k+1} \in V\} \tag{8.8}$$

carries a canonical structure of an affine space over $\mathbb{K}$, but not of a $\mathbb{K}$-module.

(C) Ideal structure. The kernel of the projection $\mathrm{pr}_{k,\ell} : J^k\mathbb{K} \to J^\ell\mathbb{K}$,

$$J_{k,\ell} = \{ \sum_{i=\ell+1}^{k} \delta^{(i)} r_i | \forall i : \, r_i \in \mathbb{K}\}, \tag{8.9}$$

is an ideal of $J^k\mathbb{K}$, and instead of a cube of ideals as for $T^k\mathbb{K}$, now we have a chain of ideals. For $\ell = 0$ we get the "augmentation ideal", i.e., the set of elements having lowest coefficient equal to zero; if $\mathbb{K}$ is a field, this is the unique maximal ideal of $J^k\mathbb{K}$. For $\ell = k-1$ we get the ideal $\mathbb{K}\delta^{(k)}$ which, on the space level, corresponds to fibers of the affine bundle $J^kM \to J^{k-1}M$.

(D) Automorphisms, endomorphisms. The only automorphisms or endomorphisms from $T^k\mathbb{K}$ that suvive in a non-trivial way come from the action (which has been defined in Section 7.3 (D)) $l^{(k)} : T^k\mathbb{K} \times T^{k+1}M \to T^{k+1}M$. In particular, there is a well-defined action $\mathbb{K} \times J^kM \to J^kM$. In a chart, it is given by

$$r.(x + \sum_{i=1}^k \delta^{(i)}v_i) = x + \sum_{i=1}^k r^i \delta^{(i)} v_i. \tag{8.10}$$

8.5. *Chart representation of* J^kf. Let $f : M \to N$ be a smooth map. Using charts, J^kM and J^kN are represented in the form (8.3), and we may ask for the chart expression of J^kf. For $k = 2$, letting $v_1 = v_2 =: w$, we get from (7.18):

$$J^2f(x + \delta v + \delta^{(2)}w) = f(x) + \delta df(x)v + \delta^{(2)}\big(df(x)w + d^2f(x)(v,v)\big), \tag{8.11}$$

which implies, with $\delta^2 = 2\delta^{(2)}$, if 2 is invertible in $\mathbb{K}$,

$$J^2f(x + \delta v_1 + \delta^2 v_2) = f(x) + \delta df(x)v_1 + \delta^2\big(df(x)v_2 + \frac{1}{2}d^2f(x)(v_1,v_1)\big).$$

For $k = 3$ we get by restriction from the formula for T^3f

$$\begin{aligned} J^3f(x + \delta v_1 + &\delta^{(2)}v_2 + \delta^{(3)}v_3) = \\ &f(x) + \delta df(x)v_1 + \delta^{(2)}\big(df(x)v_2 + d^2f(x)(v_1,v_1)\big) + \\ &\qquad \delta^{(3)}\big(df(x)v_3 + 3d^2f(x)(v_1,v_2) + d^3f(x)(v_1,v_1,v_1)\big). \end{aligned}$$

In particular, if 3 is invertible in $\mathbb{K}$, we have

$$J^3f(x + \delta v) = f(x) + \delta df(x)v + \frac{1}{2}\delta^2\, df(x)v + \frac{1}{3!}\delta^3\, d^3f(x)(v,v,v).$$

Theorem 8.6. *If $f : U \to W$ is smooth over $\mathbb{K}$, then $J^k f : J^k U \to J^k W$ is described in terms of the usual higher order differentials $d^\ell f : U \times V^\ell \to W$ by*

$$J^k f(x + \sum_{j=1}^k \delta^{(j)} v_j) = f(x) + \sum_{j=1}^k \delta^{(j)} (\sum_{\ell=1}^j \sum_{\substack{i_1+\ldots+i_\ell=j \\ i_1 \leq \cdots \leq i_\ell}} C^j_{i_1,\ldots,i_\ell} d^\ell f(x)(v_{i_1}, \ldots, v_{i_\ell})),$$

where $C^j_{i_1,\ldots,i_\ell}$ is the number of decompositions of a set of j elements into ℓ disjoint subsets containing $i_1, i_2, \ldots, i_\ell$ elements. If the integers are invertible in $\mathbb{K}$, then this can also be written

$$J^k f(x + \sum_{j=1}^k \delta^{(j)} v_j) = f(x) + \sum_{j=1}^k \delta^{(j)} (\sum_{\ell=1}^j \sum_{i_1+\ldots+i_\ell=j} \frac{j!}{i_1! \ldots i_\ell! \ell!} d^\ell f(x)(v_{i_1}, \ldots, v_{i_\ell})).$$

In particular,

$$\begin{aligned} J^k f(x + \delta v) &= f(x) + \sum_{j=1}^k \delta^{(j)} df(x)(v, \ldots, v) \\ &= f(x) + \sum_{j=1}^k \frac{1}{j!} \delta^j df(x)(v, \ldots, v), \end{aligned}$$

the latter provided the integers are invertible in $\mathbb{K}$.

Proof. We restrict the formula from Theorem 7.5 to Σ_k-invariants: two multi-indices α, $\beta \in \{0,1\}^k$ are conjugate under Σ_k if and only if $|\alpha| = |\beta|$; therefore we have to take the special case of Theorem 7.5 where, for all α with $|\alpha| = j$, the arguments v_α are equal to a given element $v_j \in V$. Recall also that, by definition, $\delta^{(j)} = \sum_{|\alpha|=j} \varepsilon^\alpha$. Then the claim follows by counting the number of terms that are equal.

Note that, by definition, the numbers $C^j_{i_1,\ldots,i_\ell}$ are integers, and hence the claim is valid in arbitrary characteristic. See [Mac79, p. 22] for another formula and some remarks on these combinatorial coefficients. They are partition numbers, and there is no "explicit formula" for these constants. But the index set of the sum in the first formula of the theorem also is a set of partitions, and these two combinatorial difficulties cancel out if we take the sum over *all* decompositions $i_1 + \ldots + i_\ell = j$ (with repetitions!) – instead of counting partitions of $\{1, \ldots, j\}$ we count ℓ-tuples of subsets, the m-th set having i_m elements (so the number of such ℓ-tuples is $\frac{j!}{i_1! \cdots i_\ell!}$), and then divide by $\ell!$ (provided the integers are invertible in $\mathbb{K}$) which gives the second formula of the claim. ■

One should note that for practical computations the first formula of the theorem is much more efficient since it has less terms.

8.7. *The composition rule for higher differentials.* Assume f and g are smooth mappings of $\mathbb{K}$-modules such that the composition $g \circ f$ is defined on a non-empty open set. As an interesting consequence of the explicit formulae from Theorem 8.6 for the k-jets of f and g we get the following explicit formula for the higher differentials of $g \circ f$ in terms of those of f and g:

$$
\begin{aligned}
& d^j(g \circ f)(x)(v, \ldots, v) \\
&\quad = \sum_{\ell=1}^{j} \sum_{\substack{i_1+\ldots+i_\ell=j \\ i_1 \leq \ldots \leq i_\ell}} C^j_{i_1,\ldots,i_\ell} d^\ell g(f(x)) \Big(d^{i_1} f(x)(v_{i_1}), \ldots, d^{i_\ell} f(x)(v_{i_\ell}) \Big) \\
&\quad = \sum_{\ell=1}^{j} \sum_{i_1+\ldots+i_\ell=j} \frac{j!}{i_1! \ldots i_\ell! \ell!} d^\ell g(f(x)) \Big(d^{i_1} f(x)(v_{i_1}), \ldots, d^{i_\ell} f(x)(v_{i_\ell}) \Big),
\end{aligned}
$$

where for simplicity we wrote $d^{i_m} f(x)(v_{i_m})$ instead of $d^{i_m} f(x)(v, \ldots, v)$. In order to prove this formula, we write the explicit formulae for both sides of the equality $J^k(g \circ f)(x + \delta v) = J^k g(J^k f(x + \delta v))$ and compare terms having coefficient $\delta^{(j)}$. The second version of the composition rule (8.19) (in the finite-dimensional real case) can be found in [Chap03, p. 172, "formule de Faa de Bruno"] and [Ho01, Satz 60.16]; the proofs given there are different, using Taylor expansions in the classical sense. See also [Mac79, p. 23, Ex. 12] for the one-dimensional version of this formula.

8.8. *Relation with classical definitions of jets.* Usually, jets are defined via an equivalence relation (cf., e.g., [KMS93, Ch. IV] or [ALV91, p. 132]): one says that two maps $f, g : M \to N$ *have the same k-jet at $x \in M$* if, in a chart, all derivatives $d^j f(x)$ and $d^j g(x)$, $j = 0, 1, \ldots, k$, coincide. Then we write $f \sim^k_x g$. We have the following relation with our concepts of $T^k f(x)$ and $J^k f(x)$:

$$
f \sim^k_x g \quad \Leftrightarrow \quad (T^k f)_x = (T^k g)_x \quad \Rightarrow \quad (J^k f)_x = (J^k g)_x,
$$

and the last implication becomes an equivalence if the integers are invertible in $\mathbb{K}$. In fact, it is clear that the first two properties are equivalent since Theorem 7.5 permits to calculate $T^k f(x)$ from the higher order derivatives and vice versa. If $J^k f(x) = J^k g(x)$, it follows from Theorem 8.5 that we can recover the values of $d^j f(x)(v, \ldots, v)$ (for $j = 0, 1, \ldots, k$), and thus, if $2, \ldots, k$ are invertible in $\mathbb{K}$, it follows by polarisation that the differentials up to order k at x agree, i.e. $T^k f(x) = T^k g(x)$ and $f \sim^k_x g$. In particular, in the real case our definition coincides with the classical one. (For the case $k = 2$, cf. also [Bes78, p. 20].) Another possible interpretation of the bundles $J^k M$ is via k-jets of *curves*: if $\alpha : I \to M$ is a curve defined on an open neighborhood I of 0 in $\mathbb{K}$ such that $\alpha(0) = x$, then $(J^k \alpha)_0(\delta 1)$ is an element of $(J^k M)_x$, and if the integers are invertible in $\mathbb{K}$, then every element $u \in (J^k M)_x$ is obtained in this way: in fact, writing, in a chart, $u = u_0 + \delta u_1 + \ldots + \delta^{(k)} u_k \in (J^k M)_x$, we let $\alpha(t) = \sum_{\ell=0}^{k} \frac{1}{\ell!} t^\ell u_\ell$. According to Theorem 8.6, we then have $J^k \alpha(0 + \delta 1) = u$. Summing up, in case of characteristic zero, both functors T^k and J^k may then be seen as "faithful representations of the abstract jet functor", and hence are equivalent. However, in practice there are important differences:

(a) At a first glance, the functor J^k seems preferable, since we get rid of the big amount of redundant information contained in T^k: the number of variables is reduced from 2^k to k.

(b) For a fixed pair $k \geq \ell$, there is just *one* jet projection $\pi_{\ell,k} : J^k M \to J^\ell M$, and thus $J, J^2, J^3, \ldots$ forms a projective system so that we can pass to the projective limit J^∞ (see Chapter 32). The ring $J^\infty\mathbb{K}$ is essentially the ring of *formal power series over* $\mathbb{K}$. For T^k, there is no nice projective limit because of the more complicated cube structure of the projections.

(c) On the other hand, the functor T^k behaves well with respect to induction procedures because $T^{k+1} = T \circ T^k$, whereas J^{k+1} and $J^1 \circ J^k$ yield quite different results.

(d) The non-existence of zero-sections makes the functors J^k conceptually more difficult than the functors T^k.

(e) Finally, in positive characteristic, T^k may contain strictly more information than J^k.

For all of these reasons, in the sequel we will work, as long as we can, with T^k, and restrict only at the very last stage to Σ_k-invariants, before we take the projective limit – see Chapter 32.

III. Second order differential geometry

This part, although strongly motivated by Section 7.3 on the second order tangent bundle TTM, is independent of Part II. In particular, it can be read without having to bother about manifolds over rings, for readers who prefer fields to rings. Then notation like $V \oplus \varepsilon V$ or $V \oplus \delta V$ should simply be interpreted as another way of writing $V \oplus V$, with a formal label added in order to distinguish both copies of V.

9. The structure of the tangent bundle of a vector bundle

9.1. *Transition functions of the tangent bundle of a vector bundle.* Recall from Section 4.7 the transition maps for TTM. Let us make a similar calculation for TM replaced by a general vector bundle $p : F \to M$. In order to keep notation simple, we use similar conventions as in [La99]: a bundle chart domain of F is denoted by $U \times W \subset V \times W$ and a bundle chart domain of TF by $T(U \times W) = U \times W \times V \times W \subset V \times W \times V \times W$ with typical element denoted by $(x, w; x', w')$ or $x + w + \varepsilon(x' + w')$. The transition maps of F are, with the notation of Section 3.3,

$$\widetilde{g}_{ij}(x, w) = (\varphi_{ij}(x), g_{ij}(x)w) = (\varphi_{ij}(x), g_{ij}(x, w)).$$

Using the rule on partial derivatives (Lemma 1.6 (iv)), it follows that the transition maps of TF are

$$\begin{aligned} T\widetilde{g}_{ij}(x, w; x', w') &= (\widetilde{g}_{ij}(x, w), d(\widetilde{g}_{ij})(x, w)(x', w')) \\ &= (\varphi_{ij}(x), g_{ij}(x)w, d\varphi_{ij}(x)x', g_{ij}(x)w' + d_1(g_{ij})(x, w)x') \end{aligned} \tag{9.1}$$

which may also be written

$$\begin{aligned} T\widetilde{g}_{ij}(x + w + \varepsilon(x' + w')) = \varphi_{ij}(x) + g_{ij}(x)w + \\ \varepsilon(d\varphi_{ij}(x)x' + g_{ij}(x)w' + d_1(g_{ij})(x, w)x'). \end{aligned}$$

9.2. *TF as a fiber bundle over M.* Let $p : F \to M$ be as above. Denote by $Tp : TF \to TM$ its tangent map and by $\pi : TF \to F$ the canonical projection. In bundle charts, these projections are given by

$$Tp(x + w + \varepsilon(x' + w')) = x + \varepsilon x', \qquad \pi(x + w + \varepsilon(x' + w')) = x + w.$$

The projections $p \circ \pi$ and $\pi_{TM} \circ Tp : TF \to M$ agree: we have the following commutative diagram of vector bundle projections:

$$\begin{array}{ccccc} & & TF & & \\ & \swarrow & & \searrow & \\ F & & & & TM \\ & \searrow & & \swarrow & \\ & & M & & \end{array}$$

9.3. *TF as a bilinear bundle.* The fiber $(TF)_x$ over $x \in M$ does not carry a canonical $\mathbb{K}$-module structure: in a chart, $(TF)_x$ is represented by $E := W \times V \times W$, and the transition maps, calculated above, are *non-linear* in (w, x', w') (the last term in (9.1) depends bilinearly on (w, x') and not linearly). Hence $TF \to M$ *is not a vector bundle* – so the question arises: what is it then? Formula (9.1) contains the answer: for fixed x, the transition functions can be written in the form

$$f(w, x', w') = (h_1(w), h_2(x'), h_3(w' + b(w, x'))), \tag{9.2}$$

with (continuous) invertible linear maps

$$h_1 = h_3 = g_{ij}(x) \in \mathrm{Gl}(W), \quad h_2 = d\varphi_{ij}(x) \in \mathrm{Gl}(V),$$

and a (continuous) bilinear map

$$b := b_x : W \times V \to W, \quad (w, x') \mapsto g_{ij}(x)^{-1}\,(d_1(g_{ij})(x, w)x'). \tag{9.3}$$

In the terminology of Appendix BA, Section BA.2, this means that the transition functions for fixed $x \in M$ belong to the *general bilinear group* $\mathrm{Gm}^{0,2}(E)$ of $E = V_1 \times V_2 \times V_3$ (where $V_1 = V_3 = W$, $V_2 = V$). Every pair of bundle charts (say, corresponding to indices $i, j \in I$) defined at x identifies the fiber $(TF)_x$ with E, in two different ways, such that transition is given by f:

$$\begin{array}{ccccc} & & (TF)_x & & \\ & \overset{(i)}{\swarrow} & & \overset{(j)}{\searrow} & \\ E & & \overset{f}{\longrightarrow} & & E \end{array}$$

The $\mathbb{K}$-module structure on $(TF)_x$ induced from the chart indexed by (i) is not an intrinsic feature since it is different from the one induced from the chart (j). But the *class* of all $\mathbb{K}$-module structures induced from all possible bundle charts is an intrinsic feature of the fiber $(TF)_x$. As we have seen, the transition map f belongs to the group $\mathrm{Gm}^{0,2}(E)$, and hence the orbit under this group of the $\mathbb{K}$-module structure induced by chart (i) contains the one induced by chart (j). Therefore this orbit is an intrinsic structure of the fiber $(TF)_x$. In the terminology of Appendix BA, this can be rephrased by saying that the *set of linear structures on $(TF)_x$ that are bilinearly related to linear structures induced from charts* is an intrinsic feature. These structures can all be described in a very explicit way: the linear part $h = (h_1, h_2, h_3)$ of the map f preserves the linear structure of E, and therefore we may forget this linear part. Only the bilinear part (9.3) really is responsible for non-linearity of the action. In other words, all linear structures L^b that are bilinearly related to the linear structure L^0 coming from the "original chart" are given by push-forward of L^0 via a map of the form

$$f_b : W \times V \times W \to W \times V \times W, \quad (w, x', w') \mapsto (w, x', w' + b(w, x')) \tag{9.4}$$

with bilinear $b = b_x$. Addition and multiplication by scalars in the $\mathbb{K}$-module (E, L^b) are then given by the explicit Formulae (BA.1). Using the terminology of Appendix BA, we may summarize:

Proposition 9.4. *If F is a vector bundle over M, then any two $\mathbb{K}$-module structures on the fibers of TF over M induced by bundle charts are bilinearly related to each other (via continuous bilinear maps). If $F = TM$, then they are related to each other via continuous* symmetric *bilinear maps.*

Proof. Only the claim on TM remains to be proved. In this case, (9.3) gives

$$b_x(w,x') = d\varphi_{ij}(x)^{-1}\, d^2\varphi_{ij}(x)(w,x'), \tag{9.5}$$

which is a *symmetric* bilinear map. ■

9.5. *"Dictionary" between bilinear geometry and differential geometry.* It follows from the preceding proposition that any intrinsic object of bilinear geometry (in the sense of Section BA.2) on the "standard fiber" $W \times V \times W$ of TF over M has an intrinsic interpretation on the fiber bundle TF over M. In the following, we present a table of such structures together with their differential geometric, chart-independent definition. Here, $z_F : M \to F$ denotes the zero section of F and $z_{TF} : F \to TF$ the zero section of TF, and

$$P : \pi \oplus Tp : TF \to F \oplus TM$$

the vector bundle direct sum of the two projections $\pi : TF \to F$ and $Tp : TF \to TM$. A typical element $u \in F$ will be represented in a chart by (x,w), $w \in W$, and 0_x denotes the zero vector in F_x. Notation concerning bilinear algebra is explained in Appendix BA.

bilinear algebra	*differential geometry*
$W \times V \times W$, bilinear space	$(TF)_x$, fiber over $x \in M$
axes:	
$W \times 0 \times 0$ (first axis)	$z_{TF}(F_x) \subset (TF)_x$
$0 \times V \times 0$ (second axis)	$T(z_f)(T_xM) \subset (TF)_x$
$0 \times 0 \times W$ (third axis)	$T_{0_x}(F_x) \subset (TF)_x$
projections:	
$\mathrm{pr}_1 : W \times V \times W \to W$	$\pi_x : (TF)_x \to F_x$
$\mathrm{pr}_2 : W \times V \times W \to V$	$(Tp)_x : (TF)_x \to T_xM$
$\mathrm{pr}_{12} : W \times V \times W \to W \times V$	$P_x = \pi_x \oplus (Tp)_x : (TF)_x \to F_x \oplus T_xM$
fibrations:	
$\{w\} \times V \times W = \mathrm{pr}_1^{-1}(w)$	$\pi_x^{-1}(u)$
$W \times \{v\} \times W = \mathrm{pr}_2^{-1}(v)$	$(Tp)^{-1}(v)$
$\{w\} \times \{v\} \times W = \mathrm{pr}_{12}^{-1}(w,v)$	$P_x^{-1}(u,v)$
special endomorphisms:	
$\varepsilon : W \times V \times W \to W \times V \times W$, $(w,v,z) \mapsto (0,v,w)$	$\mathbb{K}[\varepsilon]$-module structure on $T(F_x)$ (fiber of Tp)
if $V = W$:	case of tangent bundle, $F = TM$:
$\kappa : (w,v,z) \mapsto (v,w,z)$	the "canonical flip" $\kappa : TTM \to TTM$

Later (Section 10.7) this table will be continued to a table of *non-intrinsic* structures, i.e. structures that depend on the choice of one linear structure among all the bilinearly related structures (linear connection).

9.6. *Bilinear bundles.* A *bilinear bundle* is a fiber bundle F over M such that each fiber is a *bilinear space* in the sense of Appendix BA. This means that vector bundles F_1, F_2, F_3 over M together with bundle injections $F_i \to F$ $(i = 1, 2, 3)$ are fixed (called *axes of* F) such that the axes are *linear* subbundles of F (i.e., subbundles on which transition functions act linearly); in all chart representations of F, the fiber F_x over $x \in M$ is assumed to be a direct sum $E = (F_1)_x \oplus (F_2)_x \oplus (F_3)_x$, and transition maps in the fiber over x are assumed to belong to the general bilinear group $\mathrm{Gm}^{0,2}(E)$. Proposition 9.4 says that the tangent bundle of a vector bundle is a bilinear bundle. One may construct also bilinear bundles that are not of this type. Many definitions and statements of the subsequent sections generalize to this framework; however, we will rather stay in the framework of tangent bundles of vector bundles. – Let us just add the remark that one can characterize bilinear bundles intrinsically, without referring to chart conditions. We leave this is an exercise to the interested reader.

10. Linear connections. I: Linear structures on bilinear bundles

10.1. *Linear structures on fiber bundles.* A *linear structure* $L=(a,m)$ *on a fiber bundle* $p: E \to M$ is given by smooth bundle maps

$$a: E \times_M E \to E, \quad m: \mathbb{K} \times E \to E$$

such that, for all $x \in M$, the fiber E_x is turned into a $\mathbb{K}$-module with linear structure L_x given by addition $a_x : E_x \times E_x \to E_x$ and multiplication by scalars $m_x : \mathbb{K} \times E_x \to E_x$. Morphisms of fiber bundles with linear structure are morphisms of fiber bundles that are fiberwise linear (w.r.t. the given linear structures L, L').

10.2. *Definition of linear connections and their morphisms.* As in Section 9.1, we assume that F is a vector bundle over M with projection $p: F \to M$. A *(linear) connection* on F is a linear structure L on the fiber bundle TF over M such that, for all $x \in M$ and for every bundle chart (say, with index j) around x, the structure L_x is bilinearly related to the linear structure $L_x^{(j)}$ induced on $(TF)_x$ by the bundle chart of TF with index j. (This is well-defined: if L_x is bilinearly related to *one* structure induced from a bundle chart, then Proposition 9.4 implies that it is bilinear related to *all* structures induced from bundle charts because "bilinearly related" is an equivalence relation.) A morphism of vector bundles F, F' with linear connections L, L' is a smooth bundle map $\widetilde{f} : F \to F'$ such that $T\widetilde{f} : TF \to TF'$ is fiberwise linear with respect to L and L'; following the usual convention, we will call them also *affine (with respect to the given connections)* (although the term *linear* would be more appropriate).

10.3. *Connections on the tangent bundle.* In case $F = TM$, we say that a connection L on TM is *torsionfree* or *symmetric* if L_x is bilinearly related via a *symmetric* bilinear map to L_{φ_j}. (As above, using the last assertion of Proposition 9.4, it is seen that this condition is well-defined.) This means that the action of the canonical flip $\kappa : TTM \to TTM$ is linear with respect to L (cf. Section BA.5). A smooth map $\varphi : M \to M'$ is called *affine (with respect to linear connections on TM and on TM')* if $T\varphi : TM \to TM'$ is $L - L'$-affine; i.e. if $TT\varphi : TTM \to TTM'$ is fiberwise linear with respect to L and L'. In case φ is an affine curve, i.e. $M = I \subset \mathbb{K}$ is open and equipped with the connection induced by this chart, $\varphi : I \to M'$ is also called a *geodesic.*

10.4. *Chart representations and existence questions.* Recall that the space of all bilinearly related structures on a product of three $\mathbb{K}$-modules is an affine space over $\mathbb{K}$ (Section BA.2). Applying this fiberwisely, we see that the space of all affine connections on F, denoted by $\mathrm{Conn}(F)$, also is an affine space over $\mathbb{K}$. More precisely, assume $L, L' \in \mathrm{Conn}(F)$. Then, for all $x \in M$, L_x and L'_x are bilinearily related to each other, which means that their difference is a well-defined bilinear map

$$b_x : V_1 \times V_2 \to V_3, \quad (u,v) \mapsto b_x(u,v) = \mathrm{pr}_3^L(u +_{L'} v)$$

(Section BA.1), where the three axes are canonically given by $V_1 = F_x$, $V_2 = T_xM$, $V_3 = \varepsilon F_x$ (cf. Table 9.4). In other words, $b = L - L'$ is a tensor field whose value at x is a bilinear map $F_x \times T_x M \to F_x$. Conversely, adding such a tensor field b to a linear connection L we get another linear connection L'.

Thus $\mathrm{Conn}(F)$ is an affine space over $\mathbb{K}$ (and over the ring $\mathcal{C}^\infty(M,\mathbb{K})$). Unfortunately, this affine space may be empty. However, the space of connections of the bundle over a chart domain is never empty because, by the very definition of a connection, every chart induces a connection over the chart domain. More precisely, fix a bundle chart $\widetilde{g}_i = (\varphi_i, g_i)$; this chart induces a linear structure L_x^i on the fiber $(TF)_x$, given by the bundle chart $(TF)_x \cong W \times V \times \varepsilon W$. Then, if L is any other connection on F defined over U_i, the difference $L - L_i$ corresponds to a tensor field $b := b^i : U \times W \times V \to W$, $(x,w,v) \mapsto b_x(w,v)$, which we call the *Christoffel tensor (of the connection in the given chart)*. The linear structure L_x is then explicitly given by $+_{L_x} = +_{b_x}$ with $+_{b_x}$ given by Equation (BA.1), and similarly for multiplication by scalars:

$$\begin{aligned}(x,u,v,w) +_{L_x} (x,u',v',w') &= (x, u+u', v+v', w+w' + b_x(u,v') + b_x(u',v)),\\ r_{L_x}(x,v,w) &= (x, ru, rv, rw + (r^2 - r) b_x(u,v)).\end{aligned} \tag{10.1}$$

The calculation from Sections 9.1 and 9.2 gives us the relation between Christoffel symbols b^i, b^j belonging to different bundle charts:

$$b^j_{\varphi_{ij}(x)}(g_{ij}(x)u, d\varphi_{ij}(x)v) = g_{ij}(x) b^i_x(u,v) + \partial_v g_{ij}(x,u). \tag{10.2}$$

In the sequel we will mainly work with connections over a chart domain. Global connections, if they exist, can be constructed in various ways: a family of tensor fields $(b^i)_{i\in I}$ satisfying (10.2) defines a linear connection L; or, if our manifold M has a partition of unity subordinate to a cover by charts, then the usual arguments show that the chart connections can be patched together to define a global connection on M. Finally, connections on Lie groups and symmetric spaces will be constructed in an intrinsic way without using chart arguments (see Chapters 23 and 26). – The following result generalizes a result due to Dombrowki, cf. [La99, Th. X.4.3]:

Theorem 10.5. (The Dombrowski Splitting Theorem.) *A linear connection L on a vector bundle F induces on TF a unique structure of a vector bundle over M such that*

$$\Phi : F \oplus_M TM \oplus_M F \to TF, \quad (w_x, \delta_x, w'_x) \mapsto z(w_x) + z(\delta_x) + \varepsilon(z(w'_x))$$

becomes an isomorphism of vector bundles over M, where the sum is taken with respect to the linear structure L_x on the fiber $(TF)_x$. Conversely, given an isomorphism of bilinear bundles $\Phi : F \oplus_M TM \oplus_M F \to TF$ whose restriction to the axes F, TM is just inclusion of axes, one recovers a linear connection L simply by push-forward of the canonical linear structure L_0 on $F \oplus_M TM \oplus_M F$ by Φ.

Proof. Choose a bundle chart g_j and let $(b_x)_{x\in U}$ be the corresponding "Christoffel symbols". Then Φ is represented in the chart by

$$\Phi(x,w,x',w') = (x,w,0,0) +_{b_x} (x,0,x',0) +_{b_x} (x,0,0,w') = (x,w,x',w' + b_x(w,x')),$$

that is, in each fiber Φ is given by f_{b_x} (with f_b as in Section BA.1). Hence Φ is invertible (the inverse being given in each fiber by f_{-b_x}) and smooth in both directions. It is fiberwise linear (by construction), hence an isomorphism of linear bundles.

Conversely, if Φ is an isomorphism of bilinear bundles (in the sense of Section 9.6), then it follows directly from the definitions that the push-forward $\Phi_*(L_0)$ is a linear connection on F. ■

If a connection is fixed, then the isomorphism Φ from the Dombrowski splitting will often be considered as an identification, also by denoted by

$$TF \overset{L}{\cong} F \oplus TM \oplus \varepsilon F \tag{10.3}$$

(where εF is just a copy of F, the prefix ε being added in order to distinguish the first and the third factor). In the following two sections we explain the relation with several classically well-known concepts that are related to connections.

10.6. *Structures defined by a connection: horizontal subspace and horizontal lift.* Assume L is a linear connection on a vector bundle $p : F \to M$. Fix $x \in M$ and $u \in F_x$; in a chart we write $u = (x, w)$. Recall from the table in Section 9.5 the following intrinsic subspaces: the *tangent space* $T_uF = \{w\} \times V \times W$, the axis $T_xM = z(T_xM) = 0 \times V \times 0$, and the *vertical subspace* $V_u = \{w\} \times 0 \times W$. The parallel to the axis $0 \times V \times 0$ in $(TF)_x$ through $u = (w, 0, 0)$ depends on the linear structure L_x; it is called the *horizontal subspace* H_u and is given in a chart representation by (cf. Equation (BA.15))

$$H_u = \{\big(w, v, b_x(w, v)\big) \,|\, v \in V\} = \mathrm{pr}_{13}^{-1}(w, 0). \tag{10.4}$$

It is a subspace of the tangent space T_uF such that

$$T_uF = V_u \oplus H_u, \tag{10.5}$$

and $Tp : TF \to TM$ induces a linear isomorphism $H_u \to T_xM$ with inverse called the *horizontal lift* and denoted by

$$h_u : T_xM \to H_u \subset T_uF. \tag{10.6}$$

10.7. *Structures defined by a connection: connector and connection one-form.* The isomorphism (10.3) from the Dombrowski splitting gives rise to three bundle projectors over M:

$$\mathrm{pr}_F : TF \to F, \qquad \mathrm{pr}_{TM} : TF \to TM, \qquad \mathrm{pr}_{\varepsilon F} : TF \to \varepsilon F.$$

The first one is just π_{TF}, the second one is Tp, and the third one really depends on the connection; it is called the *connector* and in the literature is often denoted by $K : TF \to F$. We prefer the notation

$$K = \mathrm{pr}_{\varepsilon F} : TF \to \varepsilon F$$

which permits to distinguish K from the first projection. In a bundle chart, the third projection is given by

$$K(x, u, v, w) = (x, w - b_x(u, v)).$$

The connector K has the usual properties known from real finite-dimensional differential geometry (cf. e.g. [KMS93, p. 110], [Bes78, p. 38]): for all $x \in M$,

(1) $K_x : (TF)_x \to \varepsilon F_x$ is *(intrinsically) bilinear* in the sense explained in Prop. BA.7,

(2) $K_x(z) = z$ for all z belonging to the third axis εF_x (vertical space) of $(TF)_x$.

One can show that these properties characterize the third projection. Instead with the connector K, we could also work with the projector $\mathrm{pr}_{13} = \mathrm{pr}_1 \times \mathrm{pr}_3$, i.e. with the bundle projector over M

$$\Psi := \mathrm{pr}_F \oplus \mathrm{pr}_{\varepsilon F} : TF \to F \oplus \varepsilon F$$

which is called the *connection one-form*, given in a chart by

$$\Psi(x, u, v, w) = (x, u, w - b_x(u, v)).$$

Its kernel is the horizontal bundle, and its image can be seen as the vertical bundle. The map Ψ can be seen as a fiber bundle map over M or as a vector bundle map over F which is linear in fibers and hence can be interpreted as a one-form over F. Moreover, composing with the canonical injection $F \oplus \varepsilon F \to TF$ (vertical bundle), we get Ψ in the version of a projector $\widetilde{\Psi} : TF \to TF$ having as image the vertical bundle (version used in [KMS93]). Using these maps, the (inverse of the) isomorphism from the Dombrowski splitting theorem can be written

$$\Phi^{-1} = \pi \oplus Tp \oplus K = Tp \oplus \Psi : TF \to F \oplus TM \oplus \varepsilon F. \tag{10.7}$$

The “dictionary” between bilinear algebra and differential geometry from Section 9.5 can now be continued – this part concerns “non-intrinsic structures”:

bilinear algebra	*differential geometry*
fixing L^b, $b \in \mathrm{Bil}(W \times V, W)$	fixing a connection L
$f_b : W \times V \times W \to W \oplus_b V \oplus_b W$	$\Phi : F \oplus TM \oplus \varepsilon F \to TF$ (Dombrowski)
projectors:	
$\mathrm{pr}_{13} : W \oplus_b V \oplus_b W \to W \times W$	$\Psi : TF \to F \oplus \varepsilon F$ (connection one-form)
$\mathrm{pr}_3 : W \oplus_b V \oplus_b W \to W$	$K : TF \to \varepsilon F$ (connector)
$\mathrm{pr}_{23} : W \oplus_b V \oplus_b W \to V \times W$	$\mathrm{pr}_{23} : TF \to TM \oplus \varepsilon F$ (no name)
spaces, decompositions:	
$H_w = \{(w, v, b(w, v)) \mid v \in V\}$	horizontal space $H_u = \Psi^{-1}((u, 0))$
$V_w = \{(w, 0, z) \mid z \in W\}$ (intrinsic)	vertical space V_u (intrinsic)
$\{w\} \times V \times W = V_w \oplus_b H_w$	$T_uF = V_u \oplus H_u$
$V \to H_w$, $v \mapsto (w, v, b(w, v))$	horizontal lift $h_u : T_xM \to H_u \subset T_uF$
non-intrinsic exact sequences:	
$W \oplus_b V \oplus_b 0 \to W \times V \times W \xrightarrow{\mathrm{pr}_3} W$	$F \oplus TM \to TF \xrightarrow{K} \varepsilon F$
$V \to W \times V \times W \xrightarrow{\mathrm{pr}_{13}} W \times W$	$TM \to TF \xrightarrow{\Psi} F \oplus \varepsilon F$
Case $V = W$:	Case $F = TM$:
torsion of L^b $(= L^b - L^{\kappa.b})$	torsion tensor of the connection
diagonal map $V \to (V \times V) \times V$	$S : TM \to TTM$ (spray, cf. Section 11.3.)

10.8. *Connections on bilinear bundles.* The preceding definitions and facts can be generalized in a straightforward way for arbitrary bilinear bundles (cf. Section 9.6): a *linear connection on a bilinear bundle* F is given by a linear structure L on F over M which is bilinearily related to structures induced by bundle charts. The Dombrowski Spitting Theorem and the notions of connector, connection one-form, and so on, immediately carry over to this context; details are left to the reader.

10.9. *Direct sum of connections.* From the classical linear algebra constructions of connections (tensor products, hom-bundles, dual connection...), only the *direct sum of connections* survives in our general context. In order to give an intrinsic definition of the direct sum, we first describe some canonical isomorphisms of bundles. For this, assume given three vector bundles F_1, F_2, F_3 over M, with fibers modelled on $\mathbb{K}$-modules W_1, W_2, W_3.

First of all, there is a canonical isomorphism

$$T(F_1 \times_M F_2) \cong TF_1 \times_{TM} TF_2. \tag{10.8}$$

In a bundle chart with base domain U, this amounts to the canonical identification

$$T(U \times W_1 \times W_2) \cong TU \times TW_1 \times TW_2$$

which is a special case of the general isomorphism $T(M \times N) \cong TM \times TN$ for manifolds M, N.

Next, $F_1 \times_M F_2$ is a vector bundle over M, but it may also be considered as a vector bundle over F_2: the fiber over $g \in (F_2)_x$ with $x \in M$ is $\{(f,g) | \, f \in (F_1)_x\} \subset (F_1)_x \times (F_2)_x$. Then

$$(F_1 \times_M F_2) \times_{F_2} (F_3 \times_M F_2)$$

is a vector bundle over F_2, whereas over M, *a priori*, it is just a fiber bundle. But

$$F_1 \times_M F_2 \times_M F_3 \to (F_1 \times_M F_2) \times_{F_2} (F_3 \times_M F_2), \quad (x; f, g, h) \mapsto (x, g; f, h) \tag{10.9}$$

is a bijection of bundles over M – in a chart, this amounts to the canonical identification

$$U \times (W_1 \times W_2 \times W_3) \cong (U \times W_2) \times W_1 \times W_3.$$

Now, the left hand side in (10.9) is a vector bundle *over* M, and hence, via the isomorphism (10.9), also the right hand side carries a canonical vector bundle structure over M.

Next, assume that F, G are vector bundles over M with connections L_i, $i = 1,2$. Then, using the canonical isomorphisms described above and the Dombrowski splittings (10.3) corresponding to L_1 and L_2, the following isomorphism defines a linear structure on $T(F \oplus_M G)$:

$$\begin{aligned} T(F \oplus_M G) &\cong TF \oplus_{TM} TG \\ &\cong (F \oplus_M TM \oplus_M \varepsilon TF) \oplus_{TM} (G \oplus_M TM \oplus_M \varepsilon TG) \\ &\cong F \oplus_M \varepsilon F \oplus_M TM \oplus_M G \oplus_M \varepsilon G. \end{aligned} \tag{10.10}$$

In a chart, this isomorphism is described, if b_x and a_x are the Christoffel tensors on F, resp. G,

$$(x; \varepsilon v, u, w, \varepsilon u', \varepsilon w') \mapsto (x; \varepsilon v, u, w, \varepsilon(u' + b_x(v,u)), \varepsilon(w' + a_x(v,w)),$$

i.e. by the Christoffel tensor $c_x(v,(u,w)) = b_x(v,w) + a_x(v,w)$.

10.10. *Connections and the tensor bundle.* Recall from Section BA.10 that to a bilinear space $E = V_1 \times V_2 \times V_3$ we can associate in an intrinsic way the *linear* space $Z := V_1 \otimes V_2 \oplus V_3$. Doing this construction pointwise, we can associate to every bilinear bundle F an "algebraic linear bundle"

$$Z := Z(F) := F_1 \otimes F_2 \oplus F_3 \tag{10.11}$$

i.e., Z is defined as a set, but without topology and hence without manifold structure (cf. Appendix L.2). Nevertheless we may speak about "bundle charts", which are just transition functions without specified smoothness properties. As with tensor products in ordinary finite dimensional real differential geometry, it is seen that there is a canonical projection $Z(F) \to M$ such that the fibers carry a well-defined $\mathbb{K}$-module structure, and hence Z is a well-defined "algebraic linear bundle". The map

$$TF \to Z(F), \quad (x,w,v,w') \mapsto (x, w \otimes v + w')$$

does not depend on the chart and defines a bundle map from a bilinear bundle to a linear bundle; this map is fiberwise intrinsically bilinear in the sense of Prop. BA.7. The sequence $F_3 \to F \to F_1 \times_M F_2$ gives rise to an exact sequence of algebraic linear bundles over M

$$0 \to F_3 \to Z(F) \to F_1 \otimes F_2 \to 0. \tag{10.12}$$

Linear connections now give rise to linear splittings of the sequence (10.12):

$$L : F_1 \otimes F_2 \to Z(F) \tag{10.13}$$

If $F = TM$, then (10.12) reads $0 \to TM \to Z(TM) \to TM \otimes TM \to 0$, and restricting to Σ_2-invariants, we get an exact sequence $0 \to TM \to Z(TM)^\kappa \to S^2(TM) \to 0$. Torsionfree connections give rise to cross-sections $\Gamma : S^2(TM) \to Z(TM)^\kappa$ of this sequence. In the finite-dimensional real case, there is a one-to-one correspondence between such cross-sections and connections – this is the point of view on connections used in [Lo69]. See also [P62] and the remarks on differential operators and symbols in Chapter 21.

10.11. *Comments on definitions of a linear connection.* As mentioned in the introduction, one finds many different definitions of connections in the literature, and it was not our aim to add a new item to this long list, but rather to propose a way how to organize it: both the axiomatics based on the connector (see, e.g., [Bes78], [KMS93]) and the axiomatics based on the connection one-form (see, e.g., [BGV92], [KMS93], [KoNo69]) focus on specific aspects (projectors) of the linear structure L and are equivalent to ours. Similarly, the definition in [La99, p. 104] takes some features of the injection $F \times_M TM \to TF$ as axiomatics and again is equivalent to ours. Using our concept of bilinear algebra, it is fairly easy to prove that all the aforementioned definitions of a connection are equivalent, or even to give yet other axiomatic characterizations. For instance, a connection defines an isomorphism of the fibers of $(Tp)_x : (TF)_x \to T_x M$ with $F_x \oplus \varepsilon F_x \cong F_x \oplus F_x$, and hence we have a whole 2×2-matrix algebra $M(2,2;\mathbb{K})$ acting by bundle maps of TF over TM, this action depending only on the connection L:

$$\varepsilon_u = \begin{pmatrix} 0 & 0 \\ \mathbf{1} & 0 \end{pmatrix}, \quad J_u = \begin{pmatrix} 0_W & \\ & \mathbf{1}_W \end{pmatrix}, \quad H_u = \begin{pmatrix} \mathbf{1}_W & \\ & -\mathbf{1}_W \end{pmatrix}, \quad Q_u = \begin{pmatrix} 0 & \mathbf{1} \\ 0 & 0 \end{pmatrix}$$

where $W \cong F_x$ and $u = x + \varepsilon v \in TM$. We propose to call this algebra the *connector algebra*. Here, the first operator is independent of L (the intrinsic ε-tensor), whereas the other operators do depend on L (except on the fiber of TF over the zero vector $0_x \in T_x M$). The operator Q seems to be particularly interesting – it is in a sense opposed to ε: when ε is seen as sort of "annihilator" (making an object "infinitesimally small"), Q is a sort of "creator" (magnifying an "infinitesimally small" object by an "infinitely large" factor). – Finally, the relation with the definition of connections via sprays and covariant derivatives is discussed in the following two chapters.

11. Linear connections. II: Sprays

11.1. *Second order vector fields.* This chapter concerns linear connections on the tangent bundle: let $F = TM$. Recall that the second order jet bundle J^2M is the subbundle of $TF = TTM$ fixed under the action of the permutation group Σ_2, and that the two projections $p_j : TTM \to TM$ give rise to one jet projection $\pi_{1,2} : J^2M \to J^1M = TM$:

$$\begin{array}{ccc} J^2M & \subset & TTM \\ \downarrow \pi_{1,2} & & \downarrow p_j \\ J^1M & = & TM. \end{array} \tag{11.1}$$

As we have seen in 8.4 (B), the jet bundle $J^2M \to TM$ does not have a canonical "zero-section" since the zero-sections of p_j are not compatible with the canonical flip. Thus $\pi_{1,2} : J^2M \to J^1M$ is an affine bundle and not a vector bundle. We define a *second order vector field on* M to be a cross-section $X : J^1M \to J^2M$ of the jet projection $J^2M \to J^1M$. This defines a vector bundle structure on J^2M over J^1M, simply by considering X as zero-section. One may also say that $X : TM \to J^2M \subset TTM$ is a vector field on TM, taking values in the space of Σ_2-invariants. Thus a vector field $X : TM \to TTM$ is a second order vector field if and only if in a bundle chart it is given by

$$X(x + \varepsilon v) = x + (\varepsilon_1 + \varepsilon_2)v + \varepsilon_1\varepsilon_2\xi(x, v) \tag{11.2}$$

with some smooth map $\xi : U \times V \to V$.

11.2. *The second order vector field defined by a linear connection.* If L is a linear connection on TM and

$$\varepsilon_1TM \times_M \varepsilon_2TM \times_M \varepsilon_1\varepsilon_2TM \to TTM, \quad (x; \varepsilon_1u, \varepsilon_2v, \varepsilon_1\varepsilon_2v) \mapsto \varepsilon_1u +_L \varepsilon_2v +_L \varepsilon_1\varepsilon_2v$$

its corresponding Dombrowski splitting, then the diagonal imbedding

$$S : TM \to TM \times_M TM \times_M TM \overset{L}{\cong} TTM, \quad v \mapsto (v, v, 0) = \varepsilon_1v +_L \varepsilon_2v$$

is a second-order vector field. In a chart, it is given by

$$S(x + \varepsilon v) = x + (\varepsilon_1 + \varepsilon_2)v + \varepsilon_1\varepsilon_2b_x(v, v) \tag{11.3}$$

where b is the Christoffel tensor of L. (Note that S depends only on the quadratic map determined by b, and hence, if 2 is invertible in $\mathbb{K}$, we may w.l.o.g. b assume to be symmetric, i.e. L to be torsionfree.) For instance, if M is open in the topological $\mathbb{K}$-module V and L is the canonical chart connection ($b = 0$), then $S(x, v) = (x, v, v, 0)$ is the corresponding second order vector field.

11.3. *Sprays.* Not every second order vector field comes from a linear connection on TM because the smooth map ξ from (11.2) can be chosen arbitrarily. Maps of the special form $\xi(x, v) = b_x(v, v)$ can be characterized intrinsically by the following *spray condition*: let $r \in \mathbb{K}$ and recall from Section 7.3 (D) the intrinsically defined action $l^{(2)} : \mathbb{K} \times TTM \to TTM$ which is given in a chart by $r^2(x,u,v,w) = (x, ru, rv, r^2 w)$ (and which can also be defined in the purely algebraic context of bilinear spaces, see remark after Prop. BA.3). Then, if S comes from a linear connection L, we have, in a chart

$$\begin{aligned} S(x+\varepsilon rv) &= x + (\varepsilon_1+\varepsilon_2)rv + \varepsilon_1\varepsilon_2 b_x(rv,rv) \\ &= x + r(\varepsilon_1+\varepsilon_2)v + r^2\varepsilon_1\varepsilon_2 b_x(v,v) = l_r^{(2)} \circ S(x+\varepsilon v). \end{aligned}$$

By definition, a *spray* is a second order vector field satisfying the "homogenity condition" $X \circ r_{TM} = l_r^{(2)} \circ X$, for all $r \in \mathbb{K}$.

Theorem 11.4. *Assume* 2 *is invertible in* $\mathbb{K}$*. Then there is a canonical bijection between torsionfree connections* L *on* TM *and sprays* $S : TM \to J^2M$*.*

Proof. We have just seen that the second order vector field S associated to a connection is a spray. Conversely, if S is a spray, then, writing $S(x+\varepsilon v) = x + \varepsilon_1 v + \varepsilon_2 v + \varepsilon_1\varepsilon_2 q_x(v)$, we deduce that q_x is smooth and homogeneous quadratic with respect to scalars and hence is, by Cor. 1.12, a vector valued quadratic form. If 2 is invertible in $\mathbb{K}$, then the symmetric bilinear map b_x can be recovered from q_x by polarisation. Clearly, both constructions are inverse to each other. ■

The preceding result generalizes a theorem of Ambrose, Palais and Singer, cf. [MR91, 7.3]. In the case of general characteristic, the spray is equivalent to what one might call the "connection on the *quadratic* bundle TTM" (cf. Section BA.9).

11.5. *Second order differential equations and geodesics.* If $X : TM \to J^2M$ is a second order vector field, then, by definition, a *geodesic with respect to* X is a curve $\alpha : \mathbb{K} \to M$ such that $\alpha' : \mathbb{K} \to TM$ is an integral curve of S, i.e. $\alpha''(t) = S(\alpha'(t))$. Here, $\alpha'(t) = T_t\alpha \cdot 1 = T\alpha(t+\varepsilon 1)$ and $\alpha''(t) = (\alpha')'(t) = TT\alpha(t+\delta 1+\delta^{(2)}1)$. In a chart representation (11.2), the condition $\alpha''(t) = S(\alpha'(t))$ amounts to the second order differential equation $\alpha''(t) = \xi(\alpha(t), \alpha'(t))$ (cf. [La99, p. 99]).

On the other hand, if $X = S$ is the spray of a connection L, the notion of geodesic has already been defined in Section 10.3: it is an affine map $\alpha : \mathbb{K} \to M$, i.e., $J^2\alpha : J^2\mathbb{K} \to J^2M$ is fiberwise linear. This is equivalent to saying that the following diagram commutes:

$$\begin{array}{ccc} J^2\mathbb{K} & \overset{J^2\alpha}{\to} & J^2M \\ S_{\mathbb{K}} \uparrow & & \uparrow S \\ J^1\mathbb{K} & \overset{J^1\alpha}{\to} & J^1M, \end{array}$$

where $S_{\mathbb{K}} : T\mathbb{K} \to J^2\mathbb{K}$ is the spray corresponding to the canonical chart connection of $\mathbb{K}$. In a bundle chart, we have, with ξ as in (11.2), using the spray property,

$$\begin{aligned} S \circ J^1\alpha(x+\varepsilon v) &= \alpha(t) + \delta\alpha'(t)v + \delta^{(2)}\xi(\alpha(t), \alpha'(t)v) \\ &= \alpha(t) + \delta\alpha'(t)v + \delta^{(2)}v^2\xi(\alpha(t), \alpha'(t)), \\ J^2\alpha \circ S_{\mathbb{K}}(t+\varepsilon v) &= \alpha(t) + \delta\alpha'(t)v + \delta^{(2)}\alpha''(t)v^2, \end{aligned}$$

and hence the condition $J^2\alpha \circ S_{\mathbb{K}} = S \circ J^1\alpha$ reads in a chart $\alpha''(t) = \xi(\alpha(t), \alpha'(t))$. This is the same condition as above, and hence both notions of geodesic agree.

For $v \in T_x M$, we say that $S(v) \in (J^2 M)_x$ is the *geodesic* 2*-jet of* v. Note that in general we have no existence or uniqueness statements on geodesics; but the preceding arguments show that, if a geodesic α_v with $\alpha_v(0) = x$ and $\alpha_v'(0) = v$ exists, then $J^2 \alpha(0) = S(v)$.

12. Linear connections. III: Covariant derivative

12.1. *The covariant derivative associated to a connection.* We return to the case of a general vector bundle F with linear connection L. Assume $X : M \to TM$ is a vector field on M and $Y : M \to F$ a section of F. If $K : TF \to \varepsilon F$ is the connector associated to a connection L on F, we define

$$\begin{aligned} \nabla Y &:= K \circ TY : \quad TM \to \varepsilon F, \\ \nabla_X Y &:= K \circ TY \circ X : \quad M \to \varepsilon F. \end{aligned}$$

Lemma 12.2. *Let $(b_x)_{x\in U}$ be the Christoffel tensor of the connection L in the chart U. Then*

$$(\nabla_X Y)(x) = dY(x)X(x) + b_x(Y(x), X(x)).$$

Proof. In a bundle chart, by some abuse of notation, $X(x) = x + \varepsilon X(x)$, $Y(x) = (x, Y(x))$, $TY(x+\varepsilon v) = (x, Y(x)) + \varepsilon(v, dY(x)v))$, $K((x,w)+\varepsilon(x',w')) = x + \varepsilon(w' + b_x(w,x'))$, and thus

$$\begin{aligned} (\nabla_X Y)(x) &= x + \varepsilon K\big((x, Y(x)) + \varepsilon(X(x), dY(x)X(x))\big) \\ &= x + \varepsilon(dY(x)X(x) + b_x(Y(x), X(x))). \end{aligned}$$ ∎

Corollary 12.3. *If L is a torsionfree connection on TM, then for all vector fields X, Y,*

$$[X,Y](x) = (\nabla_Y X - \nabla_X Y)(x).$$

Proof. In a chart, the preceding lemma gives for the right hand side, due to the symmetry of b_x, the expression $dX(x)Y(x) - dY(x)X(x)$ which is precisely the definition of the Lie bracket. ∎

12.4. *General covariant derivatives.* A *(general) covariant derivative* is a $\mathbb{K}$-bilinear map from sections to sections,

$$\nabla : \Gamma(TM) \times \Gamma(F) \to \Gamma(F)$$

such that, for all smooth functions f and all sections X, Y,

$$\nabla_{fX} Y = f \nabla_X Y, \qquad \nabla_X (fX) = Xf \cdot Y + f \nabla_X Y.$$

It is easily proved that the covariant derivative associated to a connection L has these properties. The converse is not true: not every general covariant derivative is associated to a linear connection – see [La99, p. 202 ff and p. 279/80] for a discussion (which can, *mutatis mutandis*, be applied to our situation) of the infinite dimensional real case. In particular, if one wants to define the curvature tensor in terms of covariant derivatives, it is not sufficient to use just the properties of a general covariant derivative (see [La99, p. 231 ff]). Therefore, in Chapter 18, we will define the curvature tensor in a different way.

12.5. *Covariant derivative of a tensor field.* For a tensor field of type $(k,0)$, say $\omega : TM \times_M \ldots \times_M TM \to \mathbb{K}$, the covariant derivative $\nabla\omega$ can be defined along the lines of Remark 4.8. E.g., for a one-form $\omega : TM \to \mathbb{K}$ we define

$$(\nabla\omega)(X,Y) := L_X(\omega(Y)) - \omega(\nabla_X Y); \tag{12.1}$$

we then have to check that the value at x depends only on the value of X and Y at x and that the result depends smoothly on x, $X(x)$ and $Y(x)$. Then, by Remark 4.8, $\nabla\omega$ can be identified with a well-defined tensor field $TM \times_M TM \to \mathbb{K}$. In the sequel we will not use this construction, but see Section 13.4 for an analog of Corollary 12.3.

13. Natural operations. I: Exterior derivative of a one-form

13.1. *Bilinear maps defined on TF.* If $p : F \to M$ is a vector bundle over M, a smooth bundle map $\widetilde{f} : TF \to E$, $f : M \to N$ into a vector bundle E over N will be called *(intrinsically) bilinear* if it is linear in fibers both over F and over TM. This is equivalent to saying that, for all $x \in M$, $\widetilde{f}_x : (TF)_x \to E_{f(x)}$ is intrinsically bilinear in the sense of Proposition BA.7. From that proposition it follows that $\widetilde{f}$ is bilinear if, and only if, $\widetilde{f}$ has a chart representation of the form

$$\widetilde{f}(x, w, v, w') = (f(x), g_x(w, v) + \lambda_x(w')) \tag{13.1}$$

with bilinear g_x and linear λ_x. A bilinear map $\widetilde{f}$ is called *homogeneous* if it is constant on the vertical bundle (fibers of $\pi \times_M Tp$), and this is the case iff $\lambda = 0$ (see Cor. BA.8). Thus the homogeneous bilinear maps are in one-to-one correspondence with the fiberwise bilinear maps $g : F \times_M TM \to E$ via $\widetilde{f} = g \circ (Tp \times \pi)$:

$$\begin{array}{ccc} TF & \xrightarrow[g]{\widetilde{f}} & E \\ \downarrow & \nearrow & \\ TM \times_M F & & \end{array}$$

13.2. *The exterior derivative of a one-form.* Recall that a *tensor field of type* $(k, 0)$ is a smooth map $\omega : \times_M^k TM \to \mathbb{K}$, multilinear in fibers over M. If ω is alternating in fibers, it is called a *k-form.* A *one-form on a vector bundle* E is a smooth map $\omega : E \to \mathbb{K}$ such that $\omega_x : E_x \to \mathbb{K}$ is linear for all $x \in M$. Then

$$\mathrm{pr}_2 T\omega : TE \to T\mathbb{K} \to \varepsilon\mathbb{K}$$

is (intrinsic) bilinear in the sense of 13.1: in fact, $T\omega$ is linear on fibers over E since it is a tangent map, and $\mathrm{pr}_2 T\omega$ is linear in fibers over TM since already ω was linear in fibers over M. Equivalently, the bilinearity of $\mathrm{pr}_2 T\omega$ follows from the chart representation

$$T\omega : TF \to T\mathbb{K}, \quad (x, \varepsilon v, w, \varepsilon w') \mapsto \omega(x, v) + \varepsilon(\partial_v \omega(x, w) + \omega(x, w')).$$

Comparing with (13.1) we see that $\mathrm{pr}_2 T\omega$ is bilinear, but non-homogeneous (the term $\lambda_x(w')$ corresponds to $\omega(x, w')$). For a general vector bundle F we cannot extract a homogeneous bilinear map from this. However, if $F = TM$, then we have the alternation map from intrinsic bilinear maps $TTM \to \mathbb{K}$ to homogeneous ones:

$$\begin{aligned} &\mathrm{alt}(\mathrm{pr}_2 T\omega)(x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12}) \\ &\quad = (\mathrm{pr}_2 T\omega)(x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12}) - (\mathrm{pr}_2 T\omega)(x + \varepsilon_1 v_2 + \varepsilon_2 v_1 + \varepsilon_1\varepsilon_2 v_{12}) \\ &\quad = \partial_{v_1}\omega(x, v_2) - \partial_{v_2}\omega(x, v_1) \end{aligned}$$

is homogeneous (the value does not depend on v_{12}). As explained in 13.1, homogeneous bilinear maps $\widetilde{f}$ correspond to tensor fields g of type $(2, 0)$, and if $\widetilde{f}$ is skew-symmetric, then so is g. Hence we end up with 2-form denoted by

$$d\omega : TM \times_M TM \to \mathbb{K}, \quad (x; v_1, v_2) \mapsto \mathrm{alt}(\mathrm{pr}_2 T\omega)(x + \varepsilon_1 v_1 + \varepsilon_2 v_2). \tag{13.2}$$

In other terms, the diagram

$$\begin{array}{ccc} TTM & \stackrel{\mathrm{pr}_2 T\omega}{\longrightarrow} & \mathbb{K} \\ \downarrow & & \\ TM \times_M TM & & \end{array}$$

can in general not be completed to a commutative triangle, but

$$\begin{array}{ccc} TTM & \stackrel{\mathrm{alt}\,\mathrm{pr}_2 T\omega}{\longrightarrow} & \mathbb{K} \\ \downarrow & & \\ TM \times_M TM & & \end{array}$$

can, namely by $\mathrm{d}\omega$. The chart representation

$$\mathrm{d}\omega(x; v_1, v_2) = \partial_{v_1}\omega(x, v_2) - \partial_{v_2}\omega(x, v_1) \tag{13.3}$$

shows that $\mathrm{d}\omega$ coincides with the usual exterior derivative of a one-form. It is clear that in the preceding arguments the target space $\mathbb{K}$ can be replaced by any other fixed topological $\mathbb{K}$-module W, i.e. we can define $\mathrm{d}\omega$ for any 1-form $\omega : TM \to W$.

13.3. If $\omega = df$, i.e. $\omega(x, v) = T_x f \cdot v$, then from (13.3) we get $\mathrm{d}\omega = \mathrm{dd}f = 0$ because the ordinary second differential $d^2 f(x)$ is symmetric.

13.4. *Exterior derivative and connections.* The chart formula (13.3) has the following generalization: if ∇ is the covariant derivative of an arbitrary torsionfree connection on M and ω a one-form on M, then for all $x \in M$ and $u, v \in T_x M$, we have

$$\mathrm{d}\omega(x; u, v) = (\nabla\omega)(x; u, v) - (\nabla\omega)(x; v, u). \tag{13.4}$$

In fact, by definition of the covariant derivative of a one-form (Section 12.5),

$$\begin{aligned} \nabla\omega(X, Y) - \nabla\omega(Y, X) &= X\omega(Y) - Y\omega(X) - \omega(\nabla_X Y) + \omega(\nabla_Y X) \\ &= X\omega(Y) - Y\omega(X) - \omega([X, Y]) \end{aligned}$$

since ∇ is torsionfree (Cor. 12.3). Now fix $x \in M$ and $v, w \in T_x M$; extend, in a chart, v and w to constant vector fields X, Y on the chart domain $U \subset M$; then $[X, Y] = 0$ and hence from (13.3) we get

$$\nabla\omega(x; v, w) - \nabla\omega(x; w, v) = \partial_v\omega(x, w) - \partial_w\omega(x, v) = \mathrm{d}\omega(x; v, w).$$

13.5. *Exterior derivative of a vector bundle valued one-form.* If we replace ω by a bundle-valued one-form $\omega : TM \to E$, then in presence of a connection on E with connector $K_E : TE \to \varepsilon E$ we can define $\mathrm{d}\omega := \mathrm{alt}(K_E \circ T\omega)$ which is skew-symmetric $TM \times TM \to \varepsilon E$. Note that the preceding definition in the scalar or vector valued case is a special case of this, where pr_2 is just the connector of the canonical flat connection on $\mathbb{K}$, resp. on X. An exterior derivative of bundle valued forms, in presence of a connection, has also been defined in [Lav87, p. 165 ff].

13.6. *Remark on the infinitesimal Stoke's theorem.* In the context of synthetic differential geometry, one can prove a purely infinitesimal version of Stoke's theorem, cf. [Lav87, p. 114 ff] or [MR91, p. 134 ff]. It should be possible to state and prove such a version also in the present context – this will be taken up elsewhere.

14. Natural operations. II: The Lie bracket revisited

14.1. *Subgroups of the group of diffeomorphisms of TTM.* The group $\Gamma :=$ Diff(TTM) of diffeomorphisms (over $\mathbb{K}$) of TTM contains several subgroups which are defined via the behaviour with respect to the projections $p_i : TTM \to TM$, $i = 1, 2$ and $p : TTM \to M$:

(1) Let Γ^+ be the subgroup preserving fibers of the projection $p : TTM \to M$ and permuting the fibers of p_1 and those of p_2. In a chart, $f \in \Gamma^+$ is represented by

$$f(x+\varepsilon_1 v_1+\varepsilon_2 v_2+\varepsilon_1\varepsilon_2 v_{12}) = x+\varepsilon_1 f_1(x, v_1)+\varepsilon_2 f_2(x, v_2)+\varepsilon_1\varepsilon_2 f_{12}(x, v_1, v_2, v_{12})$$

(2) Let Γ^1 be the subgroup of Γ^+ preserving all fibers of the projection p_1. In a chart:

$$f(x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12}) = x + \varepsilon_1 v_1 + \varepsilon_2 f_2(x, v_2) + \varepsilon_1\varepsilon_2 f_{12}(x, v_1, v_2, v_{12})$$

(3) Let Γ^2 be the subgroup of Γ^+ preserving fibers of p_2. In a chart,

$$f(x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12}) = x + \varepsilon_1 f_1(x, v_1) + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 f_{12}(x, v_1, v_2, v_{12})$$

(4) Let $\Gamma^{12} = \Gamma^1 \cap \Gamma^2$ be the subgroup preserving all vertical spaces; in a chart,

$$f(x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12}) = x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 f_{12}(x, v_1, v_2, v_{12}).$$

It is clear that Γ^1 and Γ^2 are normal subgroups in Γ^+. Therefore, if $g \in \Gamma^1$ and $h \in \Gamma^2$, it follows that the group commutator $[g, h] = (ghg^{-1})h^{-1} = g(hg^{-1}h^{-1})$ belongs both to Γ^1 and to Γ^2 and hence belongs to Γ^{12}. In Theorem 14.4 we will express the Lie bracket of vector fields in terms of this commutator.

14.2. *Sections of the second order tangent bundle.* We denote by $\mathfrak{X}^2(M) :=$ $\Gamma(M, T^2M)$ the space of smooth sections of the second tangent bundle $T^2M \to M$. In a chart with chart domain $U \subset M$, we write $T^2U \cong U \times \varepsilon_1 V \times \varepsilon_2 V \times \varepsilon_1\varepsilon_2 V$, and a section $X : U \to T^2U$ is written in the form

$$X(p) = p + \varepsilon_1 X_1(p) + \varepsilon_2 X_2(p) + \varepsilon_1\varepsilon_2 X_{12}(p) \tag{14.1}$$

with (chart-dependent) vector fields $X_\alpha : U \to V$. In this notation, X is a second-order vector field (cf. Section 11.1) if, and only if, in any chart representation we have $X_1(p) = X_2(p)$ for all $p \in U$. There are three canonical injections of the space of vector fields $\mathfrak{X}(M)$ into $\mathfrak{X}^2(M)$ which simply correspond to the three canonical inclusions of axes $\iota_\alpha : TM \to TTM$, $\alpha = 01, 10, 11$, by letting $\varepsilon^\alpha Y := \iota_\alpha \circ Y : M \to TTM$ for a vector field $Y : M \to TM$. In the notation (14.1), these are the sections X having just one non-trivial component, i.e., with $X_\alpha = Y$ and $X_\beta = 0$ for $\beta \neq \alpha$.

Theorem 14.3.

(1) *There is a natural group structure on the space* $\mathfrak{X}^2(M)$*, given in a chart representation as above, by the formula*

$$(X \cdot Y)(x) = x + \varepsilon_1(X_1(x) + Y_1(x)) + \varepsilon_2(X_2(x) + Y_2(x)) + \\ \varepsilon_1\varepsilon_2(X_{12}(x) + Y_{12}(x) + dX_1(x)Y_2(x) + dX_2(x)Y_1(x)).$$

The space $\Gamma(M, J^2M)$ *of sections of* J^2M *is a subgroup of* $\mathfrak{X}^2(M)$*, and the three canonical injections* $\mathfrak{X}(M) \to \mathfrak{X}^2(M)$ *are group homomorphisms (where* $\mathfrak{X}(M)$ *is equipped with addition of vector fields).*

(2) *Every* $X \in \mathfrak{X}^2(M)$ *gives rise, in a natural way, to a diffeomorphism* $\widetilde{X} : T^2M \to T^2M$*, which, in a chart representation as above, is described by*

$$\widetilde{X}(x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12}) = x + \varepsilon_1(v_1 + X_1(x)) + \varepsilon_2(v_2 + X_2(x)) + \\ \varepsilon_1\varepsilon_2(v_{12} + X_{12}(x) + dX_1(x)v_2 + dX_2(x)v_1).$$

(3) *The map* $\mathfrak{X}^2(M) \to \mathrm{Diff}(TTM)$*,* $X \mapsto \widetilde{X}$ *is an injective group homomorphism.*

Proof. (2) We fix a point $p \in M$ and define $\widetilde{X}$ on a chart neighborhood T^2U of p by the formula given in the claim. We then have to show that this definition is independent of the chosen chart. This, in turn, amounts to proving that the map $X \mapsto \widetilde{X}$ is *natural* in the following sense: for all local diffeomorphisms f defined on a neighborhood of x, we have $f_*\widetilde{X} = \widetilde{(f_*X)}$, where f_* is the natural action of f on $\mathrm{Diff}(TTM)$, resp. on $\mathfrak{X}^2(M)$ by $f_*h = T^2f \circ h \circ (T^2f)^{-1}$, resp. by $f_*X = T^2f \circ X \circ f^{-1}$. In order to get rid of the inverse, we show, more generally, that, if $X, Y \in \mathfrak{X}^2(M)$ are f*-related* (i.e. $T^2f \circ X = Y \circ f$) for a smooth map $f : U \to M$, then $T^2f \circ \widetilde{X} = \widetilde{Y} \circ T^2f$. Now,

$$\begin{aligned}
&T^2f(\widetilde{X}(x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12})) = \\
&\quad f(x) + \varepsilon_1(df(x)(v_1 + X_1(x))) + \varepsilon_2(df(x)(v_2 + X_2(x))) + \\
&\quad \varepsilon_1\varepsilon_2 \cdot \Big(df(x)(v_{12} + X_{12}(x) + dX_1(x)v_2 + dX_2(x)v_1) + \\
&\qquad d^2f(x)(v_1 + X_1(x), v_2 + X_2(x))\Big), \\
&\widetilde{Y}(T^2f(x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12})) = \\
&\quad f(x) + \varepsilon_1(df(x)v_1 + Y_1(f(x))) + \varepsilon_2(df(x)v_2 + Y_2(f(x))) + \\
&\quad \varepsilon_1\varepsilon_2 \cdot \Big(d^2f(x)(v_1, v_2) + df(x)v_{12} + dY_1(f(x))df(x)v_2 + dY_2(f(x))df(x)v_1\Big),
\end{aligned}$$

and both expressions are equal if X and Y are f-related. Thus the map $\widetilde{X} : TTM \to TTM$ is well-defined, and the construction is natural in the sense explained above. Let us prove that $\widetilde{X}$ is a diffeomorphism. In fact, a direct check shows that its inverse is $\widetilde{X^{-1}}$ with a section $X^{-1} : M \to TTM$ defined by

$$X^{-1}(x) = x - \varepsilon_1 X_1(x) - \varepsilon_2 X_2(x) - \\ \varepsilon_1\varepsilon_2\Big(X_{12}(x) - (dX_1(x)X_2(x) + dX_2(x)X_1(x))\Big). \tag{14.2}$$

(1) By comparing the formulas from parts (1) and (2) of the claim, we see that the product is defined by $(X \cdot Y)(x) = \widetilde{X}(Y(x))$, and hence is chart-independent. By a straightforward computation, one shows that the product is associative: $((X \cdot Y) \cdot Z)(x) = (X \cdot (Y \cdot Z))(x)$, and that X^{-1} as in (14.2) is an inverse of X as in (14.1). Moreover, it is clear from the explicit formula that, if X, Y are sections of J^2M, i.e., $X_1 = X_2$ and $Y_1 = Y_2$, then $X \cdot Y$ is again a section of J^2M over U, and similarly for the inverse, hence the sections of J^2M form a subgroup of $\mathfrak{X}^2(M)$. Finally, the explicit formula also shows that, if X and Y are vector fields, seen as elements of $\mathfrak{X}^2(M)$ in one of the three canonical ways, then $X \cdot Y$ simply corresponds to the sum $X + Y$ of vector fields.

(3) Choosing Z such that $Z(x) = v$, we get

$$\widetilde{XY}(v) = ((XY)Z)(x) = (X(YZ))(x) = \widetilde{X} \circ \widetilde{Y}(v).$$

Clearly, $\widetilde{0} = \mathrm{id}_{TTM}$, and hence we have a group action. The action is faithful since, if $\widetilde{X}(x) = x$ for all $x \in U$, then $X_1(x) = 0 = X_2(x)$ and thus also $X_{12}(x) = 0$ for all $x \in U$. ■

For a conceptual, chart-independent definition of the group structure on the space $\Gamma(M, T^kM)$ of sections of T^kM over M, see Theorem 28.2.

Theorem 14.4. *The group commutator in the group $\mathfrak{X}^2(M)$ and the Lie bracket in $\mathfrak{X}(M)$ are related via*

$$[\varepsilon_1 X, \varepsilon_2 Y]_{\mathfrak{X}^2(M)} = \varepsilon_1\varepsilon_2[X, Y]$$

for all $X, Y \in \mathfrak{X}(M)$. Equivalently, the group commutator in the group $\mathrm{Diff}(TTM)$ *and the Lie bracket in $\mathfrak{X}(M)$ are related via*

$$[\widetilde{\varepsilon_1 X}, \widetilde{\varepsilon_2 Y}] = \varepsilon_1 \widetilde{\varepsilon_2 [X, Y]}.$$

Proof. We fix $x \in M$. Let $X : M \to TM$ be a vector field. It gives rise, via the three imbeddings $\mathfrak{X}(M) \to \mathfrak{X}^2(M)$, to three diffeomorphisms of TTM. Specialising the chart formula from Theorem 14.3 (2), these three diffeomorphisms are described by their values at the point $u = x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12}$:

$$\begin{aligned}
\widetilde{\varepsilon_1 X}(u) &= x + \varepsilon_1(v_1 + X(x)) + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2(v_{12} + dX(x)v_2) \\
\widetilde{\varepsilon_2 X}(u) &= x + \varepsilon_1 v_1 + \varepsilon_2(v_2 + X(x)) + \varepsilon_1\varepsilon_2(v_{12} + dX(x)v_1) \\
\widetilde{\varepsilon_1\varepsilon_2 X}(u) &= x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2(v_{12} + X(x)).
\end{aligned} \tag{14.3}$$

We observe that these diffeomorphisms belong to the groups Γ^2, Γ^1 resp. Γ^{12} defined in Section 14.1. (Without using Theorem 14.3, these diffeomorphisms may be defined as follows: the first two maps may also be seen as the tangent maps $T_{\varepsilon_i}\widetilde{X}$ of the infinitesimal automorphism $\widetilde{X} : TM \to TM$, and the third map is defined by translation in direction of the vertical bundle $VM \subset TTM$, and such translations are canonical in the bilinear space $(TTM)_x$, cf. Section BA.2.) Thus, if X and Y are vector fields, both diffeomorphisms $\widetilde{\varepsilon_1 X}$ and $\widetilde{\varepsilon_2 Y}$ preserve the fiber $(TTM)_x$

and are represented in a chart $V \times V \times V$ of this fiber by bijections g and h of the fiber given by

$$g(v, x', v') = (v + a, x', v' + \alpha(x')), \quad h(v, x', v') = (v, x' + b, v' + \beta(v))$$

with $a = X(x)$, $b = Y(x)$, $\alpha = dX(x)$, $\beta = dY(x)$. Then we obtain for the commutator

$$\begin{aligned}[g, h](v, x', v') &= gh(v - a, x' - b, v' - \alpha(x' - b) - \beta(v)) \\ &= g(v - a, x', v' - \alpha(x' - b)\beta(v) + \beta(v - a)) \\ &= g(v, x', v' - \alpha(x') + \alpha(b) - \beta(a) + \alpha(x')) \\ &= (v, x', v' + \alpha(b) - \beta(a)).\end{aligned}$$

The value at 0_x is therefore

$$\begin{aligned}[g, h](0, 0, 0) &= (0, 0, \alpha(b) - \beta(a)) = (0, 0, dX(x)Y(x) - dY(x)X(x)) \\ &= (0, 0, [X, Y](x)) = \varepsilon_1\varepsilon_2[X, Y](x)\end{aligned}$$

by definition of the Lie bracket in Theorem 4.2. ■

14.5. *Remark on definition of the Lie bracket and the Jacobi identity.* The definition of the Lie bracket via Theorem 4.2 is not very conceptual. Theorem 14.4 offers a more conceptual way of defining it: *for two vector fields $X, Y \in \mathfrak{X}(M)$, there exists a unique vector field $[X, Y] \in \mathfrak{X}(M)$ such that the group commutator $[\widetilde{\varepsilon_1 X}, \widetilde{\varepsilon_2 Y}]$ agrees with $\widetilde{\varepsilon_1\varepsilon_2[X, Y]}$.* It is then easily seen that $[X, Y]$ depends $\mathbb{K}$-bilinearly on X and Y and that $[X, X] = 0$. One would like, then, to prove the Jacobi identity by intrinsic arguments not involving chart computations. For this, it is necessary to invoke the third order tangent bundle T^3M and the natural group structure on the space $\mathfrak{X}^3(M)$ of its sections (Theorem 28.2). Then the Jacobi identity can be proved in the same way as will be done for the Lie algebra of a Lie group in Section 24.4.

Essentially, this strategy of defining the Lie bracket and proving the Jacobi identity is the one used in synthetic differential geometry, and it also corresponds to the definition of the Lie algebra of a *formal group* (see [Se65]). In the framework of formal groups, the Jacobi identity is obtained by a third order computation from Hall's identity for iterated group commutators,

$$[[u, v], w^u][[w, u], v^w][[v, w], u^v] = 1,$$

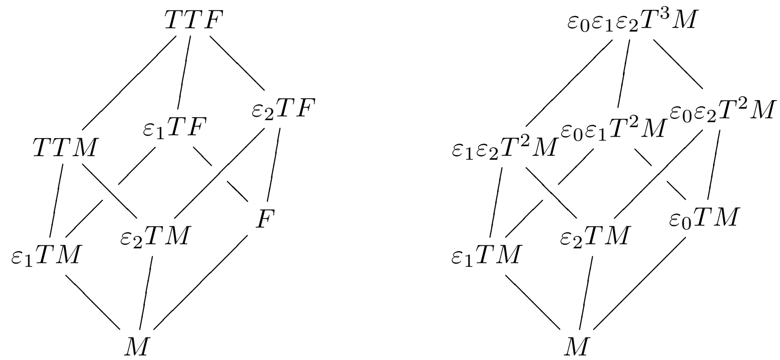
(here, $x^y = yxy^{-1}$) which is valid in any group (cf. [Se65, p. LG 4.18]). Compared to our framework, this proof rather corresponds to using the bundle J^3M which is more complicated than T^3M since it does not have "axes". Similar arguments are used for the proof of the Jacobi identity in synthetic differential geometry, cf. [Lav87, p. 74 ff] and [MR91, p. 187/88].

IV. Third and higher order differential geometry

Having linearized the bundle TF over M by means of a linear connection, our plan of attack is to linearize in a similar way all higher tangent bundles T^kF over M. It will turn out that there is an essentially canonical procedure to do so, starting with a connection on F and one on TM. The procedure is canonical up to a permutation acting on T^kF; in particular, for $k=2$, we get two versions of this procedure. Their difference can be interpreted as the curvature of L.

15. The structure of T^kF: Multilinear bundles

15.1. *Higher order models of F: projection and injection cube.* For a vector bundle $p: F \to M$, we consider the fiber bundle $T^kF \to M$ as a "k-th order model of F." The case $F = TM$ deserves special attention: $T^kF = T^{k+1}M$ may be seen as the model of order $k+1$ of M. The bundle T^kF comes with a natural "$k+1$-dimensional cube of projections": e.g., for TTF, resp. for T^3M, this is the commutative diagram of projections



(15.1)

where the notation $\varepsilon_i TM$ and $\varepsilon_i TF$ has the same meaning as $T_{\varepsilon_i}M$ and $T_{\varepsilon_i}F$ in Section 7.4. In case $F = TM$ we write also $F = \varepsilon_0 TM$. Each vertice of the cube is the projection of a vector bundle onto its base, but the composition of more than one of such projections is just the projection of a fiber bundle onto its base. In particular, T^kF is a fiber bundle over M. The fiber over $x \in M$ will be denoted by $(T^kF)_x$. For general k, the cube for $T^{k+1}M$ will have vertices $\varepsilon^\alpha T^{|\alpha|}M$ (notation for multi-indices $\alpha \in \{0,1\}^{k+1}$ being as in Section 7.4, resp. MA.1), and the cube for T^kF will have vertices $\varepsilon^\alpha T^{|\alpha|}M$ and $\varepsilon^\alpha T^{|\alpha|}F$ with $\alpha \in \{0,1\}^k$; it contains the k-dimensional cube for T^kM as a sub-cube. Diagram (15.1) can also be interpreted as a diagram of zero sections, just by reading the vertices as upward arrows. In particular, inclusion of spaces of the second row from below in the top space defines $k+1$ injections

$$\varepsilon_i TM \to T^kF \quad (i = 1, \ldots, k), \qquad F \to T^kF \tag{15.2}$$

which, for $x \in M$ fixed, define $k+1$ "basic axes" in $(T^kF)_x$ which will be denoted by $\varepsilon_i T_x M$, F_x. Altogether there are $2^{k+1}-1$ "axes in $(T^kF)_x$" whose definition will be given below.

15.2. *The axes-bundle.* For a multi-index $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_k) \in I_{k+1} = \{0,1\}^{k+1}$, we denote by

$$F^\alpha := \begin{cases} \varepsilon^\alpha F & \text{if} \quad \alpha_0 = 1 \\ \varepsilon^\alpha TM & \text{if} \quad \alpha_0 = 0 \end{cases} \tag{15.3}$$

a copy of F, resp. of TM (for the moment, ε^α is just a label to distinguish the various copies of F and of TM) and we define the *axes-bundle* to be the direct sum (over M) of all of these copies of F and of TM:

$$A^kF := \bigoplus_{\substack{\alpha \in I_{k+1} \\ \alpha > 0}} F^\alpha \tag{15.4}$$

We let also

$$A^{k+1}M := \bigoplus_{\substack{\alpha \in I_{k+1} \\ \alpha > 0}} \varepsilon^\alpha TM. \tag{15.5}$$

These are direct sums of vector bundles over M and hence are itself a vector bundles over M, in contrast to T^kF and to $T^{k+1}M$ which do not carry a canonical vector bundle structure.

15.3. *Isomorphism $T^kF \cong A^kF$ over a chart domain.* Assume $\varphi_i : M \supset U_i \to V$ is a chart such that $\varphi_i(x) = 0$. Then $T^k\varphi_i : T^kM \supset T^kU_i \to T^kV$ is a bundle chart of T^kM defining a bijection of fibers

$$(\Phi_i)_x := (T^k\varphi_i)_x : (T^kM)_x \to (T^kV)_0.$$

Taking another chart φ_j with $\varphi_j(x) = 0$, we have the following commutative diagram of bijections, describing the transition functions of the atlas of T^kM induced from the one of M:

$$\begin{array}{ccccc} & & (T^kM)_x & & \\ & {\scriptstyle (T^k\varphi_j)_x} \swarrow & & \searrow {\scriptstyle (T^k\varphi_i)_x} & \\ (T^kV)_0 & & \xrightarrow{(T^k\varphi_{ij})_0} & & (T^kV)_0 \end{array} \tag{15.6}$$

The transition functions of T^kM,

$$(\Phi_{ij})_0 := (T^k\varphi_{ij})_0 : (T^kV)_0 = \oplus_{\alpha>0}\varepsilon^\alpha V_\alpha \to \oplus_{\alpha>0}\varepsilon^\alpha V_\alpha$$

are not linear: they are given by the Tangent Map Formula (Theorem 7.5), with $f = \varphi_{ij}$, $f(0) = 0$,

$$T^kf(0 + \sum_\alpha \varepsilon^\alpha v_\alpha) = \sum_\alpha \varepsilon^\alpha \sum_{\ell=1}^{|\alpha|} \sum_{\lambda \in \mathcal{P}_\ell(\alpha)} b^\lambda(v_{\lambda^1}, \dots v_{\lambda^\ell})$$

where, for a partition λ of α of length ℓ, we have written b^λ for the multilinear map $d^\ell f(0)$. This shows that $(\Phi_{ij})_0$ is a *multilinear map* in the sense of Section

MA.5. Thus transition functions of T^kM are multilinear maps, and T^kM is a *multilinear bundle* in the following sense:

15.4. *Multilinear bundles.* A *multilinear bundle (with base M, of degree k)* is a fiber bundle E over M together with a family $(E_\alpha)_{\alpha \in I_k}$ of vector bundles and a bundle atlas such that, in a bundle chart around $x \in M$, the fiber E_x carries a structure of a multilinear space over $\mathbb{K}$ whose axes are given by natural inclusions $E_\alpha \to E$, and change of charts, restricted to fibers, is by isomorphisms of multilinear spaces. For $k = 1$, a multilinear bundle is simply a vector bundle, and for $k = 2$ it is bilinear bundle in the sense of Section 9.6.

Theorem 15.5. *If F is a vector bundle over M, then the iterated tangent bundle T^kF carries a canonical structure of a multilinear bundle over M. In particular, for all $p \in M$, the general multilinear group $\mathrm{Gm}^{0,k+1}((T^kF)_p)$ of the fiber over p is intrinsically defined, and there are canonical inclusions of axes $F^\alpha \subset T^kF$.*

Proof. The arguments are essentially the same as the ones used above in the special case $F = TM$: bundle charts of F are of the form $\widetilde{g}_i : F \supset \widetilde{U}_i \to V \times W$, with transition functions $\widetilde{g}_{ij}(x,w) = (\varphi_{ij}(x), g_{ij}(x)w)$, where $g_{ij}(x) : W \to W$ is a linear map. We assume $\varphi_i(p) = \varphi_j(p) = 0$, whence $\varphi_{ij}(0) = 0$. Bundle charts of T^kF are of the form $T^k\widetilde{g}_{ij}$, and as an analog of (15.6) we have the following commutative diagram of bijections:

$$\begin{array}{ccccc} & & (T^kF)_p & & \\ & {\scriptstyle (T^k\widetilde{g}_j)_p} \swarrow & & \searrow {\scriptstyle (T^k\widetilde{g}_i)_p} & \\ (T^kV)_0 \times T^kW & & \xrightarrow{(T^k\widetilde{g}_{ij})_0} & & (T^kV)_0 \times T^kW. \end{array} \tag{15.7}$$

For $k = 1$, the chart expression has been calculated in Section 9.1. In order to give the general chart expression, let $\widetilde{U} \subset F$ be a bundle chart domain with base $U \subset M$ and let W be the model space for the fibers of F. Then $T^k(\widetilde{U}) \subset T^kF$ is a bundle chart domain for T^kF, and elements of T^kF_p are represented in the form

$$u = \sum_{\alpha > 0} \varepsilon^\alpha v_\alpha \qquad \varepsilon^\alpha v_\alpha \in V_\alpha = \begin{cases} \varepsilon^\alpha W & \text{if} \quad \alpha_0 = 1 \\ \varepsilon^\alpha V & \text{if} \quad \alpha_0 = 0. \end{cases} \tag{15.8}$$

For $k = 2$, we may also use the following two equivalent notations (cf. Chapter 7):

$$\begin{aligned} u &= \varepsilon^{001} w_{001} + \varepsilon^{010} v_{010} + \varepsilon^{011} w_{011} + \varepsilon^{100} v_{100} + \varepsilon^{101} w_{101} + \varepsilon^{110} v_{110} + \varepsilon^{111} w_{111} \\ &= \varepsilon_0 w_0 + \varepsilon_1 v_1 + \varepsilon_0\varepsilon_1 w_{01} + \varepsilon_2 v_2 + \varepsilon_0\varepsilon_2 w_{02} + \varepsilon_1\varepsilon_2 v_{12} + \varepsilon_0\varepsilon_1\varepsilon_2 w_{012} \end{aligned} \tag{15.9}$$

where we replaced the letter v by w whenever the argument belongs to W. Now, the general formula for the transition functions is gotten from Theorem 7.5: we let

$$\begin{aligned} c_x^m &:= d^m \varphi_{ij}(x), \\ b_x^{m+1}(w, v_1, \ldots, v_m) &:= \ldots := b_x^{m+1}(v_1, \ldots, v_m, w) := \partial_{v_1} \cdots \partial_{v_m} g_{ij}(x, w). \end{aligned}$$

For $x=p$, we get

$$\begin{aligned} T^k\widetilde{g}_{ij}(\sum_{\alpha_0=1}\varepsilon^\alpha w_\alpha + \sum_{\alpha_0=0}\varepsilon^\alpha v_\alpha) &= \sum_{\alpha_0=1}\varepsilon^\alpha g_{ij}(0)w_\alpha + \sum_{\alpha_0=0}\varepsilon^\alpha d\varphi_{ij}(0)v_\alpha \\ &+ \sum_{m=2}^{k+1}\Big(\sum_{\substack{\alpha_0=1\\|\alpha|=m}}\varepsilon^\alpha\sum_{\ell=2}^{m}\sum_{\substack{\beta^1+\ldots+\beta^\ell=\alpha\\ \beta^1<\ldots<\beta^\ell}} b_0^\ell(w_{\beta^0},v_{\beta^1},\ldots,v_{\beta^\ell}) \\ &+ \sum_{\substack{\alpha_0=0\\|\alpha|=m}}\varepsilon^\alpha\sum_{\ell=2}^{m}\sum_{\substack{\beta^1+\ldots+\beta^\ell=\alpha\\ \beta^1<\ldots<\beta^\ell}} c_0^\ell(v_{\beta^1},\ldots,v_{\beta^\ell})\Big) \end{aligned} \tag{15.10}$$

(For $m=k+1$, only terms of the first sum over α actually appear, and if $V=TM$, then this is the expression for $T^{k+1}\varphi_{ij}$, with ε_0 being an additional infinitesimal unit.) Since $g_{ij}(0)$ and $d\varphi_{ij}(0)$ are invertible linear maps, comparing with Section MA.5, we see that $T^k\widetilde{g}_{ij}$ belongs to the general multilinear group $\mathrm{Gm}^{0,k+1}(E)$ of the fiber $E=\sum_{\alpha>0}V_\alpha$. From this our claim follows in the way already described in Section 9.3. ∎

As for any multilinear space, the *special multilinear group* $\mathrm{Gm}^{1,k+1}(E)$ of the fiber E is now well-defined. In terms of the preceding proof this means that, for a fixed point $p\in M$, we may, "without loss of generality", reduce to the case where $g_{ij}(0)=\mathrm{id}_W$ and $d\varphi_{ij}(0)=\mathrm{id}_V$. The special multilinear group will play the most important rôle in the sequel since it is the group acting simply transitively on the space of multilinearly related linear structures (cf. Section MA.7, MA.9). More generally, all the groups $\mathrm{Gm}^{j,k+1}((T^k)_x)$ with $j=0,1,\ldots,k+1$ (see Section MA.5 and Theorem MA.6) are well-defined.

16. The structure of T^kF: Multilinear connections

16.1. *Definition of multilinear connections and their homomorphisms.* Recall that a linear structure L on T^kF over M is simply given by two smooth structure maps, $T^kF\times_M T^kF \to T^kF$ and $\mathbb{K}\times T^kF \to T^kF$, defining a $\mathbb{K}$-module structure on each fiber of T^kF over M. A *multilinear connection on* T^kF is a linear structure L that is *multilinearly related* (i.e. conjugate under $\mathrm{Gm}^{0,k}(T^kF)_x$) to all linear structures induced from charts. (As in case $k=2$ it is clear from Theorem 15.4 that this notation is well-defined.) Assume K is a multilinear connection on T^kF and L a multilinear connection on T^kG and $f: F \to G$ a vector bundle homomorphism. We say that f is a K-L *homomorphism* if $T^kf : T^kF \to T^kG$ is fiberwise linear (with respect to K_x and $L_{f(x)}$). Clearly, the usual functorial rules hold, and thus vector bundles with multilinear connections (of a fixed order k) form a category. For $k=0$ this is just the category of vector bundles, and for $k=1$ it is the category of vector bundles with linear connection.

Theorem 16.2. (The generalized Dombrowski Splitting Theorem.) *For any vector bundle F over M, the following objects are in bijection to each other:*

(1) *multilinear connections L on T^kF*
(2) *isomorphisms $\Phi : A^kF \to T^kF$ of multilinear bundles over M such that the restriction of Φ to each axis F^α is the identity map on this axis (here and in the sequel, the corresponding axes F^α of A^kF and of T^kF are identified with each other).*

The bijection is described as follows: given Φ, the linear structure $L = L^\Phi$ is just the push-forward of the canonical linear structure L_0 on A^kF by Φ . Given L, we let

$$\Phi := \Phi^L : A^kF = \bigoplus_{\alpha>0} F^\alpha \to T^kF, \quad (x; v_\alpha)_{\alpha>0} \mapsto \sum_{\alpha>0} v_\alpha .$$

In this formula, the sum on the right hand side is taken in the fiber $(T^kF)_x$ with respect to the linear structure L_x.

Proof. Assume first that $\Phi : A^kF \to T^kF$ is any isomorphism of multilinear bundles over M. By definition of an isomorphism, the push-forward L^Φ of L_0 is then a multilinear connection on T^kF.

Conversely, assume a multilinear connection L on T^kF be given and define $\Phi = \Phi^L$ as in the claim. In order to prove that Φ is a diffeomorphism, we need the chart representation of multilinearly related linear structures: with respect to a fixed chart, both the fiber $E := (T^kF)_x$ and the fiber $(A^kF)_x$ are represented in the form $E = \oplus_{\alpha>0} V_\alpha$ with V_α as in Equation (15.8), V being the model space for the manifold M and W the one for the fiber F_x. Since A^kF is a vector bundle, the linear structure of the fiber agrees with the one induced from the chart, and the linear structure induced by L is multilinearly related to this linear structure: it is obtained by push-forward via a map f_b which is the chart representation of Φ, where f_b is defined as in Section MA.5: $b = (b_\lambda)_{\lambda\in\mathrm{Part}(I)}$ is a family of multilinear maps

$$b^\lambda := b_x^\lambda : V_{\lambda^{(1)}} \times \ldots \times V_{\lambda^{\ell(\lambda)}} \to V_\alpha, \tag{16.1}$$

λ is a partition of α, and $V_\beta = \varepsilon^\beta V$ if $\beta_0 = 0$, $V_\beta = \varepsilon^\beta W$ if $\beta_0 = 1$, and for $v = \sum_{\alpha>0} v_\alpha$,

$$f_b(v) = v + \sum_{|\alpha|\geq 2} \sum_{\substack{\lambda\in\mathcal{P}(\alpha)\\ \ell(\lambda)\geq 2}} b^\lambda(v_{\lambda^{(1)}}, \ldots, v_{\lambda^{\ell(\lambda)}}). \tag{16.2}$$

(The map $f_b = f_{b_x}$, resp. their components $(b^\lambda) = (b^\lambda_x)$, will be called the *Christoffel tensors of L in the given chart*; for $k = 1$, there is just one component, and it agrees with the classical Christoffel symbols of a connection, cf. Section 10.4.) We claim that the linear structure L is smooth if and only if, for all bundle charts, the Christoffel tensors b depend smoothly on x in the sense that

$$x \mapsto b^\lambda_x(v_{\lambda^{(1)}}, \ldots, v_{\lambda^{\ell(\lambda)}})$$

is smooth for all choices of the $v_{\lambda^{(i)}}$. In fact, it is clear that the linear structure is smooth if it is related to chart structures via smooth f_b's, and conversely, f_b and hence b can be recovered from the vector addition: in a chart, addition of elements from the axes is given by $\sum^{(b)}_{\alpha>0} v_\alpha = f_b(v)$, where the supersript (b) in the first sum means "addition with respect to the linear structure given by b_x". Hence the map f_b defined by a smooth linear structure is also smooth. We have proved that Φ is smooth. It is bijective with smooth inverse: in fact, f_{b_x} belongs to the group $\mathrm{Gm}^{1,k}(E)$ and hence is invertible (Theorem MA.6), and its inverse is of the form f_{c_x} with another multilinear map c_x which can be calculated by the algorithm used in the proof of Theorem MA.6. Since this algorithm consists of a sequence of compositions of multilinear maps and multiplications by -1, $c_x(v)$ depends again smoothly on x, and hence Φ^{-1} is smooth.

Finally, it is immediate from the definitions that the constructions $L \mapsto \Phi^L$ and $\Phi \mapsto L^\Phi$ are inverse of each other. ■

The isomorphism $\Phi = \Phi^L$ is called the *linearization map of* L. In case $F = TM$ is the tangent bundle, we have $F^\alpha = \varepsilon^\alpha TM$ for all α, and the linearization map can be written

$$\Phi^L_k : \bigoplus_{\alpha\in I_k, \alpha>0} \varepsilon^\alpha TM \to T^k M.$$

16.3. *Multilinear differential calculus on manifolds with multilinear connection.* We can define a "matrix calculus" for multilinear maps between higher order tangent bundles T^kM, T^kN equipped with multilinear connections, simply by using fiberwise the purely algebraic "matrix calculus" from Section MA.12. Basically, the choice of a multilinear connection on T^kM is the analog of the choice of a basis in a (say, n-dimensional) vector space E; the splitting map $\Phi^L : A^kF \to T^kF$ corresponds to the induced isomorphism $\mathbb{K}^n \to E$, the "matrix" is the induced homorphism $\mathbb{K}^n \to \mathbb{K}^m$, individual "matrix coefficients" are defined by composing with the various base-depending projections $\mathbb{K}^m \to \mathbb{K}$ and injections $\mathbb{K} \to \mathbb{K}^n$, and finally there is a "matrix multiplication rule" describing the matrix coefficients of the matrix of a composition.

Let us describe these concepts in more detail. Assume K is a multilinear connection on T^kM and L a multilinear connection on T^kN (the case of general vector bundles is treated similarly), and $f : M \to N$ is an arbitrary smooth map. We define its *multilinear higher differentials (with respect to K and L)*

by $d_K^L f := (\Phi^L)^{-1} \circ T^k f \circ \Phi^K : A^k M \to A^k N$, i.e., by the following commuting diagram:

$$\begin{array}{ccc} T^k M & \stackrel{\Phi^K}{\leftarrow} & A^k M = \bigoplus_{\alpha>0} \varepsilon^\alpha TM \\ T^k f \downarrow & & \downarrow d_K^L f \\ T^k N & \stackrel{\Phi^L}{\leftarrow} & A^k N = \bigoplus_{\alpha>0} \varepsilon^\alpha TN \end{array} . \tag{16.3}$$

As for matrices of linear maps, we have the functorial rules

$$d_{L_2}^{L_3} g \circ d_{L_1}^{L_2} f = d_{L_1}^{L_3}(g \circ f), \quad d_L^L(\mathrm{id}) = \mathrm{id} \tag{16.4}$$

which are immediate from the definitions. In each fiber, the maps

$$(d_K^L)_x f : (A^k M)_x \to (A^k N)_{f(x)}$$

are homomorphisms of multilinear spaces (in the sense of MA.5) – in fact, with respect to chart structures, this is true by Theorem 7.5, and hence it is true with respect to any other structure that is multilinearly related to chart structures. It follows that the "matrix coefficients", indexed by partitions Λ and Ω,

$$\begin{aligned} d_x^\Lambda f := (d_K^L f)_x^\Lambda &:= \mathrm{pr}_\alpha \circ T^k f \circ \iota_\Lambda : \bigoplus_{i=1}^{\ell(\Lambda)} \varepsilon^{\Lambda^i} TM \to \varepsilon^\alpha TM \\ (d_K^L f)_x^{\Omega|\Lambda} &:= \mathrm{pr}_\Omega \circ d_K^L f \circ \iota_\Lambda : \bigoplus_{i=1}^{\ell(\Lambda)} \varepsilon^{\Lambda^i} TM \to \bigoplus_{i=1}^{\ell(\Omega)} \varepsilon^{\Omega^i} TN \end{aligned} \tag{16.5}$$

are fiberwise multilinear maps in the usual sense; here pr and ι denote the canonical projections and injections in the direct sum bundle $A^k N$ defined by

$$\mathrm{pr}_\Omega : A^k N \to \bigoplus_{i=1}^{\ell(\Omega)} \varepsilon^{\Omega^i} TN, \qquad \iota_\Lambda : \bigoplus_{i=1}^{\ell(\Lambda)} \varepsilon^{\Lambda^i} TM \to T^k M.$$

The whole map is recovered from the matrix entries via

$$T^k f(x + \sum_\alpha \varepsilon^\alpha v_\alpha) = f(x) + \sum_\alpha \varepsilon^\alpha \sum_{\Lambda \in \mathcal{P}(\alpha)} (d^\Lambda f)_x (v_{\Lambda^1}, \ldots, v_{\Lambda^\ell}), \tag{16.6}$$

where the sum on the left-hand side is taken with respect to the linear structure K, and on the right-hand side with respect to L. Homomorphisms of multilinear connections in the sense of 16.1 are characterized by the property that the linearization maps commute with the natural action $f_* : A^k M \to A^k N$, $\sum_\alpha v_\alpha \mapsto \sum_\alpha Tf(v_\alpha)$:

$$\begin{array}{ccc} T^k M & \stackrel{\Phi^K}{\leftarrow} & A^k M = \bigoplus_{\alpha>0} \varepsilon^\alpha TM \\ T^k f \downarrow & & \downarrow f_* \\ T^k N & \stackrel{\Phi^L}{\leftarrow} & A^k N = \bigoplus_{\alpha>0} \varepsilon^\alpha TN \end{array} . \tag{16.7}$$

This, in turn, means that the matrix $d_K^L f$ is a "diagonal matrix", i.e. all components $(d_K^L f)_x^\Lambda$ for $\ell(\Lambda) > 1$ vanish, and for $\ell(\Lambda) = 1$ they agree with Tf.

16.4. *Composition rules and "Matrix multiplication".* Let us return to Equation (16.5). Assume $g : M_1 \to M_2$ and $f : M_2 \to M_3$ are smooth maps and L^i, for $i = 1,2,3$, are multilinear connections on $T^k M_i$. The "matrix coefficients" of $f \circ g$ with respect to L^3 and L^1 can be calculated from those of f with respect to L^3, L^2 and those of g with respect to L^2, L^1 by using the composition rule for multilinear maps (MA.19), resp. Prop. MA.11:

$$(d^{L_1}_{L_3}(f \circ g))^\Lambda = \sum_{\Omega \preceq \Lambda} (d^{L_1}_{L_2} f)^\Omega \circ (d^{L_2}_{L_3} g)^{\Omega|\Lambda}. \tag{16.8}$$

The notation d^Λ generalizes the usual differential d since, taking $M = V$ with the flat connection of a linear space, the matrix coefficient $d^\Lambda f$ as defined here coincides with the usual higher differential $d^\ell f$ with ℓ being the length of Λ, cf. Theorem 7.5.

16.5. *On existence of multilinear connections.* For $k > 1$, the space of all multilinear connections on $T^k F$ is no longer an affine space over $\mathbb{K}$ in any natural way (since $\mathrm{Gm}^{1,k}(E)$ is no longer a vector group), and hence it is more difficult to construct multilinear connections than to construct linear ones. They may be defined in charts by brute force by choosing arbitrary "Christoffel symbols" b^Λ which then have to behave under chart changes according to the composition rule in the general multilinear group (16.8). In the sequel we will see that there are canonical procedures to construct multilinear connections out of linear ones.

17. Construction of multilinear connections

17.1. *Deriving linear structures.* Assume $L = (a, m)$ is a linear structure on a fiber bundle E over M, with structural maps $a : E \times_M E \to E$, $m : \mathbb{K} \times E \to E$. Then $TL = (Ta, Tm)$ defines a linear structure on TE over TM, called the *derived linear structure* (this is even a linear structure with respect to $T\mathbb{K}$, but we will only consider it as a linear structure with respect to $\mathbb{K}$). However, in general there is no canonical way to see the derived linear structure TL as a linear structure on the bundle TE over the smaller base space M.

17.2. *Deriving multilinear connections.* Assume we are given the following data: a multilinear connection L on T^kF with corresponding isomorphism of linear bundles $\Phi : A^kF = \bigoplus_{\alpha>0} F^\alpha \to T^kF$, as well as linear connections L' on TM and L'' on F. Then we can construct, in a canonical way, an isomorphism $T^{k+1}F \cong A^{k+1}F$. This is obtained in two steps:

(a) $T^{k+1}F \cong T(A^kF)$. This obtained by deriving Φ, which gives a bundle isomorphism

$$T\Phi : T(\bigoplus_{\alpha>0} F^\alpha) = T(A^kF) \to T(T^kF) = T^{k+1}F.$$

(a) $T(A^kF) \cong A^{k+1}F$. Since A^kF is a direct sum of certain copies of F and of TM, we can equip A^kF with the direct sum of the connections L' (on the copies of TM) and L'' (on the copies of F) – see Section 10.10. This gives us a linear structure on $T(A^kF)$ and an isomorphism of linear bundles $T(A^kF) \cong A^{k+1}F$.

Putting (a) and (b) together, we end up with a bundle isomorphism $T^{k+1}F \cong A^{k+1}F$ which, by transport of structure, can be used to define a linear structure on $T^{k+1}F$. Let us make this isomorphism more explicit: in the following equalities, we use first the isomorphism $T\Phi$, then the canonical isomorphism (10.8) $T(A \oplus_M B) \cong TA \oplus_{TM} TB$, the isomorphisms Φ_1 corresponding to L' and to L'' and finally the canonical isomorphism (10.9) $(A \oplus_M TM) \oplus_{TM} (B \oplus_M TM) \cong A \oplus_M TM \oplus_M B$. By $\oplus^{TM}$ we denote the direct sum of vector bundles over TM and by $\oplus = \oplus_M$ the direct sum of vector bundles over M, and $e_1, \ldots, e_k$ is the canonical basis of $I_k = \{0,1\}^k$:

$$\begin{aligned}
T^{k+1}F \overset{T\Phi}{\cong} T(\bigoplus_{\substack{\alpha\in I_k\\ \alpha>0}} F^\alpha) &\overset{(10.8)}{\cong} \bigoplus_{\substack{\alpha\in I_k\\ \alpha>0}}^{TM} TF^\alpha \\
&\overset{\Phi_1}{\cong} \bigoplus_{\substack{\alpha\in I_k\\ \alpha>0}}^{TM} (F^\alpha \oplus \varepsilon_{k+1}TM \oplus \varepsilon_{k+1}F^\alpha) \\
&\overset{(10.9)}{\cong} \varepsilon_{k+1}TM \oplus \bigoplus_{\substack{\alpha\in I_k\\ \alpha>0}} F^\alpha \oplus \bigoplus_{\substack{\alpha\in I_k\\ \alpha>0}} \varepsilon_{k+1}F^\alpha \overset{\text{def.}}{\cong} \bigoplus_{\substack{\alpha\in I_{k+1}\\ \alpha>0}} F^\alpha.
\end{aligned} \tag{17.1}$$

Let us write $D\Phi$ for the isomorphism (17.1) of bundles over M. Suppressing the canonical isomorphisms (10.8) and (10.9), we may write

$$D\Phi = T\Phi \circ (\times_{\alpha \in I_k} \Phi_1^{e_{k+1},\alpha}) : \bigoplus_{\substack{\alpha \in I_{k+1} \\ \alpha > 0}} F^\alpha \to T^{k+1}F, \tag{17.2}$$

where $\Phi_1^{e_{k+1},\alpha}$ is the splitting isomorphism Φ_1 for TF^α. By transport of structure, $D\Phi$ defines on $T^{k+1}F$ the structure of a linear bundle over M which we denote by DL.

Theorem 17.3. *The linear structure DL defined on $T^{k+1}F$ by $D\Phi$ is a multilinear connection on $T^{k+1}F$.*

Proof. As seen in the proof of the Generalized Dombrowski Splitting Theorem 16.2, in a chart, Φ is given by (with $v = \sum_{\substack{\alpha \in I_k \\ \alpha > 0}} \varepsilon^\alpha v_\alpha$)

$$\begin{aligned} \Phi(x+v) &= x + v + \sum_{\substack{\alpha \in I_k \\ \alpha > 0}} \varepsilon^\alpha \sum_{\substack{\lambda \in \mathcal{P}(\alpha) \\ \ell(\lambda) \geq 2}} b_x^\lambda(v_{\lambda^1}, \ldots, v_{\lambda^{\ell(\lambda)}}) \\ &= x + \sum_{\substack{\alpha \in I_k \\ \alpha > 0}} \varepsilon^\alpha \sum_{\lambda \in \mathcal{P}(\alpha)} b_x^\lambda(v_{\lambda^1}, \ldots, v_{\lambda^{\ell(\lambda)}}) \end{aligned}$$

with multilinear maps b_x^λ, depending smoothly on x, and $b_x^\lambda(u) = u$ if $\ell(\lambda) = 1$. The tangent map $T\Phi$ is calculated by taking ε_{k+1} as new infinitesimal unit:

$$\begin{aligned} T\Phi(x + \sum_{\substack{\alpha \in I_{k+1} \\ \alpha > 0}} \varepsilon^\alpha v_\alpha) &= \Phi(x + \sum_{\substack{\alpha \in I_k \\ \alpha > 0}} \varepsilon^\alpha v_\alpha) + \\ &\qquad \varepsilon^{k+1}\Big(d\Phi(x + \sum_{\substack{\alpha \in I_k \\ \alpha > 0}} \varepsilon^\alpha v_\alpha) \cdot \sum_{\alpha \in I_k} \varepsilon^\alpha v_{\alpha + e_{k+1}} \Big) \\ &= x + \sum_{\substack{\alpha \in I_k \\ \alpha > 0}} \varepsilon^\alpha \sum_{\lambda \in \mathcal{P}(\alpha)} b_x^\lambda(v_{\lambda^1}, \ldots, v_{\lambda^{\ell(\lambda)}}) + \\ &\qquad \sum_{\substack{\alpha \in I_k \\ \alpha > 0}} \varepsilon^{\alpha + e_{k+1}} \sum_{\lambda \in \mathcal{P}(\alpha)} \Big(\sum_{i=1}^{\ell(\lambda)} b_x^\lambda(v_{\lambda^1}, \ldots, v_{\lambda^i + e_{k+1}}, \ldots, v_{\lambda^{\ell(\lambda)}}) \\ &\qquad\qquad + \partial_{v_{e_{k+1}}} b_x^\lambda(v_{\lambda^1}, \ldots, v_{\lambda^{\ell(\lambda)}}) \Big) \end{aligned}$$

(the partial derivative in the last term is with respect to x; the dependence on the other variables is linear and hence deriving amounts to replace the argument v_{λ^i} by the direction vector $v_{\lambda^i + e_{k+1}}$). Next, we have to insert the chart representation of the isomorphisms Φ_1: we have to replace, for all $\beta \in I_k$, $v_{\beta + e_{k+1}}$ by $v_{\beta + e_{k+1}} + d_x(v_\beta, v_{e_{k+1}})$ where, if $\beta_0 = 0$, d_x is the Christoffel tensor of the connection L' on TM, and, if $\beta_0 = 1$, d_x is the Christoffel tensor of the connection L'' on F. The resulting formula for $D\Phi$ is:

$$D\Phi(x+\sum_{\substack{\alpha\in I_{k+1}\\ \alpha>0}}\varepsilon^\alpha v_\alpha) = x+\sum_{\substack{\alpha\in I_k\\ \alpha>0}}\varepsilon^\alpha\sum_{\lambda\in\mathcal{P}(\alpha)} b_x^\lambda(v_{\lambda^1},\dots,v_{\lambda^{\ell(\lambda)}})+$$
$$\sum_{\substack{\alpha\in I_k\\ \alpha>0}}\varepsilon^{\alpha+e_{k+1}}\sum_{\lambda\in\mathcal{P}(\alpha)}\Big(\sum_{i=1}^{\ell(\lambda)} b_x^\lambda(v_{\lambda^1},\dots,v_{\lambda^i+e_{k+1}},\dots,v_{\lambda^{\ell(\lambda)}})+ \tag{17.3}$$
$$b_x^\lambda(v_{\lambda^1},\dots,d_x(v_{\lambda_i},v_{e_{k+1}}),\dots,v_{\lambda^{\ell(\lambda)}})+\partial_{v_{e_{k+1}}}b_x^\lambda(v_{\lambda^1},\dots,v_{\lambda^{\ell(\lambda)}})\Big)$$

This expression is again a sum of multilinear terms and hence defines a new multilinear map $f_B=f_{B_x}$ where B_x depends smoothly on x. Thus DL is a multilinear connection on $T^{k+1}F$. ∎

For later use, let us spell out the formula in case $k=2$, writing $b_x(w,v)$ for the Christoffel tensor of L and $c_x(u,v)$ for the one of L' in the chart, und using the more traditional notation $v=\varepsilon_0 w_0+\varepsilon_1 v_1+\varepsilon_2 v_2+\varepsilon_0\varepsilon_1 w_{01}+\varepsilon_0\varepsilon_2 w_{02}+\varepsilon_1\varepsilon_2 v_{12}+\varepsilon_0\varepsilon_1\varepsilon_2 w_{012}$ (see Eqn. (15.9)) for elements in the fiber $(TTF)_x$:

$$D\Phi_1(x,v)=v+\varepsilon_0\varepsilon_1 b_x(w_0,v_1)+\varepsilon_0\varepsilon_2 b_x(w_0,v_2)+\varepsilon_1\varepsilon_2 c_x(v_1,v_2)+$$
$$\varepsilon_0\varepsilon_1\varepsilon_2\big(b_x(w_0,v_{12})+b_x(w_{01},v_2)+b_x(w_{02},v_1)+ \tag{17.4}$$
$$b_x(w_0,c_x(v_1,v_2))+b_x(b_x(w_0,v_2),v_1)+\partial_{v_2}b_x(w_0,v_1)\big).$$

If $F=TM$, then this formula describes a linear structure on T^3M, depending on two linear connections L,L' on TM. In this case, unless otherwise stated, we will choose $L=L'$, i.e. $b_x=c_x$.

17.4. The preceding construction has the property that the zero section $z:T^kF\to T^{k+1}F=T_{\varepsilon_{k+1}}T^kF$ is a L-DL-linear homomorphism of linear bundles over M; more precisely, by construction L and DL agree on the image of the zero section:

$$\begin{array}{ccc} T^kF & \stackrel{z}{\to} & T^{k+1}F \\ \Phi\uparrow & & \uparrow D\Phi \\ A^kF & \subset & A^{k+1}F \end{array}$$

(this may also be seen from the chart representation (17.4), where the image of the zero section corresponds to points v with $v_{\beta+e_{k+1}}=0$ for all β.)

17.5. *The sequence of derived linear structures.* Assume L is a linear connection on the vector bundle $p:F\to M$ and L' a linear connection on TM. Then, using the preceding construction, we can derive L and construct a connection $L_2:=DL$ on TTF over M, and so on: we get a sequence of derived linear structures $L_k=D^{k-1}L$ on T^kF over M, $k=2,3,\dots$ with corresponding isomorphisms $\Phi_k=D^{k-1}\Phi$ of linear bundles over M. In case of tangent bundles, $F=TM$, then (unless otherwise stated) we will take $L=L'$ in this construction, and the construction is functorial in the following sense: if $f:M\to N$ is affine for given connections L on TM, L' on TN, then $T^kf:T^kM\to T^kN$ is an L_k-L_k' homomorphism in the sense of 16.1 (and conversely). This is seen by an easy induction using that the construction involves only natural isomorphisms and the tangent functor. For instance, $\gamma:\mathbb{K}\to M$

is a geodesic if and only if, for some $k > 1$, $T^k\gamma : T^k\mathbb{K} \to T^kM$ is a L_k-L'_k homomorphism, where L_k is the natural connection on $\mathbb{K}$ given by the canonical identification $T^k\mathbb{K} = A^k\mathbb{K}$.

However, the "operator D of covariant derivative" has not the property of the usual derivative d that higher derivatives are independent of the order of the ε_i: in general, Φ_k and L_k will depend on the order $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_k$ of the subsequent scalar extensions leading from F to T^kF, and we denote by L_k^σ, $\sigma \in \Sigma_k$, the linear structure obtained with respect to the order $\varepsilon_{\sigma(1)}, \ldots, \varepsilon_{\sigma(k)}$. Then the push-forward of L_k by the canonical action of σ on T^kF is given by

$$\sigma \cdot L_k = L_k^\sigma. \tag{17.5}$$

These remarks make it necessary to study in more detail the behavior of the linear structures L_k with respect to the action of the symmetric group (see next chapter).

17.6. *Example: trivializable tangent bundles.* We will see later that the tangent bundle of a Lie group is always trivializable, and thus the following example can be applied to the case of an arbitrary Lie group: assume that the tangent bundle TM is trivializable over M, i.e. that there is a diffeomorphism $\Phi_0 : M \times V \to TM$ which is fiber-preserving and linear in fibers over M. We define a diffeomorphism $\Psi_1 : (M \times V) \times (V \times V) \to TTM$ by the following diagram:

$$\begin{array}{ccc} M \times V \times V \times V & \overset{\Psi_1}{\to} & TTM \\ \downarrow & & \uparrow T\Phi_0 \\ M \times V \times TV & \overset{\Phi_0 \times \mathrm{id}}{\to} & TM \times TV \end{array}$$

Then Ψ_1 is a trivialization of the bundle TTM over M. We can continue in this way and trivialize all higher tangent bundles T^kM over M via

$$\Psi_{k+1} := T^k\Phi_0 \circ (\Psi_k \times \mathrm{id}) = T\Psi_k \circ (\Phi_0 \times \mathrm{id}) : \quad M \times \bigoplus_{\substack{\alpha \in I_{k+1} \\ \alpha > 0}} \varepsilon^\alpha V \to T^{k+1}M.$$

Via push-forward, this defines linear structures L_k on all higher order tangent bundles T^kM. By our assumption on Φ_0, L_1 is just the vector bundle structure on TM.

Lemma 17.7. *Assume the tangent bundle TM is trivializable, and retain notation from* Section 17.6. *Then the linear structure L_2 is a connection on the tangent bundle TM, and $L_k = D^{k-2}L_2$ for $k > 2$, i.e., the linear structures L_k agree with the derived connections of L_2 defined in the preceding section.*

Proof. In order to prove that L_2 is a connection, in a bundle chart, Φ_0 is represented by $\Phi_0(x, w) = (x, g_x(w)) = (x, g(x, w))$; then

$$T\Phi_0((x, w) + \varepsilon(x', w')) = (x, g_x(w)) + \varepsilon\big(g_x(w'), d_1g(x, w)x'\big)$$

(cf. Section 9.1) where the last term depends bilinearly on (w, x'), proving that L_2 is a connection. Explicitly, the Christoffel tensor of this connection is given by Equation (9.3):

$$b_x(w, x') = g_x^{-1}d_1g(x, w)x'$$

(Note that the bilinear map b_x is in general not symmetric.) Next, it follows that $L_3 = DL_2$ since both L_3 and DL_2 are constructed by deriving L_2 and composing with Ψ_2, resp. with Φ_2, and Ψ_2 and Φ_2 induce the same linear structure L_2 on TTM. By induction, this argument shows that $L_k = D^{k-1}L_2$. ∎

18. Curvature

18.1. *Totally symmetric multilinear connections.* The symmetric group Σ_k acts on T^kF; in the notation (15.8) this action is by permuting the last k coordinates of $\{0,1\}^{k+1}$. If $F = TM$, then $T^kF = T^{k+1}M$, and ε_0 may be considered as another infinitesimal unit, and the permutation group Σ_{k+1} acts on all $k+1$ coordinates. We say that a multilinear connection L is *totally* or *very symmetric* if it is invariant under the action of Σ_k, resp. of Σ_{k+1}. Equivalently, this means that Σ_k acts linearly with respect to L, or that L is fiberwise conjugate to all chart-structures under the action of *very symmetric group* $\mathrm{Vsm}^{1,k}((T^kF)_x) = (\mathrm{Gm}^{1,k}(T^kF))^{\Sigma_k}$ defined in Section SA.2. In terms of the Dombrowski Splitting Map $\Phi : A^kF \to T^kF$, the condition is equivalent to saying that Φ is equivariant with respect to the natural Σ_k-actions on both sides.

18.2. *Curvature operators and curvature forms.* Curvature is the obstruction for a multilinear connection L to be very symmetric, or, in other words, it measures the difference between L and the linear structure $L^\sigma = \sigma.L$ obtained by push-forward via $\sigma \in \Sigma_k$. In order to formalize this idea, let us denote the canonical linear structure on the axes-bundle A^kF by L^0; it is Σ_k-invariant. The linear structure L is obtained as push-forward of L^0 via Φ, i.e. $L = \Phi.L^0$. It follows that $\sigma.L = \sigma.\Phi.L^0 = \sigma\Phi\sigma^{-1}.L^0$, and hence the linear structure L^σ is given by push-forward via $\Phi^\sigma = \sigma \circ \Phi \circ \sigma^{-1}$. It is proved in Prop. SA.3 that Φ^σ is again a multilinear map, and thus we see that $\sigma.L$ is again a multilinear connection. Now we define the *curvature operator* $R := R^\sigma := R^{\sigma,L}$ by

$$R^{\sigma,L} := \Phi^\sigma \circ \Phi^{-1} = \sigma \circ \Phi \circ \sigma^{-1} \circ \Phi^{-1} : T^kF \to T^kF, \tag{18.1}$$

respectively its action on the axes-bundle $\widetilde{R} := \widetilde{R}^\sigma := \widetilde{R}^{\sigma,L}$ given by

$$\widetilde{R}^{\sigma,L} := \Phi^{-1} \circ R \circ \Phi = \Phi^{-1} \circ \sigma \circ \Phi \circ \sigma^{-1} = \Phi^{-1} \circ \Phi^{\sigma^{-1}} : A^kF \to A^kF \tag{18.2}$$

Both operators are compositions of multilinear maps and hence are multilinear automorphisms of T^kF, resp. of A^kF. The curvature operator R can be characterized as the unique element of $\mathrm{Gm}^{1,k}(T^kF)$ relating the linear structures L and L^σ:

$$R \cdot L = L^\sigma. \tag{18.3}$$

In fact, $R \cdot L = \sigma \circ \Phi \circ \sigma^{-1} \circ \Phi^{-1} \cdot L = \sigma \circ \Phi \circ \sigma^{-1} \circ L_0 = \sigma \circ \Phi \cdot L_0 = \sigma \cdot L$. Immediately from the definition we get the following "cocycle relations" : for fixed L and for all $\sigma, \tau \in \Sigma_k$,

$$R^{\mathrm{id}} = \mathrm{id}, \quad R^{\sigma\tau} = \sigma R^\tau \sigma^{-1} \circ R^\sigma. \tag{18.4}$$

In fact, it is also possible to interprete the curvature operators as higher differentials ("matrices") in the sense of Section 16.3, and then the cocycle relations are seen as a special case of the matrix multiplication rule (16.5): from (18.2) it follows that

$$\begin{array}{rcl} T^kF & \overset{\Phi^{\sigma^{-1}}}{\leftarrow} & A^kF \\ T^k\,\mathrm{id}_F \downarrow & & \downarrow \widetilde{R} \\ T^kF & \overset{\Phi}{\leftarrow} & A^kF \end{array}$$

commutes, which shows that

$$\widetilde{R}^{\sigma} = d^L_{\sigma^{-1}.L}(\mathrm{id}_F) \tag{18.5}$$

is a higher differential in the sense of 16.3. The matrix coefficients of the curvature operators are called the *curvature forms*,

$$\widetilde{R}^{\Omega|\Lambda} := \mathrm{pr}_\Omega \circ \widetilde{R} \circ \iota_\Lambda : F_{\Lambda^1} \oplus_M \ldots \oplus_M F_{\Lambda^l} \to F_{\Omega^1} \oplus_M \ldots \oplus_M F_{\Omega^r}. \tag{18.6}$$

Of particular interest are the *elementary curvature forms*

$$\widetilde{R}^{\Lambda} := \widetilde{R}^{\underline{\Lambda}|\Lambda} : F_{\Lambda^1} \oplus_M \ldots \oplus_M F_{\Lambda^l} \to F_{\underline{\Lambda}}$$

which are ordinary tensor fields, i.e. fiberwise multilinear maps in the usual sense. For the longest partition $\Lambda = (e_1, \ldots, e_k)$ we get the *top curvature form.* In particular, for $k = 3$ one finds the classically known curvature tensor as a top curvature form:

Theorem 18.3. *Assume L is a linear connection on the vector bundle F over M and L' a torsionfree connection on TM. Let $k = 2$ and $R = R^{\kappa,DL}$ where $\kappa : T^2F \to T^2F$ is the canonical flip and DL is the connection derived from L as described in* Theorem 17.3. *Then R does not depend on the choice of L'. Moreover, R belongs to the subgroup* $\mathrm{Gm}^{2,3}(TTF)$*, i.e. $\widetilde{R}_x$ has only one non-trivial component, namely the top curvature form*

$$\widetilde{R}^{\Lambda} : \varepsilon_1 TM \times \varepsilon_2 TM \times \varepsilon_0 F \to \varepsilon_0\varepsilon_1\varepsilon_2.F$$

where $\Lambda = (e_1, e_2, e_3)$. In other words, for $v = (v_\alpha) \in (A^2F)_x$,

$$\widetilde{R}_x(v) = v + \widetilde{R}^{\Lambda}_x(\varepsilon_1 v_1, \varepsilon_2 v_2)\varepsilon_0 v_0.$$

In a chart, with Christoffel tensor b_x of L, the top curvature form $\widetilde{R}^{\Lambda}_x$ is given by

$$\widetilde{R}^{\Lambda}_x(v_1, v_2)w_0 = b_x(b_x(w_0, v_2), v_1) - b_x(b_x(w_0, v_1), v_2) + \partial_{v_2} b_x(w_0, v_1) - \partial_{v_1} b_x(w_0, v_2).$$

Proof. We use the chart representation (17.4) of the map $\Phi_2 = D\Phi$ for $L_2 = DL$: the chart formula for L_2^κ is obtained from this formula by exchanging $\varepsilon_1 \leftrightarrow \varepsilon_2$, $v_1 \leftrightarrow v_2$, $w_{01} \leftrightarrow w_{02}$. Note that the resulting formula differs from the one of Φ_2 only in the $\varepsilon_0\varepsilon_1\varepsilon_2$-component and in the $\varepsilon_1\varepsilon_2$-component, and the difference in the $\varepsilon_1\varepsilon_2$-component vanishes if $c(v_1, v_2) = c(v_2, v_1)$, i.e. if L' is torsionfree. Let us assume that this is the case. Then from Formula (17.4) we get the following chart formula for the difference between L_2 and L_2^κ:

$$\begin{aligned} &\Phi_2(v) - \Phi_2^\kappa(v) = \varepsilon_0\varepsilon_1\varepsilon_2 \cdot \\ &\Big(b_x(b_x(w_0, v_2), v_1) - b_x(b_x(w_0, v_1), v_2) + \partial_{v_2} b_x(w_0, v_1) - \partial_{v_1} b_x(w_0, v_2)\Big) \end{aligned} \tag{18.7}$$

since c_x is assumed to be symmetric. This formula proves all claims of the theorem: if we denote by $A_x(v_1, v_2)w_0$ the right-hand side of (18.7) and by f_A the multilinear map given by $f_A(v) = v + A(v_1, v_2)w_0$, then from the composition rule of multilinear maps it follows that $f_A \cdot L_2 = L_2^\kappa$; thus (18.3) is satisfied, and by uniqueness in (18.3) we get $R = f_A$, and A is the only non-trivial component of the curvature maps R, resp. $\widetilde{R}$. Formula (18.7) shows that it is indeed independent of c and hence of L'. ■

The chart representation (18.7) of the top curvature form coincides with the usual expression of the curvature tensor of a connection in a chart by its Christoffel tensor (cf. e.g. [La99, p. 232]), thus showing that our interpretation of the curvature agrees with all other known concepts.

18.4. *Case of the tangent bundle.* Assume we are in the case of the tangent bundle: $F = TM$; then we usually choose $L = L'$. If L has torsion, then also L' has torsion, and the assumptions of Theorem 18.3 are not satisfied. However, Formula (17.4) still can be used to establish a relation between the classical curvature tensor and the curvature forms of L; we will not go into the details and leave apart the topic of connections with torsion. Let us just add the remark that in this case the Σ_3-orbit of the linear structure L_3 on T^3M has 6 elements which we may represent by a hexagon. None of the 6 structures is invariant under a transposition, thus we may represent the three transpositions by symmetries with respect to three axes that do not pass through the vertices.

18.5. *Bianchi's identity.* Now assume that $F = TM$ is the tangent bundle and that L is torsionfree. Then, by torsionfreeness, L and DL are invariant under the transposition (12), and the usual curvature tensor is the top curvature form of $R = R^{(23)}$. Let us look at the Σ_3-orbit of the linear structure $K := DL$ on T^3M. This orbit contains at most 3 elements since K is stabilized by the transposition (12). We denote these elements by $K^1 := K = K^{\mathrm{id}} = K^{(12)}$, $K^2 := K^{(23)} = K^{(132)}$ and $K^3 := K^{(13)} = K^{(123)}$. From the "matrix multiplication rule" (16.5), we get

$$\mathrm{id} = d_{K^1}^{K^1}\,\mathrm{id} = d_{K^2}^{K^1}\,\mathrm{id} \circ d_{K^3}^{K^2}\,\mathrm{id} \circ d_{K^1}^{K^3}\,\mathrm{id}\,.$$

Taking the top curvature forms, we get the *Bianchi identity* for the curvature form $R = R^{(23)}$

$$R + R^{(132)} + R^{(123)} = 0.$$

In a similar way, we get the usual skew-symmetry of the curvature form $R = R^{(23)}$. More generally, one can show that the top degree curvature forms $R^{(23),D^kL}$ for the derived linear structures L_{k+1} agree with the covariant derivatives of the curvature tensor in the usual sense, and under suitable assumptions one can prove algebraic relations for these.

The following result is included for the sake of completeness. It relates the curvature as defined here to the usual definition in terms of covariant derivatives:

Theorem 18.6. *Assume ∇ is the covariant derivative associated to the linear connection L on the vector bundle F. Then, for all sections X, Y of TM and Z of F and all $x \in M$, the curvature is given by*

$$R_x(X(x), Y(x))Z(x) = ([\nabla_X, \nabla_Y]Z - \nabla_{[X,Y]}Z)(x).$$

In particular, the right hand side depends only on the values of X, Y, Z at x.

Proof. The right hand side has the following chart expression (the calculation given in [La99, p. 232] can be repeated verbatim):

$$b_x(b_x(Z,Y),X) - b_x(b_x(Z,X),Y) + \partial_Y b_x(Z,X) - \partial_X b_x(Z,Y) \tag{18.8}$$

where b_x is the Christoffel tensor of L. This coincides with the chart expression for the curvature R given by Eqn. (18.7). ∎

18.7. *Structure equations.* Assume L, L' are two connections on F with covariant derivatives ∇, ∇' and curvature tensors R, R'; let $A := L' - L = \nabla' - \nabla$ (tensor field of type $(2,1)$). For $X, Y \in \mathfrak{X}(M)$, we write $A_X(Y) := A(X,Y)$. Then the "new" curvature is calculated from the "old" one via

$$\begin{aligned} R'(X,Y) &= [(\nabla + A)_X, (\nabla + A)_Y] - (\nabla + A)_{[X,Y]} \\ &= R(X,Y) + [A_X, \nabla_Y] - [A_Y, \nabla_X] - A_{[X,Y]} + [A_X, A_Y]. \end{aligned}$$

Thus R' depends on A in a quadratic way. In particular, the set of flat connections can be seen as the zero set of a quadratic map. If we assume that $F = TM$ and that ∇ is torsionfree, then we can write $[\nabla_X, A_Y] - A_{\nabla_X Y} = (\nabla A)(X,Y)$, and the preceding equation reads

$$R'(X,Y) = R(X,Y) + (\nabla A)((X,Y) - (Y,X)) + [A_X, A_Y]. \tag{18.9}$$

If, moreover L was the flat connection of a chart domain, then $R = 0$, and the tensor $A : TM \times W \to W$ can be interpreted as an $\mathrm{End}(W)$-valued one-form (where W is the model space for the fibers of F). The exterior derivative of A is $\mathrm{d}A(X,Y) = (\nabla A)((X,Y) - (Y,X))$, which gives the second term. The third term is symbollically written $\frac{1}{2}[A,A]$, and we get

$$R' = \mathrm{d}A + \frac{1}{2}[A,A] \tag{18.10}$$

which is a version of the structure equations of E. Cartan. If also L' is flat, we get the condition $\mathrm{d}A + \frac{1}{2}[A,A] = 0$ which is another version of the structure equations.

19. Linear structures on jet bundles

19.1. *Jet-compatible multilinear connections.* We say that a multilinear connection L on T^kF is *weakly symmetric* or *jet-compatible* if the jet-bundle $J^kF = (T^kF)^{\Sigma^k}$ is a *linear* subbundle of the linear bundle (T^kF, L). Equivalently, L is conjugate to all chart-structures under the *symmetric multilinear group* $\mathrm{Sm}^{1,k}((T^kF)_x)$ defined in Section SA.2. In terms of the Dombrowski Splitting Map $\Phi : A^kF \to T^kF$, the condition is equivalent to saying that

$$\Phi((A^kF)^{\Sigma_k}) = (T^kF)^{\Sigma_k}.$$

A totally symmetric multilinear connection is weakly symmetric (since $\mathrm{Vsm}(E) \subset \mathrm{Sm}(E)$) but the converse is not true: for instance, every connection on TM is weakly symmetric since the chart representation of Φ is $\Phi(x; u, v, w) = (x, u, v, w + b_x(v, w))$, and this preserves the diagonal $\{(x, u, u, w)|x \in U, u, w \in V\}$. On the other hand, Φ is not always totally symmetric (it is totally symmetric if and only if b_x is a symmetric bilinear map). More generally, a connection may be weakly symmetric even if it has non-trivial curvature:

Theorem 19.2. *Assume L is a linear connection on F and L' a linear connection on TM. Then the derived multilinear connections $L_k = D^{k-1}L$ on T^kF are weakly symmetric.*

Proof. The claim is proved by induction on k. The case $k = 1$ is trivial since $\Sigma_1 = \{\mathrm{id}\}$.

The case $k = 2$ can be proved by using the explicit formula (17.4) for Φ_2 in a chart: the vector v is fixed under Σ_2 iff $v_1 = v_2$ and $w_{01} = w_{02}$; then (17.4) shows that also $\Phi(v)$ is fixed under Σ_2.

Now assume that the property $\Phi_k((A^kF)^{\Sigma_k}) = (T^kF)^{\Sigma_k}$ holds for the linear structure L_k on T^kF. In order to prove that $\Phi_{k+1}((A^{k+1}F)^{\Sigma_{k+1}}) \subset (T^{k+1}F)^{\Sigma_{k+1}}$, note first that the set $R = \Sigma_k \cup \{(k, k+1)\}$ generates Σ_{k+1}, where Σ_k is imbedded in Σ_{k+1} as the subgroup permuting $\varepsilon_1, \ldots, \varepsilon_k$. Thus the fixed point spaces are

$$(A^{k+1}F)^{\Sigma_{k+1}} = (A^{k+1}F)^{\Sigma_k} \cap (A^{k+1}F)^{(k+1,1)},$$
$$(T^{k+1}F)^{\Sigma_{k+1}} = (T^{k+1}F)^{\Sigma_k} \cap (T^{k+1}F)^{(k+1,1)},$$

and it suffices to show that

(a) $\Phi_{k+1}((A^{k+1}F)^{\Sigma_k}) = (T^{k+1}F)^{\Sigma_k}$ and
(b) $\Phi((A^{k+1}F)^{(k+1,k)}) = (T^{k+1}F)^{(k+1,k)}$.

In order to prove (a), recall from Formula (17.2) that Φ_{k+1} is defined as a composition of two maps: $\Phi_{k+1} = T\Phi_k \circ \oplus_\alpha \Phi_1^{e_{k+1},\alpha}$. We claim that both maps preserve the subspace of Σ_k-invariants, i.e., the following diagram commutes:

$$\begin{array}{ccccc} A^{k+1}F & \overset{\oplus_\alpha \Phi_1^{e_{k+1},\alpha}}{\to} & T(A^kF) & \overset{T\Phi_k}{\to} & T(T^kF) \\ \uparrow & & \uparrow & & \uparrow \\ (A^{k+1}F)^{\Sigma_k} & \overset{\oplus_\alpha \Phi_1^{e_{k+1},\alpha}}{\to} & (T(A^kF))^{\Sigma_k} & \overset{T\Phi_k}{\to} & T(T^kF)^{\Sigma_k} \end{array}$$

In fact, for the square on the right of the diagram this follows by applying the tangent functor to the induction hypothesis since the action of $\tau \in \Sigma_k$ on $T^{k+1}F$ is simply given by the tangent map $T\tau$ of the action on T^kF. The map on the left of the diagram commutes also with the Σ_k-action; this follows from the fact that Σ_k does not act on ε_{k+1}, hence simply interchanges the various copies of Φ_1 used in the construction, and it can also be directly seen from the chart representation

$$\oplus_\alpha \Phi_1^{e_{k+1},\alpha} : \sum_{\alpha \in I_{k+1}} \varepsilon^\alpha v_\alpha \mapsto \sum_{\alpha_{k+1}=0} \varepsilon^\alpha v_\alpha + \varepsilon_{k+1} v_{k+1} + \sum_{\alpha_{k+1}=0} \varepsilon^{\alpha+e_{k+1}} (v_{\alpha+e_{k+1}} + d_x(v_\alpha, v_{e_{k+1}})).$$

Next, we will show that the space of invariants of the transposition $(k, k+1)$ is preserved. In order to prove this, we will use another representation of the map Φ_{k+1}, introducing first a more detailed notation for (17.2): we write $\Phi_j^{e_{i_j},\ldots,e_{i_1};\beta}$ for the map $\Phi_j : A^jF^\beta \to T^kF^\beta$ taken with respect to the sequence $\varepsilon_{i_1}, \ldots, \varepsilon_{i_j}$ of scalar extensions of the bundle F^β. Then our recursion formula for $\Phi_{k+1} = \Phi_k^{e_{k+1},\ldots,e_1;e_0}$ reads

$$\Phi_{k+1} = T\Phi_k^{e_k,\ldots,e_1;e_0} \circ (\times_{\substack{\alpha \in I_k \\ \alpha > 0}} \Phi_1^{e_{k+1},\alpha}). \tag{19.1}$$

Applying twice this recursion formula, we get another recursion formula for Φ_{k+1}, this time in terms of Φ_{k-1} and Φ_2:

$$\begin{aligned} \Phi_{k+1} &= T\Phi_k^{e_k,\ldots,e_1;e_0} \circ (\times_{\substack{\alpha \in I_k \\ \alpha > 0}} \Phi_1^{e_{k+1};\alpha}) \\ &= T^2\Phi_{k-1}^{e_{k-1},\ldots,e_1;e_0} \circ (\times_{\substack{\alpha \in I_{k-1} \\ \alpha > 0}} T\Phi_1^{e_k;\alpha}) \circ (\times_{\substack{\alpha \in I_k \\ \alpha > 0}} \Phi_1^{e_{k+1};\alpha}) \\ &= T^2\Phi_{k-1}^{e_{k-1},\ldots,e_1;e_0} \circ (\times_{\substack{\alpha \in I_{k-1} \\ \alpha > 0}} \Phi_2^{e_{k+1},e_k;\alpha}). \end{aligned} \tag{19.2}$$

Now, $T^2\Phi_{k-1}$, being a second tangent map with respect to scalar extensions $\varepsilon_{k+1}, \varepsilon_k$, commutes with the transposition $(k, k+1)$. As seen for the case $k = 2$ (beginning of our induction), $\Phi_2^{e_{k+1},e_k;\alpha}$ preserves the subspace fixed under $(k, k+1)$. Altogether, it follows that also Φ_{k+1} preserves the subspace fixed under $(k, k+1)$, as had to be shown. ■

19.3. *Multilinear differential calculus for jets.* If L is a jet-compatible connection on T^kF, then structures such as the linearization map or the higher differentials from Chapter 16 can be restricted to the subbundle J^kF. We give a short description of the theory, where, for simplicity, in the following we only consider the case of the tangent bundle, $F = TM$, leaving to the reader the generalization to arbitrary vector bundles. The model for the jet bundle J^kM is the subbundle of A^kM fixed under Σ_k,

$$(A^kM)^{\Sigma^l} = (\bigoplus_\alpha \varepsilon^\alpha TM)^{\Sigma_k} = \bigoplus_{j=1}^k \delta^{(j)} TM$$

with the $\delta^{(j)}$ as in (8.1). If L is a jet-compatible connection on T^kM, then it induces a bundle isomorphism

$$J\Phi : \bigoplus_{j=1}^k \delta^{(j)} TM \to J^kM,$$

and $J\Phi$ is the same for all connections L^σ, $\sigma \in \Sigma_k$. In particular, if $L = L_k = D^{k-1}K$ is derived from some linear connection K on TM, then $J\Phi$ depends only on K and not on the order of the scalar extensions in the definition of the iterated tangent functor T^k.

Now let $f : M \to N$ be a smooth map between manifolds with linear connections K on TM and L on TN. Restricting the covariant derivative $d_K^L f$ from Section 16.3 to a map $J_K^L f : J^k M \to J^k N$,

$$\begin{array}{ccc} J^k M & \overset{\Phi_k}{\leftarrow} & \bigoplus_{j=1}^k \delta^{(j)} TM \\ J^k f \downarrow & & \downarrow J_K^L f \\ J^k N & \overset{\Phi_k}{\leftarrow} & \bigoplus_{j=1}^k \delta^{(j)} TN \end{array} ,$$

we get a jet version of the higher covariant derivatives that no longer depends on the standard order of scalar extensions but only on the connections themselves. Restricting the expression (16.6) of $T^k f$ by its matrix coefficients to spaces of Σ_k-invariants and recalling from Section MA.3 that $\lambda \sim \mu$ iff $\ell(\lambda) = \ell(\mu) =: \ell$ and $|\lambda^j| = |\mu^j| =: i_j$ for all j, we get

$$J^k f(x + \sum_j \delta^{(j)} v_j) = f(x) + \sum_{j=1}^k \delta^{(j)} \sum_{\ell=1}^j \sum_{\substack{i_1 \leq \ldots \leq i_\ell \\ i_1 + \ldots + i_\ell = j}} (d^{(i_1,\ldots,i_\ell)} f)_x(v_{i_1}, \ldots, v_{i_\ell})$$

with the "matrix coefficient"

$$(d^{(i_1,\ldots,i_\ell)} f)_x(v_{i_1}, \ldots, v_{i_\ell}) = \sum_{\substack{\lambda : \forall r = 1,\ldots,\ell : \\ |\lambda^r| = i_r}} (d^\lambda f)_x(v_{i_1}, \ldots, v_{i_\ell}).$$

This formula should be compared with the one from Theorem 8.6 which concerns the "flat case" (chart connection). One expects that also in this general case there exist symmetry properties of the matrix coefficients and combinatorial relations between different coefficients, depending of course on the curvature operators of the connections. This seems to be an interesting topic for future work.

20. Shifts and symmetrization

20.1. *Totally symmetrizable multilinear connections.* We say that a multilinear connection L on T^kF is *(totally) symmetrizable* if all its curvature operators $R^{\sigma,L}$ belong to the central vector group $\mathrm{Gm}^{k,k+1}(T^kF)$ of $\mathrm{Gm}^{1,k+1}(T^kF)$. Equivalently, all curvature forms $\widetilde{R}^\Lambda$ for $\ell(\Lambda) < k$ vanish, and only the top curvature form possibly survives.

Theorem 20.2. *Assume that L is a totally symmetrizable multilinear connection on T^kF and that $2,\ldots,k$ are invertible in $\mathbb{K}$. Then all connections $\sigma \cdot L$, $\sigma \in \Sigma_k$, belong to the orbit $\mathrm{Gm}^{k,k+1}(E)\cdot L$, which is an affine space over $\mathbb{K}$, and their barycenter in this affine space,*

$$\mathrm{Sym}(L) := \frac{1}{k!}\sum_{\sigma\in\Sigma_k} \sigma \cdot L,$$

is a very symmetric multilinear connection on T^kF. Moreover, if L is invariant under Σ_{k-1}, then

$$\mathrm{Sym}(L) = \frac{1}{k}\sum_{j=1}^{k}(12\ldots k)^j \cdot L.$$

Proof. This follows by applying pointwise the corresponding purely algebraic result (Cor. SA.9). ■

Corollary 20.3. *Assume L is a torsionfree connection on TM. Then the linear structure DL on T^3M is symmetrizable.*

Proof. According to Theorem 18.3, the only non-vanishing curvature form of DL is the top curvature form, which means that DL is symmetrizable. ■

In the situation of the corollary, another proof of Bianchi's identity (cf. Section 17.5) is obtained by observing that the curvature of $\mathrm{Sym}(L)$ vanishes – see Section SA.10 for the purely algebraic argument.

20.4. *Shift invariance.* Besides the action of Σ_k, the iterated tangent bundle T^kF carries another canonical family of endomorphisms: recall from Section 7.3 (D) that, deriving the canonical action $\mathbb{K}\times F \to F$ of $\mathbb{K}$ on the vector bundle F, we get an action $T^k\mathbb{K}\times T^kF \to T^kF$ which is trivial on the base T^kM and linear in the fibers of the vector bundle $T^kF \to T^kM$. The chart representation is

$$\sum_{\alpha\in I_k}\varepsilon^\alpha r_\alpha \cdot \sum_{\alpha\in I_k}\varepsilon^\alpha(v_\alpha + w_\alpha) = \sum_\alpha \varepsilon^\alpha(v_\alpha + \sum_{\beta+\gamma=\alpha} r_\beta w_\gamma) \tag{20.1}$$

where $v_\alpha \in V$, $w_\alpha \in W$ and $v_0 := x \in U$. (This generalizes (7.11).) In particular, for $j = 1,\ldots,k$, the action of the "infinitesimal unit" ε_j on T^kF, denoted by

$$S_{0j} : T^kM \to T^kM, \quad u \mapsto \varepsilon_j.u, \tag{20.2}$$

will be called a *shift operator*. This terminology is explained by the chart representation of S_{0j}: letting $r_\alpha = 1$ if $\alpha = e_j$ and $r_\alpha = 0$ if $\alpha \neq e_j$, we see from (20.1) that ε_j acts in a fiber of the bundle $T^kF \to T^kM$ by "shifting" the index,

$$\varepsilon_j \cdot \sum_{\alpha \in I_k} \varepsilon^\alpha (v_\alpha + w_\alpha) = \sum_\alpha \varepsilon^\alpha v_\alpha + \sum_{\alpha_j = 1} \varepsilon^\alpha w_{\alpha - e_j} = \sum_\alpha \varepsilon^\alpha v_\alpha + \sum_{\alpha_j = 0} \varepsilon^{\alpha + e_j} w_\alpha. \quad (20.3)$$

All terms $\varepsilon^\alpha w_\alpha$ with $\alpha_j = 1$ are annihilated since $\varepsilon_j^2 = 0$. The action $T^k\mathbb{K} \times T^kF \to T^kF$ is Σ_k-invariant in the sense that, for all $\sigma \in \Sigma_k$ and $r \in T^k\mathbb{K}$, $u \in T^kF$, $\sigma(r.u) = \sigma(r).\sigma(u)$. In particular, the shift operators S_{0j} are permuted among each other by the Σ_k-action.

But now let $F = TM = T_{\varepsilon_0}M$. Then the action $T^k\mathbb{K} \times T^{k+1}M \to T^{k+1}M$ is not invariant under Σ_{k+1}. In particular, conjugating by the transposition $(0i)$ for $i = 1, \ldots, k$, we get further shift operators

$$S_{ij} := (0i) \circ S_{0j} \circ (0i) : T^{k+1}M \to T^{k+1}M.$$

In a chart, they are represented by

$$S_{ij} : T^{k+1}M \to T^{k+1}M, \quad (x; \sum_{\alpha > 0} \varepsilon^\alpha v_\alpha) \mapsto (x; \sum_{\substack{\alpha_i = 1 \\ \alpha_j = 0}} \varepsilon^{\alpha + e_j} v_\alpha + \sum_{\alpha_i = 0} \varepsilon^\alpha v_\alpha).$$

We say that S_{ij} is a shift *in direction* j *with basis* i, and if moreover $i > j$, we say that the shift is *positive*. For $k = 1$, the two shifts S_{10} and S_{01} are precisely the two almost dual structures on TTM (Section 4.7).

We say that a multilinear connection L on T^kF is *(positively) shift-invariant* if all (positive) shifts are linear operators with respect to L. Equivalently, the splitting map $\Phi : A^kF \to T^kF$ is equivariant under the groups $\mathrm{Shi}(T^kM)$, resp. $\mathrm{Shi}_+(T^kM)$ defined by (SA.20).

Note that shift invariance is a natural condition: chart connections are shift invariant. More generally, if F is a $\mathbb{K}$-vector bundle over M, then T^kF is a $T^k\mathbb{K}$-vector bundle over T^kM, and if $\widetilde{f} : F \to F'$ is a $\mathbb{K}$-vector bundle morphism, then $T^k\widetilde{f} : T^kF \to T^kF'$ is a $T^k\mathbb{K}$-vector bundle morphism. By $T^k\mathbb{K}$-linearity in fibers, we deduce that $T^k\widetilde{f}$ commutes with all shifts. Similarly, if $f : M \to N$ is smooth, then $T^{k+1}f$ commutes with all shifts S_{ij}. Invariance of a multilinear connection under S_{ij} means that the shifts act linearly with respect to the linear structure, and the fibers are actually modules over the ring obtained by adjoining to $\mathbb{K}$ an infinitsimal unit ε_j whose action is specified by i.

Theorem 20.5. *Assume L is a torsionfree connection on TM.*

1. *The linear structure DL on T^3M is invariant under all shifts.*
2. *The linear structure D^kL on $T^{k+2}M$ is invariant under the shifts $S_{k+2,k+1}$ and $S_{k+1,k+2}$.*

Proof. (1) We use the chart representation (17.4) for the map $\Phi_2 = D\Phi_1$:

$$\Phi_2(x, v) = v + \varepsilon_0\varepsilon_1 b_x(w_0, v_1) + \varepsilon_0\varepsilon_2 b_x(w_0, v_2) + \varepsilon_1\varepsilon_2 b_x(v_1, v_2) + \\ \varepsilon_0\varepsilon_1\varepsilon_2\big(b_x(w_0, v_{12}) + b_x(w_{01}, v_2) + b_x(w_{02}, v_1) + t(w_0, v_1, v_2)\big) \quad (20.4)$$

with trilinear component

$$t(w_0,v_1,v_2) = b_x(w_0,b_x(v_1,v_2)) + b_x(b_x(w_0,v_2),v_1) + \partial_{v_2} b_x(w_0,v_1).$$

Clearly, all bilinear components of $\Phi_2(x,\cdot)$ are conjugate to each other, and hence the algebraic Shift Invariance Condition (Prop. SA.13) is fulfilled.

(2) The claim is proved by induction. For $k=1$ it follows from Part (1). Now assume the claim is proved for $k>1$. We use the recursion formula (19.2) for Φ_{k+1}:

$$\Phi_{k+1}: \ A^{k+1}F \quad \overset{\oplus_\alpha \Phi_2^{e_{k+1},e_k;\alpha}}{\longrightarrow} \quad TT(A^{k-1}F) \quad \overset{TT\Phi_{k-1}}{\longrightarrow} \quad TT(T^{k-1}F) = T^{k+1}F$$

The map $TT\Phi_{k-1}$, like any second tangent map, commutes with both shifts $S_{k+1,k}$ and $S_{k,k+1}$. By the case $k=1$ of the induction, $\Phi_2^{e_{k+1},e_k;\alpha}$ also commutes with both shifts $S_{k+1,k}$ and $S_{k,k+1}$, whence the claim. ■

As seen in the algebraic part (Theorem SA.14), the condition that L be invariant under *all* shifts is a rather strong condition: it implies in turn that L is symmetrizable. – Part (1) of the preceding theorem does not carry over in this form to arbitrary k: using the recursion formula (17.3), one can see that D^2L is in general no longer invariant under the shifts S_{12} and S_{21}. On the other hand, the statement (2) is not the strongest possible result: combinatorial arguments, using again the recursion formula (17.3), show that D^kL is invariant under all shifts of the form $S_{i,k+2}$, $i<k+2$. One would like to understand this situation in a conceptual way.

Corollary 20.6. *Under the assumptions of the preceding theorem, and if* 3 *is invertible in* $\mathbb{K}$*, then the linear structure* $\mathrm{Sym}(DL)$ *on* T^3M *is totally symmetric and invariant under all shifts.*

Proof. By Cor. 20.3, DL is indeed symmetrizable, and by Theorem 20.2, $\mathrm{Sym}(DL)$ is then totally symmetric. We have to show that it is invariant under all shifts. In a chart, symmetrization amounts to replace in Equation (19.1) the trilinear map t by its symmetrized version $\frac{1}{3}(t+(123).t+(132).t)$, whereas the other components do not change (recall that b_x is already symmetric, by torsionfreeness). Thus $\mathrm{Sym}(DL)$ still satisfies the shift invariance condition SA.13. ■

Theorem 20.7. *Assume L is a* flat *linear structure on TM, i.e., without torsion and such that all curvature forms of DL vanish. Then all linear structures D^kL, $k\in\mathbb{N}$, are totally symmetric and invariant under all shifts.*

Proof. The proof is by induction on k: for $k=1$, DL is totally symmetric by assumption and shift-invariant by Theorem 20.5. For the induction step, we use exactly the same arguments as in the proof of Theorem 19.2, replacing the property of preserving the space of Σ_k-invariants by full Σ_k-invariance. In this way it is seen that D^kL also is Σ_{k+1}-invariant, i.e. all D^kL are totally symmetric. By Theorem 20.5 they are invariant under one shift (namely $S_{k+1,k}$), but since they are totally symmetric and all other shifts are conjugate among each other under the permutation group, D^kL is then invariant under all shifts. ■

For a complete understanding of the structure of $T^k M$, in relation with connections, it seems of basic importance to find a canonical construction of a sequence of totally symmetric and shift-invariant multilinear structures coming from the linear structures $D^k L$, $k = 1, 2, \ldots$. Unfortunately, already $D(\mathrm{Sym}(DL))$ is in general no longer totally symmetrizable in the sense of 20.1, and thus an appropriate symmetrization procedure has to be more subtle than the one used in Theorem 20.2. We will solve this problem for Lie groups (Chapter 25) and symmetric spaces (Chapter 30), where the solution is closely related to the *exponential mapping of polynomial groups* (Chapter PG), and then come back to the general problem in Chapter 31. – In the following chapter we recall a similar construction in a context which is "dual" to the one treated here, namely in the context of linear differential operators.

21. Remarks on differential operators and symbols

21.1. *Overview over definitions of differential operators.* In the literature several definitions of differential operators can be found. They agree for finite-dimensional real manifolds but may lead to quite different concepts in our much more general setting. All have in common that a *linear differential operator* between vector bundles $F \to M$ and $F' \to M$ should be a $\mathbb{K}$-linear map

$$D : \Gamma(F) \to \Gamma(F')$$

from sections of F to sections of F' satisfying some additional conditions such as:

(1) D is *support-decreasing*. Since there is no hope to prove an analog of the classical Peetre theorem (cf., e.g., [Hel84, Th. II.1.4] or [KMS93, Ch. 19]) in our setting, this cannot serve as a possible definition in the general case.

(2) The value $Df(x)$ depends only on the k-jet of the section $f : M \to F$ at x. More precisely, let $J^k(M, F, \mathrm{sec})$ be the space whose fibers over M are given by k-jets of sections of F (if $F = M \times \mathbb{K}$ is the trivial bundle, then $J^1(M, F, \mathrm{sec})$ essentially is the cotangent bundle T^*M). Following [KMS93, p. 143], we say that D is *local and of order k* if there is a map $D_k : J^k(M, F, \mathrm{sec}) \to F'$ such that $Df(x) = D_k(J^k_x f)$. In case of the trivial line bundle, i.e., $F = F' = M \times \mathbb{K}$, a differential operator of order one is thus seen as a section of the dual bundle of T^*M, and not as a section of TM. Already this example shows that, when trying to generalize this definition, we inevitably run into difficulties related to double dualization.

(3) Following the elegant algebraic definition of differential operators given e.g. in [ALV91, p. 85] or [Nes03, p. 125, p. 131], we may require that there exists $k \in \mathbb{N}$ such that, for all smooth functions $f_0, \ldots, f_k \in \mathcal{C}^\infty(M, \mathbb{K})$,

$$[\ldots[[D, f_0], f_1] \ldots f_k] = 0,$$

where, for an operator A from sections to sections, $[A, f]X = A(fX) - fA(X)$. Moreover, one would add some smoothness condition. It is not hard to see that (2) implies (3), whereas already in the infinite-dimensional Banach case the converse fails in general (cf. remarks in Section 4.6). Hence it seems that this definition is best suited for the finite-dimensional case.

(4) In the special case $F = F'$, we may assume that there exists $k \in \mathbb{N}$ such that D can (at least locally over a chart domain) be written as a finite product of operators $A\nabla_{X_1} \circ \ldots \circ \nabla_{X_k}$, where ∇ runs over the covariant derivatives associated to connections on F and A is a field of endomorphisms (cf. [BGV92, p.64]). In case $F = M \times \mathbb{K}$ is the trivial bundle over a finite-dimensional manifold, this means that, in a chart, $D = g_0 + \sum_\alpha g_\alpha \partial^\alpha$ is a sum a partial derivatives multiplied by functions. This is the most classical notion; it implies (2) and hence (3).

All of these concepts are useful, and hence one should distinguish between different classes of differential operators. In any case, k is the *degree* or *order of D*, and the degree defines a natural filtration on the space of differential operators. The second approach seems to be the most conceptual one, and probably a correct formalization of concepts of differential operators in our context will need the use of (algebraic) double duals. However, this is a topic for another work, and therefore the approach outlined in the sequel will be closest to the most classical definition (4). Things will be most clear if we start with differential operators on the trivial line bundle, i.e. with scalar differential operators.

21.2. *Pure differential operators.* A *pure scalar differential operator of order at most* k $(k \geq 1)$ on a manifold M is a smooth section $D : M \to T^k M$ of the higher order tangent bundle $T^k M$ over M. The space of such sections is denoted by $\mathfrak{X}^k(M)$. (See Chapter 28 for a more systematic theory.) If $f : M \to \mathbb{K}$ is a smooth function, then $T^k f \circ D : M \to T^k \mathbb{K}$ is smooth, and we let

$$Df := \mathrm{pr}_k \circ T^k f \circ D : M \to \mathbb{K}, \tag{21.1}$$

where $\mathrm{pr}_k : T^k \mathbb{K} \to \mathbb{K}$, $\sum_{\alpha \geq 0} a_\alpha \varepsilon^\alpha \mapsto a_{(1\ldots1)}$ is the projection onto the last factor. Thus D induces a map

$$D^{op} : \mathcal{C}^\infty(M) \to \mathcal{C}^\infty(M), \quad f \mapsto Df$$

which clearly is $\mathbb{K}$-linear. Let $\mathcal{D}_k(M) \subset \mathrm{End}_{\mathbb{K}}(\mathcal{C}^\infty(M))$ be the $\mathbb{K}$-span of all such operators, called the space of *linear scalar differential operators of degree at most* k. In a chart, using Formula (7.19), the "constant coefficient" operator $D = \sum_{\alpha > 0} \varepsilon^\alpha v_\alpha$ acts by

$$(\sum_{\alpha>0} \varepsilon^\alpha v_\alpha \cdot f)(x) = \mathrm{pr}_k(T^k f(x + \sum_\alpha v_\alpha)) = \sum_{\ell=1}^{k} \sum_{\beta \in \mathcal{P}_\ell(1,\ldots,1)} d^\ell f(x)(v_{\beta^1}, \ldots, v_{\beta^\ell}), \tag{21.2}$$

In other words,

$$(\sum_{\alpha>0} \varepsilon^\alpha v_\alpha)^{op} = \sum_{\ell=1}^{k} \sum_{\beta \in \mathcal{P}_\ell(1,\ldots,1)} \partial_{v_{\beta^1}} \cdots \partial_{v_{\beta^\ell}}$$

is a scalar differential operator in the most classical sense. For instance, $\varepsilon_1 v_1 + \varepsilon_2 v_2$ acts as $\partial_{v_1} \partial_{v_2}$ and $\delta^{(2)} v$ as ∂_v^2. The chart formula also shows that, for all $\sigma \in \Sigma_k$ acting on $T^k M$ in the canonical way, $\sigma \circ D$ and D give rise to the same operator; therefore, if $2, \ldots, k$ are invertible in $\mathbb{K}$, then we could as well work with sections of $J^k M$. Also, if $T^k z : T^k M \to T^{k+1} M$ is the injection obtained by deriving some zero section $z = z_j : M \to T_{\varepsilon_j} M$, then D and $z \circ D$ give rise to the same operator on functions. Somewhat loosely, we thus may say that a differential operator of order k is also one of degree $k+1$. (Taking in the definition of Df instead of pr_k the projection onto the "augmentation ideal" of $T^k \mathbb{K}$, one gets another kind of operators satisfying an analog of the Leibniz rule, called *expansions* in [KMS93, Section 37.6].)

The *composition* of two pure differential operators $D_i : M \to T^{k_i} M$, $i = 1, 2$, is defined by

$$D_2 \cdot D_1 := T^{k_1} D_2 \circ D_1 : M \to T^{k_1} M \to T^{k_1 + k_2} M. \tag{21.3}$$

It is easily checked that this composition is associative, and that $D \mapsto D^{op}$ is a "contravariant" representation:

$$\begin{aligned}(D_2 \cdot D_1)^{op} f &= \mathrm{pr}_{k_1+k_2} \circ T^{k_1+k_2} f \circ (D_2 \cdot D_1) \\ &= \mathrm{pr}_{k_1+k_2} \circ T^{k_1+k_2} f \circ T^{k_1} D_2 \circ D_1 \\ &= \mathrm{pr}_{k_1+k_2} \circ T^{k_1} (T^{k_2} f \circ D_2) \circ D_1 \\ &= \mathrm{pr}_{k_1} \circ T^{k_1} (\mathrm{pr}_{k_2} \circ T^{k_2} f \circ D_2) \circ D_1 \\ &= (D_1^{op} \circ D_2^{op}) f.\end{aligned}$$

It follows that $\mathcal{D}^k(M) \cdot \mathcal{D}^\ell(M) \subset \mathcal{D}^{k+\ell}(M)$, and hence $\mathcal{D}^\infty(M) := \bigcup_{i=1}^\infty \mathcal{D}^i(M)$ is an associative filtered algebra over $\mathbb{K}$.

21.3. *Principal symbol of a scalar differential operator.* Let $D : M \to T^k M$ be a section and

$$\mathrm{pr} : T^k M \to \varepsilon_1 TM \times \ldots \times \varepsilon_k TM$$

be the product of the canonical projections, and let

$$t : \varepsilon_1 TM \times \ldots \times \varepsilon_k TM \to \varepsilon_1 TM \otimes \ldots \otimes \varepsilon_k TM$$

be the tensor product map (where $\otimes = \otimes_M$ is the algebraic tensor product of vector bundles over M, i.e. given by tensor products in fibers, without considering a topology). We say that

$$\sigma^k(D) := t \circ \mathrm{pr} \circ D : M \to \varepsilon_1 TM \otimes \ldots \otimes \varepsilon_k TM$$

is the *principal symbol of* D. It is a section of the (algebraic) bundle $\otimes^k TM$ over M. Comparison with the chart representation (21.2) shows that the principal symbol of $\sum_{\alpha>0} \varepsilon^\alpha v_\alpha$ is $v_1 \otimes \ldots \otimes v_k$, coming from the "highest term" $\partial_{v_1} \cdots \partial_{v_k}$. If D happens to be a section of $J^k M$, then $\sigma^k(D)$ is a section of the algebraic bundle $S^k(TM)$ over M (k-th symmetric power of TM). If the "operator representation" $D \mapsto D^{op}$ is injective, then one can show that the principal symbol map factors through a $\mathbb{K}$-linear map

$$\widetilde{\sigma}^k : \mathcal{D}^k(M) \to \Gamma(S^k TM)$$

(where $\Gamma(S^k TM)$ is the space of sections that locally are finite sums of t composed with smooth sections of $TM \times \ldots \times TM$), i.e. $\sigma^k(D) = \widetilde{\sigma}^k(D^{op})$, and that the sequence

$$0 \quad \to \quad \mathcal{D}^{k-1}(M) \quad \to \quad \mathcal{D}^k(M) \quad \stackrel{\widetilde{\sigma}^k}{\to} \quad \Gamma(S^k(TM)) \quad \to 0 \tag{21.4}$$

is exact. Details are left to the reader.

21.4. *Correspondence between total symbols and differential operators (scalar case).* For real finite dimensional manifolds it is known that, in presence of a connection on TM, one can associate to a differential operator a *total symbol* (which is a section of $S^k TM$) in such a way that this correspondence becomes a bijection between operators and their total symbols.[1] For the case of second order operators, the construction is essentially described in [Lo69] (see also [P62] and [Be00, I.B]); the general construction seems to be folklore – the few references we know about include [BBG98] and [Bor02] (where the construction is called *préscription d'ordre standard*). Let us briefly describe the construction: clearly, the problem is equivalent to find $C^\infty(M)$-linear sections of the exact sequences (21.4) for all k. Using the covariant derivative ∇ of L, for a function f one defines $\nabla^k f : \times^k TM \to \mathbb{K}$, the "$k$-th order Hessian of f" (defined as usual in differential geometry by $\nabla^1 f := \mathrm{d}f$, $\nabla^{j+1} f := \nabla(\nabla^j f)$). Then to a section s of $S^k TM$ we associate a differential operator $D := D(s)$: if $s(x) = v_1 \otimes \ldots \otimes v_k$ with all $v_j \in T_x M$, let

$$Df(x) = (\nabla^k f)(x; v_1, \ldots, v_k).$$

One easily proves that $\widetilde{\sigma}^k D(s) = s$, hence $s \mapsto D(s)$ gives the desired section of the exact sequence (21.4). This construction should be seen as an analog of the construction of linear structures on higher order tangent bundles (Chapter 17) ; it can be adapted to the case of differential operators acting on sections of (finite dimensional) vector bundles over finite dimensional real manifolds (cf. [Bor02]).

21.5. *Differential operators on general vector bundles.* First of all, differential operators on a trivial bundle $M \times W$ may be defined as in the scalar case. For general vector bundles, the problem arises that we need a good class of differential operators of degree zero (which can be avoided on trivial vector bundles). The reason for this is that, although vector bundles can be trivialized locally, the preceding definition of differential operator over chart domains would no longer be chart-independent. The smallest class of differential operators of degree zero which one could take here are fields of endomorphisms that arise precisely from chart changes. Taking a somewhat bigger class, one arrives at Definition (4) of a differential operator mentioned in Section 21.1.

Finally, in order to define differential operators acting from sections of one vector bundles F, to sections of *another* bundle F', one needs to single out some subspace in $\mathrm{Hom}_{\mathbb{K}}(F_x, F'_x)$ giving the differential operators of degree zero; but there is no natural candidate for such a space (in the general case).

[1] [ALV91, p.87]: "Obviously, to one and only one symbol there correspond many differential operators. However, it is a remarkable fact that in presence of a connection this correspondence can be made one-to-one." On the same page, the authors add the remark: "This procedure of restoring a differential operator from its symbol is completely analogous to the basic procedure of quantum mechanics – quantization."

22. The exterior derivative

22.1. *Skewsymmetric derivatives.* In the preceding chapters, we have focused on symmetric derivatives. They depend on an additional structure (connection) and thus may be considered as "not natural". The skewsymmetric version of the derivative (the exterior derivative), in contrast, is completely natural, and therefore plays a dominant role in differential geometry[1]. Since the exterior derivative is so natural, we wish to give also a definition which, in our context, is as natural as possible. Recall (Section 4.5) that we consider a differential k-form ω as a smooth map $TM \times_M \ldots \times_M TM \to \mathbb{K}$ which is multilinear and alternating in fibers. One could proceed as in [La99, Ch. V], where the exterior derivative of a differential form ω is defined in a chart by

$$\begin{aligned} \mathrm{d}\omega(x; v_1, \ldots, v_{k+1}) &= \sum_{\sigma \in \Sigma_{k+1}} \operatorname{sgn}(\sigma) \partial_{v_{\sigma(1)}} \omega(\widehat{v_{\sigma(1)}}, v_{\sigma(2)}, \ldots, v_{\sigma(k+1)}) \\ &= \sum_{i=1}^{k+1} (-1)^i \partial_{v_i} \omega(v_1, \ldots, \widehat{v_i}, \ldots, v_{k+1}). \end{aligned} \tag{22.1}$$

Then one shows that evaluation of $\mathrm{d}\omega$ on vector fields $X_1, \ldots, X_k$ is given by the formula

$$\begin{aligned} \mathrm{d}\omega(X_1, \ldots, X_{k+1}) = &\sum_{i=1}^{k+1} (-1)^i X_i \omega(X_1, \ldots, \widehat{X_i}, \ldots, X_{k+1}) + \\ &\sum_{1 \le i < j \le k+1} (-1)^{i+j} \omega([X_i, X_j], X_1, \ldots, \widehat{X}_i, \ldots, \widehat{X}_j, \ldots, X_{k+1}), \end{aligned} \tag{22.2}$$

which implies that (22.1) is chart-independent. The usual properties of the exterior derivative then follow by standard calculations. A "natural definition" in our sense should not rely on chart formulas, and it should be built only on the natural operators T (tangent functor) and alt (alternation operator). For the exterior derivative of one-forms this has been carried out in Chapter 13. Here is the pattern for the general case:

(1) We define the notion of an *(intrinsic) k-multilinear map* $f : T^k M \to \mathbb{K}$, and we define also *homogeneous* such maps.

(2) There is a canonical bijection between tensor fields $\underline{\omega} : \times_M^k TM \to \mathbb{K}$ and homogeneous k-multilinear maps $\overline{\omega} : T^k M \to \mathbb{K}$.

(3) If $f : T^k M \to \mathbb{K}$ is k-multilinear, then $\mathrm{pr}_2 \circ Tf : T^{k+1} M \to \varepsilon\mathbb{K}$ is $k+1$-multilinear.

(4) The alternation operator alt maps intrinsic k-multilinear maps onto homogeneous ones.

[1] Cf., e.g., [Sh97, p. 54]: "As usual in modern differential geometry, we shall be concerned only with the *skew-symmetric* part of the higher derivatives."

(5) Finally, we let $\mathrm{d}\underline{\omega} := \underline{\mathrm{alt}(\mathrm{pr}_2 T(\overline{\omega}))}$. In other words, the natural operator d is defined by the following commutative diagram:

$$\begin{array}{ccccc} \mathcal{T}^{k,0}(M) & \cong & \mathcal{M}_h(T^kM) & \subset & \mathcal{M}(T^kM) \\ \mathrm{d}\downarrow & & & & \downarrow T \\ \mathcal{T}^{k+1,0}(M) & \cong & \mathcal{M}_h(T^{k+1}M) & \overset{\mathrm{alt}}{\leftarrow} & \mathcal{M}(T^{k+1}M) \end{array}$$

where $\mathcal{T}$ denotes tensor fields and $\mathcal{M}$ (resp. $\mathcal{M}_h$) intrinsic multilinear (homogeneous) maps.

(6) The property $\mathrm{d} \circ \mathrm{d} = 0$ follows from the fact that $\mathrm{alt} \circ T \circ T = 0$ due to symmetry of second differentials.

In the following, we carry out this program.

22.2. *Multilinear maps on T^kM.* We say that a smooth map $\omega : T^kM \to \mathbb{K}$ is *(intrinsically) multilinear* if $\omega : T(T^{k-1}M) \to \mathbb{K}$ is linear in fibers, for all the various ways $p_j : T^kM \to T^{k-1}M$, $j = 1, \dots, k-1$ of seeing T^kM as a vector bundle over $T^{k-1}M$. Equivalently, for all $x \in M$, the restriction $\omega_x : (T^kM)_x \to \mathbb{K}$ to the fiber of T^kM over x is intrinsically multilinear in the sense of Section MA.15. We denote, for fixed $x \in M$, by H_j the kernel of p_j (tangent spaces $T_{0_x}(T^{k-1}M) \subset (T^kM)_x$ with first $T = \varepsilon_j T$). We say that ω is *homogeneous* if ω_x is homogeneous in the sense of MA.15, i.e. the restriction of ω_x to intersections $H_{ij} := H_i \cap H_j$ ($i \neq j$) is constant. In a chart, an intrinsically multilinear map is given by

$$\omega(x + \sum_\alpha \varepsilon^\alpha v_\alpha) = \sum_{\ell=1}^{k} \sum_{\Lambda \in \mathcal{P}_\ell(1,\dots,1)} b_x^\Lambda(v_{\Lambda^1}, \dots, v_{\Lambda^\ell}), \tag{22.3}$$

where $b_x^\Lambda : V^\ell \to \mathbb{K}$ is a multilinear map in the usual sense, and it is homogeneous if and only if the only non-zero term in (22.3) is the one belonging to $\ell = k$, i.e., to the longest partition $\Lambda = (e_1, \dots, e_k)$ of the full multi-index $(1, \dots, 1)$ (see Prop. MA.16). This implies as in Prop. MA.16 that we have a bijection between tensor fields and homogeneous multilinear maps:

$$\begin{aligned} \mathcal{T}^{k,0}(M, \mathbb{K}) \to \mathcal{M}_h(T^kM, \mathbb{K}), \quad & \omega \mapsto \overline{\omega} = \omega \circ \mathrm{pr}, \\ \mathcal{M}^h(T^kM, \mathbb{K}) \to \mathcal{T}^{k,0}(M, \mathbb{K}), \quad & \overline{\omega} \mapsto \underline{\omega} := \overline{\omega} \circ \iota^L, \end{aligned} \tag{22.4}$$

where $\mathrm{pr} : T^kM \to TM \times_M \ldots \times_M \times TM$, $v \mapsto (\mathrm{pr}_1(v), \dots, \mathrm{pr}_k(v))$ is the vector bundle direct sum of the various projections $\mathrm{pr}_i : T^kM \to TM$. The definition of the first map in (22.4) does not involve connections or charts, whereas the definition of the second map does – the injection $\iota^L : \times^k TM \to T^kM$ obtained from the Splitting Map $\Phi_k : \oplus_\alpha \varepsilon^\alpha TM \to T^kM$ depends on a connection L. The situation is summarized by the following diagram (cf. (MA.36)):

$$\begin{array}{ccc} T^kM & & \\ \mathrm{pr} \downarrow\uparrow \iota^L & \overset{\overline{\omega}}{\searrow} & \\ \times_M^k TM & \overset{\underline{\omega}}{\to} & \mathbb{K} \end{array}$$

Later on, one may consider $\overline{\omega}$ and $\underline{\omega}$ as the same object ω.

Lemma 22.3. *If* $\overline{\omega} : T^k M \to \mathbb{K}$ *is* k*-multilinear, then* $\mathrm{pr}_2 \circ T\overline{\omega} : T^{k+1}M \to \varepsilon\mathbb{K}$ *is* $k+1$*-multilinear.*

Proof. This can be proved by a direct computation using (22.3). More intrinsically, the lemma is proved as follows. Let the last tangent functor correspond to scalar extension with ε_{k+1}. Then we have two kinds of tangent spaces in $T^{k+1}M$: the first kind is given by spaces that are stable under ε_{k+1}; they correspond to projections p_i, $i = 1, \ldots, k$, and are tangent bundles of tangent spaces in $T^k M$. The restriction of Tf to such tangent spaces is the tangent map of f restricted to tangent spaces in $T^k M$ and hence is linear. The second kind of tangent spaces corresponds to the last projection p_{k+1}; but since Tf is defined with respect to ε_{k+1}, the restriction of Tf to such tangent spaces is linear as it is a tangent map.■

Lemma 22.4. *If* $\overline{\omega} : T^k M \to \mathbb{K}$ *is* k*-multilinear, then*

$$\mathrm{alt}\,\overline{\omega} : T^k M \to \mathbb{K}, \quad z \mapsto \sum_{\sigma \in \Sigma_k} \mathrm{sgn}(\sigma)\overline{\omega}(\sigma(z))$$

is homogeneous k*-multilinear.*

Proof. It is clear that $\mathrm{alt}(\omega)$ is k-multilinear since it is a sum of multilinear functions. It remains to prove homogeneity: notation being as in MA.15, we have to show that $\mathrm{alt}\,\overline{\omega}(v) = 0$ for $v \in H_{ij}$ with $i \neq j$. Let $\tau = (ij)$ be transposition between i and j; then τ acts trivially an H_{ij} since $v = \sum_\alpha \varepsilon^\alpha v_\alpha$ where $\alpha_i = 1 = \alpha_j$ for all α contributing effectively to the sum. It follows that $\tau v = v$ and hence

$$\mathrm{alt}\,\overline{\omega}(v) = \sum_{\sigma \in \Sigma_n} \mathrm{sgn}(\sigma)\overline{\omega}(\sigma v) = \sum_{\sigma \in A_n} (\overline{\omega}(\sigma v) - \overline{\omega}(\sigma\tau v)) = 0. \qquad ■$$

22.5. *Exterior derivative.* The exterior derivative of a k-form $\underline{\omega} : \times^k TM \to \mathbb{K}$ is now defined by

$$\mathrm{d}\underline{\omega} := \underline{\mathrm{alt}_{k+1}\,\mathrm{pr}_2\,T\overline{\omega}} = \underline{\mathrm{pr}_2\,\mathrm{alt}_{k+1}\,T\overline{\omega}}.$$

In fact, the same definition can be given for any tensor field $\underline{\omega} : T^k M \to \mathbb{K}$ (not necessarily skew-symmetric), but then $\mathrm{d}\underline{\omega} = \mathrm{d}(\mathrm{alt}\,\underline{\omega})$, so this gives nothing really new. Similarly, we define the exterior derivative of $\overline{\omega} \in \mathcal{M}_h(T^k M)$ by

$$\mathrm{d}\overline{\omega} := \mathrm{alt}_{k+1}(\mathrm{pr}_2\,T\overline{\omega}) = \mathrm{pr}_2(\mathrm{alt}_{k+1}\,T\overline{\omega}) \in \mathcal{M}_h(T^{k+1}M).$$

Theorem 22.6. *For all* k*-forms* $\underline{\omega}$*, we have* $\mathrm{d}\mathrm{d}\underline{\omega} = 0$.

Proof. We show, equivalently, that $\mathrm{d}\mathrm{d}\overline{\omega} = 0$. The operator $\mathrm{d} = \mathrm{d}_k$ is defined as

$$\mathrm{d}_k = \mathrm{alt}_{k+1} \circ \mathrm{pr}_2 \circ T = \mathrm{pr}_2 \circ \mathrm{alt}_{k+1} \circ T,$$

and hence

$$\mathrm{d}_{k+1} \circ \mathrm{d}_k = \mathrm{alt}_{k+2} \circ \mathrm{pr}_2 \circ T \circ \mathrm{alt}_{k+1} \circ \mathrm{pr}_2 \circ T = \mathrm{pr}_4 \circ \mathrm{alt}_{k+2} \circ T \circ \mathrm{alt}_{k+1} \circ T.$$

Now, we have $T \circ \mathrm{alt}_k = \mathrm{alt}_k \circ T$ for all k, because for all $\sigma \in \Sigma_k$, acting on $T^k M$, we have $T(\overline{\omega} \circ \sigma) = T\overline{\omega} \circ T\sigma$, where $T\sigma$ is the action of σ on $T^{k+1}M$ if we inject Σ_k into Σ_{k+1} in the natural way. Hence $T \circ \mathrm{alt}_k = \mathrm{alt}_k \circ T$ for all k, and from this follows $\mathrm{pr}_2\,T \circ \mathrm{alt}_k = \mathrm{alt}_k \circ \mathrm{pr}_2\,T$, and we get for the operator $\mathrm{d}_{k+1} \circ \mathrm{d}_k$:

$$\mathrm{d}_{k+1} \circ \mathrm{d}_k = \mathrm{pr}_4 \circ \mathrm{alt}_{k+2} \circ T \circ \mathrm{alt}_{k+1} \circ T \circ \mathrm{pr}_2 \circ T = \mathrm{pr}_4 \circ \mathrm{alt}_{k+2} \circ \mathrm{alt}_{k+1} \circ T \circ T.$$

But, by symmetry of the second differential, $TT\overline{\omega}$ is invariant under the permutation $(k+2, k+1) \in \Sigma_{k+2}$ which corresponds to exchanging ε_{k+2} and ε_{k+1}, and hence applying the alternation operator alt_{k+2} annihilates the whole expression. ■

The *de Rham cohomology* is now defined as usual:

$$H^k(M) := \ker(\mathrm{d}_{k+1})/\operatorname{im}(\mathrm{d}_k).$$

22.7. *Exterior derivative and covariant derivative.* If L is a linear connection of TM, we define the covariant derivative of a tensor field $\underline{\omega} : \times^k TM \to \mathbb{K}$ by

$$\nabla\underline{\omega} := \underline{\mathrm{pr}_2 T\overline{\omega}} \in \mathcal{T}^{k+1,0}(M),$$

where $\underline{f} := f \circ \iota^L$ is defined as in (22.4) and depends on the connection L unless f is homogeneous.

Proposition 22.8. *For all k-forms $\underline{\omega} : \times^k M \to \mathbb{K}$,*

$$\mathrm{d}\underline{\omega} = \mathrm{alt}_{k+1}(\nabla\underline{\omega}).$$

Proof.

$$\mathrm{alt}_{k+1}(\nabla\underline{\omega}) = \mathrm{alt}_{k+1}\, \underline{\mathrm{pr}_2 T\overline{\omega}} = \underline{\mathrm{alt}_{k+1}\, \mathrm{pr}_2 T\overline{\omega}} = \mathrm{d}\underline{\omega}. \qquad \blacksquare$$

Corollary 22.9. *In a bundle chart,* $\mathrm{d}\omega$ *is given by* (22.1)*, and evaluation of* $\mathrm{d}\omega$ *on vector fields* $X_1, \ldots, X_k$ *is given by* (22.2)*.*

Proof. Let ∇ be the covariant derivative induced by the chart, i.e. $\nabla = d$ is the ordinary derivative. Then

$$(\nabla\omega)(x; v_1, \ldots, v_{k+1}) = \partial_{v_{k+1}}\omega(x; v_1, v_2, \ldots, v_k)$$

Applying alt_{k+1} gives the first line of (22.1), and it implies the second line since already ω was supposed to be skew-symmetric.

In order to prove the formula for evaluation on vector fields, we first have to check that the result of the right-hand side at a point $p \in M$ depends only on the values of the X_i at p; then take $X_i = \widetilde{v}_i$, the extension to a constant vector field in a chart, and the formula reduces to the one established before. $\blacksquare$

Of course, one may ask now: what about the Poincaré lemma and de Rham's theorem? Clearly, these results do not carry over in their classical forms (not even for the infinite-dimensional real case). Already determining the solution space of the "trivial" differential equation $df = 0$ for a scalar function f, as a first step towards the Poincaré lemma, is in general not possible (in the p-adic case there are non-locally constant solutions, see [Sch84] for an example), and one may expect that there is some cohomological theory giving a better insight into the structure of the solution space of this "trivial" differential equation. For general k-forms, one should then combine such a theory with the usual de Rham cohomology.

V. Lie theory

In this part we develop the basic "higher order theory" of iterated tangent bundles of Lie groups and symmetric spaces. It can be read without studying most of the preceding results – see the "Leitfaden" (introduction). Strictly speaking, not even Theorem 4.4 on the Lie bracket of vector fields is needed since we give, in Chapter 23, another, independent definition of the Lie bracket associated to a Lie group, see below.

23. The three canonical connections of a Lie group

23.1. In the following theorem we will give another characterization of the Lie bracket of a Lie group. One might as well use it as *definition* of the Lie bracket; this has the advantage that one can develop Lie theory without speaking about vector fields, and in this way the theory becomes simpler and more transparent. In the following, G is a Lie group with multiplication $m : G \times G \to G$ (cf. Chapter 5). Then TTG is a Lie group with multiplication TTm, and the fiber $(TTG)_e$ is a normal subgroup. We let $\mathfrak{g} := T_eG$; then the three "axes" $\varepsilon_1\mathfrak{g}$, $\varepsilon_2\mathfrak{g}$, $\varepsilon_1\varepsilon_2\mathfrak{g}$ of $(TTG)_e$ are canonical, and the fiber of the axes-bundle over the origin is $(A^2G)_e = \varepsilon_1\mathfrak{g} \oplus \varepsilon_2\mathfrak{g} \oplus \varepsilon_1\varepsilon_2\mathfrak{g} \cong \mathfrak{g} \times \mathfrak{g} \times \mathfrak{g}$. (Cf. Chapter 15 for this notation.)

Theorem 23.2. (Canonical commutation rules in $(TTG)_e$.) *The group commutator $[g, h] = ghg^{-1}h^{-1}$ of an element $g = \varepsilon_1 v$ from the first and an element $h = \varepsilon_2 w$ from the second axis in $(TTG)_e$, where $v, w \in \mathfrak{g}$, can be expressed by the Lie bracket $[v, w]$ in the following way:*

$$[\varepsilon_1 v, \varepsilon_2 w] = \varepsilon_1\varepsilon_2[v, w], \tag{23.1}$$

and the third axis $\varepsilon_1\varepsilon_2\mathfrak{g}$ is central in $(TTG)_e$, i.e. for $j = 1, 2$ and all $u, v \in \mathfrak{g}$,

$$[\varepsilon_1\varepsilon_2 u, \varepsilon_j v] = 0. \tag{23.2}$$

Proof. First we prove (23.2). For any Lie group H, the tangent space T_eH is an abelian subgroup of TH whose group law is vector addition; this follows from Equation (5.1). Now, taking $H := TG$, $\varepsilon_1\varepsilon_2 u$ and $\varepsilon_j v$ both belong to the fiber over the origin in $T_{\varepsilon_j}TH$ and hence commute. Moreover, we see that

$$\varepsilon_i v \cdot \varepsilon_i w = \varepsilon_i(v + w) \quad (i = 1, 2), \qquad \varepsilon_1\varepsilon_2 v \cdot \varepsilon_1\varepsilon_2 w = \varepsilon_1\varepsilon_2(v + w). \tag{23.3}$$

Next, we prove the fundamental relation (23.1): by definition, the Lie bracket of $v, w \in \mathfrak{g}$ is given by evaluating $[v^R, w^R]$ at e, where v^R, w^R are the right-invariant vector fields extending v, w (cf. our convention concerning the sign of the Lie bracket from Section 5.3). Recall that v^R is given by left multiplication by v in TG, so the corresponding infinitesimal automorphism is left translation $l_v : TG \to TG$,

$u \mapsto v \cdot u$. The vector field v^R gives rise to two sections of TTG over G, with the notation from Section 14.2 denoted by $\varepsilon_i v^R$ ($i = 1, 2$), and each section gives rise to a diffeomorphism $\widetilde{\varepsilon_i v^R}$ of TTG (Theorem 14.3). These are nothing but the two versions of the tangent map $Tl_v : TTG \to TTG$, and hence both can be written as left multiplications in TTG, namely $\widetilde{\varepsilon_1 v^R} = l_{\varepsilon_1 v}$ and $\widetilde{\varepsilon_2 v^R} = l_{\varepsilon_2 v}$ with $\varepsilon_j v \in (TTG)_e$, $j = 1, 2$, defined as above. Now, applying first the definition of the Lie bracket in $\mathfrak{g}$ and then Theorem 14.4, we get

$$\begin{aligned}\varepsilon_1\varepsilon_2[v,w] &= \varepsilon_1\varepsilon_2[v^R, w^R](e)\\ &= [\widetilde{\varepsilon_1 v^R}, \widetilde{\varepsilon_2 w^R}]_{\mathrm{Diff}(TTG)}\,(e) = [l_{\varepsilon_1 v}, l_{\varepsilon_2 w}]_{\mathrm{Diff}(TTG)}\,(e)\\ &= (\varepsilon_1 v)(\varepsilon_2 w)(\varepsilon_1 v)^{-1}(\varepsilon_2 w)^{-1} = [\varepsilon_1 v, \varepsilon_2 w]_{TTG}\end{aligned}$$

where we have added as index the group in which the respective commutators are taken. ■

We will often use (23.1) in the following form:

$$\varepsilon_1 v_1 \cdot \varepsilon_2 w_2 = \varepsilon_1\varepsilon_2[v_1, w_2] \cdot \varepsilon_2 w_2 \cdot \varepsilon_1 v_1. \tag{23.4}$$

Another way of stating (23.1) is by saying that the second order tangent map of the commutator map

$$c := c_G : G \times G \to G, \quad (g,h) \mapsto [g,h] = ghg^{-1}h^{-1}$$

is determined by

$$T^2 c(\varepsilon_1 v, \varepsilon_2 w) = \varepsilon_1\varepsilon_2[v,w];$$

this follows simply by noting that $c_{TTG} = TTc_G$. One may also work with the conjugation map $G \times G \to G$, $(g,h) \mapsto ghg^{-1}$; then the fundamental relation (23.1) can also be written in the following way

$$\mathrm{Ad}(\varepsilon_1 v)\varepsilon_2 w = \varepsilon_2 w \cdot \varepsilon_1\varepsilon_2[v,w] = (\mathrm{id} + \varepsilon_1 \,\mathrm{ad}(v))\varepsilon_2 w.$$

This relation is equivalent to the the fact that the differential of $\mathrm{Ad} : G \times \mathfrak{g} \to \mathfrak{g}$ is $\mathrm{ad} : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$, $(X,Y) \mapsto [X,Y]$ (cf. [Ne02b, Prop. III.1.6] for the real locally convex case).

23.3. *Left- and right trivialization of TG and of TTG.* So far we have avoided to use the well-known fact that the tangent bundle TG of a Lie group G can be trivialized: the sequence of group homomorphisms given by injection of $\mathfrak{g} = T_e G$ in TG and projection $TG \to G$,

$$0 \;\to\; \mathfrak{g} \;\to\; TG \;\to\; G \;\to\; 1,$$

is exact and splits via the zero-section $z : G \to TG$. Thus, group theoretically, TG is a semidirect product of G and of $\mathfrak{g}$. Explicitly, the map

$$\Psi_1^R : \mathfrak{g} \times G \to TG, \quad (v,g) \mapsto v \cdot g = Tm(v, 0_g) = T_e(r_g)v = l_v(0_g) \tag{23.5}$$

is a smooth bijection with smooth inverse $TG \to \mathfrak{g} \times G$, $u \mapsto (u\pi(u)^{-1}, \pi(u))$. The diffeomorphism $TG \cong \mathfrak{g} \times G$ thus obtained is called the *right trivialization of* *the tangent bundle*. Similarly, the left trivialization Ψ_1^L is defined. The kernel of the canonical projection $TG \to G$ is $\mathfrak{g} = T_eG$ on which G acts via the *adjoint representation*

$$\mathrm{Ad} : G \times \mathfrak{g} \to \mathfrak{g}, \quad (g, v) \mapsto \mathrm{Ad}(g)v = gvg^{-1}$$

(product taken in TG; then as usual $\mathrm{Ad}(g) : \mathfrak{g} \to \mathfrak{g}$ is the differential of the conjugation by g at the origin). From Formula (5.1) we see that the normal subgroup $\mathfrak{g}$ of TG is abelian with product being vector addition. It follows that the group structure of TG in the right trivialization is given by

$$(v, g) \cdot (w, h) = v \cdot g \cdot w \cdot h = v \cdot (gwg^{-1}) \cdot gh = (v + \mathrm{Ad}(g)w, gh), \tag{23.6}$$

which is the above mentioned realization of TG is a semidirect product. Similarly, TTG is described as an iterated semidirect product: the right trivialization map of the group $TTG = T_{\varepsilon_2}T_{\varepsilon_1}G$ is obtained by replacing in the right trivialization map for TG, $\Psi_1^R : \mathfrak{g} \times G \to TG$, $(\varepsilon v, g) \mapsto \varepsilon v \cdot g$, the group G by $TG = T_{\varepsilon_1}G$ and letting $\varepsilon = \varepsilon_2$. We get a diffeomorphism

$$\begin{aligned} &\Psi_2^R : \varepsilon_1\varepsilon_2\mathfrak{g} \times \varepsilon_2\mathfrak{g} \times \varepsilon_1\mathfrak{g} \times G \to TTG, \\ &\qquad (\varepsilon_2\varepsilon_1 v_{12}, \varepsilon_2 v_2, \varepsilon_1 v_1, g) \mapsto \varepsilon_2\varepsilon_1 v_{12} \cdot \varepsilon_2 v_2 \cdot \varepsilon_1 v_1 \cdot g, \end{aligned} \tag{23.7}$$

where the products are taken in the group TTG. Similarly, we get the left trivialization

$$\begin{aligned} &\Psi_2^L : G \times \varepsilon_1\mathfrak{g} \times \varepsilon_2\mathfrak{g} \times \varepsilon_1\varepsilon_2\mathfrak{g} \to TTG, \\ &\qquad (g, \varepsilon_1 v_1, \varepsilon_2 v_2, \varepsilon_2\varepsilon_1 v_{12}) \mapsto g \cdot \varepsilon_1 v_1 \cdot \varepsilon_2 v_2 \cdot \varepsilon_2\varepsilon_1 v_{12}. \end{aligned} \tag{23.8}$$

Clearly, both trivializations define linear structures on the fiber $(TTG)_e$ and hence on all other fibers. We will show below (Theorem 23.5) that both linear structures indeed define linear connections on TG.

Theorem 23.4. *With respect to the right trivialization Ψ_2^R, the group structure of the subgroup $(TTG)_e \cong \varepsilon_1\varepsilon_2\mathfrak{g} \times \varepsilon_2\mathfrak{g} \times \varepsilon_1\mathfrak{g}$ is given by the product*

$$\begin{aligned} &\Psi_2^R(\varepsilon_1\varepsilon_2 v_{12}, \varepsilon_2 v_2, \varepsilon_1 v_1) \cdot \Psi_2^R(\varepsilon_1\varepsilon_2 w_{12}, \varepsilon_2 w_2, \varepsilon_1 w_1) \\ &= \Psi_2^R(\varepsilon_1\varepsilon_2(v_{12} + w_{12} + [v_1, w_2]), \varepsilon_2(v_2 + w_2), \varepsilon_1(v_1 + w_1)) \end{aligned} \tag{23.9}$$

and inversion

$$(\Psi_2^R(\varepsilon_1\varepsilon_2 v_{12}, \varepsilon_2 v_2, \varepsilon_1 v_1))^{-1} = \Psi_2^R(\varepsilon_1\varepsilon_2([v_1, v_2] - v_{12}), -\varepsilon_2 v_2, -\varepsilon_1 v_1). \tag{23.10}$$

Proof. Using the Commutation Rules (23.1) – (23.4), we get

$$\begin{aligned} &\Psi_2^R(\varepsilon_1\varepsilon_2 v_{12}, \varepsilon_2 v_2, \varepsilon_1 v_1) \cdot \Psi_2^R(\varepsilon_1\varepsilon_2 w_{12}, \varepsilon_2 w_2, \varepsilon_1 w_1) \\ &\qquad = \varepsilon_1\varepsilon_2 v_{12} \cdot \varepsilon_2 v_2 \cdot \varepsilon_1 v_1 \cdot \varepsilon_1\varepsilon_2 w_{12} \cdot \varepsilon_2 w_2 \cdot \varepsilon_1 w_1 \\ &\qquad = \varepsilon_1\varepsilon_2(v_{12} + w_{12}) \cdot \varepsilon_2 v_2 \cdot \varepsilon_1\varepsilon_2[v_1, w_2] \cdot \varepsilon_2 w_2 \cdot \varepsilon_1 v_1 \cdot \varepsilon_1 w_1 \\ &\qquad = \varepsilon_1\varepsilon_2(v_{12} + w_{12} + [v_1, w_2]) \cdot \varepsilon_2(v_2 + w_2) \cdot \varepsilon_1(v_1 + w_1)) \\ &\qquad = \Psi_2^R(\varepsilon_1\varepsilon_2(v_{12} + w_{12} + [v_1, w_2]), \varepsilon_2(v_2 + w_2), \varepsilon_1(v_1 + w_1)) \end{aligned}$$

and

$$\begin{aligned} (\Psi_2^R(\varepsilon_1\varepsilon_2 v_{12}, \varepsilon_2 v_2, \varepsilon_1 v_1))^{-1} &= (\varepsilon_1\varepsilon_2 v_{12} \cdot \varepsilon_2 v_2 \cdot \varepsilon_1 v_1)^{-1} \\ &= \varepsilon_1(-v_1) \cdot \varepsilon_2(-v_2) \cdot \varepsilon_1\varepsilon_2(-v_{12}) \\ &= \varepsilon_2(-v_2) \cdot \varepsilon_1(-v_1) \cdot \varepsilon_1\varepsilon_2([-v_1, -v_2] - v_{12}) \\ &= \varepsilon_1\varepsilon_2([v_1, v_2] - v_{12}) \cdot \varepsilon_2(-v_2) \cdot \varepsilon_1(-v_1) \\ &= \Psi_2^R(\varepsilon_1\varepsilon_2([v_1, v_2] - v_{12}), \varepsilon_2(-v_2), \varepsilon_1(-v_1)) \end{aligned}$$

■

The formula for the group structure of the whole group TTG in the right trivialization is easily deduced from the theorem: considering Ψ_2^R as an identification, it reads

$$(\varepsilon_1\varepsilon_2 v_{12}, \varepsilon_2 v_2, \varepsilon_1 v_1, g) \cdot (\varepsilon_1\varepsilon_2 w_{12}, \varepsilon_2 w_2, \varepsilon_1 w_1, h) =$$
$$(\varepsilon_1\varepsilon_2(v_{12} + \mathrm{Ad}(g)w_{12} + [v_1, \mathrm{Ad}(g)w_2]), \varepsilon_2(v_2 + \mathrm{Ad}(g)w_2), \varepsilon_1(v_1 + \mathrm{Ad}(g)w_1), gh) \tag{23.11}$$

In particular, the group G acts (by conjugation) *linearly* with respect to the linear structure on $(T^2G)_e$, and hence we can transport this linear structure by left- or right translations to all other fibers $(T^2g)_x$, $x \in G$. Clearly, this depends smoothly on x, and hence we have defined two linear structures on TTG which will be denoted by L^L and L^R, called the *left* and *right linear structures.*

Theorem 23.5. *The left and right linear structures L^L and L^R are connections on TG, i.e. they are bilinearly related to all chart structures. The difference $L^L - L^R$ is the tensor field of type $(2,1)$ on G given by the Lie bracket, and this is also the torsion of L^L.*

Proof. Let us prove first that L^L and L^R are bilinearly related to all chart structures. Recall from Section 17.6 that, for any manifold M, a trivialization $TM \cong M \times V$ of the tangent bundle defines a connection on TM. More specifically, in the present case, deriving $\Psi_1^R : \mathfrak{g} \times G \to TG$,

$$T\Psi_1^R : T\mathfrak{g} \times TG \to TTG, \quad (X, h) \mapsto X \cdot h,$$

and composing again with Ψ_1^R, we get a trivialization of TTG:

$$T(\Psi_1^R) \circ (\mathrm{id} \times \Psi_1^R) : \varepsilon_1\varepsilon_2\mathfrak{g} \times \varepsilon_2\mathfrak{g} \times \varepsilon_1\mathfrak{g} \times G = T\mathfrak{g} \times (\mathfrak{g} \times G) \to TTG, \tag{23.12}$$

and the linear structure L induced by this trivialization is in fact a connection (see Lemma 17.7). We claim that $L = L^L$. In fact, writing $\Psi_1 = Tm \circ (\iota \times z)$, where $\iota : \mathfrak{g} \to TG$ and $z : G \to TG$ are the natural inclusions, we get $T\Psi_1 = TTm \circ (T\iota \times Tz)$, which means with $T = T_{\varepsilon_2}$,

$$\begin{aligned} T(\Psi_1^R)(\varepsilon_1\varepsilon_2 v_{12}, \varepsilon_2 v_2, \Psi_1^R(\varepsilon_1 v_1, g)) &= (\varepsilon_1\varepsilon_2 v_{12} \cdot \varepsilon_1 v_1) \cdot (\varepsilon_2 v_2 \cdot g) \\ &= \varepsilon_1 v_1 \cdot \varepsilon_2 v_2 \cdot \varepsilon_1\varepsilon_2 v_{12} \cdot g. \end{aligned}$$

For $g = e$, this agrees with Ψ_2^L, and hence we have $L_e = (L^L)_e$. But both linear structures are invariant under left- and right translations and hence agree on all fibers. We have proved that L^L is a connection. In the same way we see that L^R is a connection.

We show that, on the fiber $(TTG)_e$, the difference between L^L and L^R is the Lie bracket. In fact, using (23.1), we can calculate the map relating both linear structures:

$$\begin{aligned} (\Psi_2^L)^{-1} \circ \Psi_2^R(\varepsilon_2\varepsilon_1 v_{12}, \varepsilon_1 v_2, \varepsilon_1 v_1) &= (\Psi_2^L)^{-1}(\varepsilon_2\varepsilon_1 v_{12} \cdot \varepsilon_2 v_2 \cdot \varepsilon_1 v_1) \\ &= (\Psi_2^L)^{-1}(\varepsilon_1 v_1 \cdot \varepsilon_2 v_2 \cdot \varepsilon_2\varepsilon_1(v_{12} + [v_2, v_1])) \\ &= (\varepsilon_1 v_1, \varepsilon_2 v_2, \varepsilon_1\varepsilon_2(v_{12} + [v_2, v_1])). \end{aligned} \tag{23.13}$$

Identifying $\varepsilon_1\mathfrak{g}\times\varepsilon_2\mathfrak{g}\times\varepsilon_1\varepsilon_2\mathfrak{g}$ and $\varepsilon_1\varepsilon_2\mathfrak{g}\times\varepsilon_2\mathfrak{g}\times\varepsilon_1\mathfrak{g}$ in the canonical way, this means that L^L and L^R are related via the map

$$f_b(u,v,w) = (u,v,w+[u,v]), \tag{23.14}$$

i.e. they are bilinearly related via the Lie bracket.

In order to prove that $L^R - L^L$ is the torsion tensor of L^R, we have to show that $L^R = \kappa \cdot L^L$, where $\kappa : TTG \to TTG$ is the canonical flip. But, since κ commutes with second tangent maps, κ is a group automorphism of (TTG, TTm), whence

$$\begin{aligned} \kappa \circ \Psi_2^L(\varepsilon_1 v_1, \varepsilon_2 v_2, \varepsilon_1\varepsilon_2 v_{12}) &= \kappa(\varepsilon_1 v_1 \cdot \varepsilon_2 v_2 \cdot \varepsilon_1\varepsilon_2 v_{12}) \\ &= \kappa(\varepsilon_1 v_1)\cdot\kappa(\varepsilon_2 v_2)\cdot\kappa(\varepsilon_1\varepsilon_2 v_{12}) \\ &= \varepsilon_2 v_1 \cdot \varepsilon_1 v_2 \cdot \varepsilon_1\varepsilon_2 v_{12} \\ &= \Psi_2^R(\varepsilon_1 v_2, \varepsilon_2 v_1, \varepsilon_1\varepsilon_2 v_{12}) \\ &= \Psi_2^R(\kappa(\varepsilon_1 v_1, \varepsilon_2 v_2, \varepsilon_1\varepsilon_2 v_{12})) \end{aligned} \tag{23.15}$$

and hence $\Psi_2^R = \kappa \circ \Psi_2^L \circ \kappa$, which is equivalent to $L^R = \kappa \cdot L^L$. ■

23.6. *Remarks on the symmetric connection.* If $\frac{1}{2} \in \mathbb{K}$, then the connection $L^S := \frac{L^R + L^L}{2}$ is called the *symmetric connection of* G; it is torsionfree and invariant under left- and right-translations. Moreover, it is invariant under the inversion map $j : G \to G$ because $j.L^L = L^R$. In fact, this is the canonical connection of G seen as a symmetric space, cf. Chapter 26.

Theorem 23.7. (Structure of $(J^2G)_e$.) *For $F = L, R, S$ (the latter in case $\frac{1}{2} \in \mathbb{K}$), the bijections $\Psi_2^F : \varepsilon_1\mathfrak{g}\times\varepsilon_2\mathfrak{g}\times\varepsilon_1\varepsilon_2\mathfrak{g} \to (TTG)_e$ restrict to bijections*

$$J\Psi_2^F : \delta\mathfrak{g}\times\delta^{(2)}\mathfrak{g} \to (J^2G)_e.$$

These three maps coincide and are given by the explicit formula

$$\Psi_2(\delta^{(2)}v, \delta w) = \delta^{(2)}v\cdot\delta w = \varepsilon_1\varepsilon_2 v\cdot\varepsilon_1 w\cdot\varepsilon_2 w,$$

and the group structure of $(J^2G)_e$ is given by

$$\Psi_2(\delta^{(2)}v', \delta v)\cdot\Psi_2(\delta^{(2)}w', \delta w) = \Psi_2(\delta^{(2)}(v'+w'+[v,w]), \delta(v+w)).$$

Inversion in $(J^2G)_e$ is multiplication by the scalar -1.

Proof. We have remarked above (Equation (23.15)) that the canonical flip κ acts on $(TTG)_e$ by $\kappa(\varepsilon_1 v_1\cdot\varepsilon_2 v_2\cdot\varepsilon_1\varepsilon_2 v_{12}) = \varepsilon_1\varepsilon_2 v_{12}\cdot\varepsilon_2 v_1\cdot\varepsilon_1 v_2$. Now all claims follow from Formulae (23.7) – (23.10) by observing that $[v,v]=0$. ■

Finally, for the sake of completeness we add some results that are well-known to hold in the real (say, Banach) case ; these results will not be needed in the sequel.

Lemma 23.8. *The covariant derivatives associated to the three canonical connections of the Lie group G can be described as follows: fix $x \in G$; for $v \in T_xG$ denote by v^L, resp. v^R the unique left- (resp. right-) invariant vector field taking value v at x. Then, for all vector fields $X, Y \in \mathfrak{X}(G)$,*

$$(\nabla^L_X Y)(x) = [Y, (X(x))^R](x),$$
$$(\nabla^R_X Y)(x) = [Y, (X(x))^L](x),$$
$$(\nabla^S_X Y)(x) = \frac{1}{2}[Y, (X(x))^L + (X(x))^R](x).$$

Proof. Note that v^R is constant in the right trivialized picture, and hence $K^R \circ Tv^R = 0$ for the connector of L^R. It follows that

$$(\nabla^R_v Y)(x) = K^R \circ TY \circ v^R(x) = K^R \circ (TY \circ v^R - Tv^R \circ Y)(x)$$
$$= K^R \circ [Y, v^R](x) = [Y, v^R](x).$$

Similarly for ∇^L. The last equality follows by taking the arithmetic mean of the preceding ones. ∎

Theorem 23.9. *The curvature tensor of the connections L^L and L^R vanishes.*

Proof. Let R^R be the curvature tensor of the connection L^R. We use the "classical" characterization of R^R via $R^R(X,Y)Z = ([\nabla_X, \nabla_Y] - \nabla_{[X,Y]})Z$ (Theorem 18.6). Taking values at the origin e and using Lemma 23.8, we see that the right hand side equals $(([\mathrm{ad}(X^R), \mathrm{ad}(Y^R)] - \mathrm{ad}[X^R, Y^R])Z^R)(e)$ which is zero since the space of right-invariant vector fields is a Lie algebra. ∎

Since L^L and L^R have torsion, the preceding result should not be interpreted in the sense that L^L and L^R be flat.

23.10. *The adjoint Maurer-Cartan form.* We denote by ω the tensor field of type $(2,1)$ on G given at each point by the Lie bracket (torsion of L^R). Using the right trivialization of TG, it can also be seen as a $\mathrm{End}(\mathfrak{g})$-valued one-form

$$\omega : TG \cong G \times \mathfrak{g} \to \mathrm{End}(\mathfrak{g}), \quad (g, v) \mapsto \mathrm{ad}(v).$$

We call this the *adjoint Maurer-Cartan form* since our $\mathrm{End}(\mathfrak{g})$-valued form is obtained from the usual $\mathfrak{g}$-valued form by composing with the adjoint representation $\mathrm{ad} : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$, cf. e.g. [Sh97, p. 108].

Theorem 23.11. *The adjoint Maurer-Cartan form satisfies the structure equation*

$$\mathrm{d}\omega(X,Y) + [\omega_X, \omega_Y] = 0$$

for all $X, Y \in \mathfrak{X}(G)$ (where the bracket is a commutator in $\mathrm{End}(\mathfrak{g})$*). The curvature R^S of the connection L^S is given by taking pointwise triple Lie brackets:*

$$R^S(X,Y)Z = \frac{1}{4}\mathrm{d}\omega(X,Y)Z = \frac{1}{4}\omega(\omega(X,Y),Z) = \frac{1}{4}[[X,Y],Z].$$

Proof. We have seen that the curvatures R^L and R^R of the connections L^L, resp. L^R vanish. Thus we can apply the structure equation (18.9), with the difference term $A = \omega$, which leads to the first claim.

Next we apply the structure equation (18.10), but this time with $R = R^L = 0$ and $R' = R^S$, the symmetric curvature. The difference term is now $A = \frac{1}{2}\omega$. Thus $R^S = \frac{1}{2}\mathrm{d}\omega + \frac{1}{8}[\omega, \omega] = -\frac{1}{8}[\omega, \omega] = \frac{1}{4}\mathrm{d}\omega$. ∎

24. The structure of higher order tangent groups

We retain the assumptions from the preceding chapter: G is a Lie group with multiplication $m : G \times G \to G$. We are going to investigate the structure of the groups $T^k G$ for $k \geq 3$.

24.1. *Fundamental commutation relations for elements of the axes.* Let $\varepsilon^\alpha v_\alpha$, $\varepsilon^\beta w_\beta \in (T^k G)_e$ be elements of axes in $(T_k G)_e$. Then the commutator of these two elements in the group $(T^k G)_e$ is given by the following fundamental commutation rules:

$$[\varepsilon^\alpha v_\alpha, \varepsilon^\beta w_\beta] = \begin{cases} \varepsilon^{\alpha+\beta}[v_\alpha, w_\beta] & \text{if } \alpha \perp \beta, \\ 0 & \text{else.} \end{cases} \tag{24.1}$$

In fact, if $\operatorname{supp}(\alpha) \cap \operatorname{supp}(\beta) \neq \emptyset$, then both elements commute: $[\varepsilon^\alpha v_\alpha, \varepsilon^\beta v_\beta] = 0$, because both belong to the tangent space $T_{\varepsilon_i}(T^{k-1}G)$ at the origin for some ε_i (namely for i such that $e_i \in \operatorname{supp}(\alpha) \cap \operatorname{supp}(\beta)$), and the group structure on the tangent space is just vector addition. This proves the second relation. If $\alpha \perp \beta$, then we apply (23.1) with ε_1 replaced by ε^α and ε_2 replaced by ε^β, and we get the first relation. In a unified way, (24.1) may be written

$$[\varepsilon^\alpha v_\alpha, \varepsilon^\beta w_\beta] = \varepsilon^\alpha \varepsilon^\beta [v_\alpha, w_\beta] \tag{24.2}$$

because $\varepsilon^\alpha \varepsilon^\beta = 0$ if $\operatorname{supp}(\alpha) \cap \operatorname{supp}(\beta) \neq \emptyset$. This can also be written

$$\varepsilon^\alpha v_\alpha \cdot \varepsilon^\beta w_\beta = \varepsilon^\alpha \varepsilon^\beta [v_\alpha, w_\beta] \cdot \varepsilon^\beta w_\beta \cdot \varepsilon^\alpha v_\alpha = \varepsilon^\beta w_\beta \cdot \varepsilon^\alpha v_\alpha \cdot \varepsilon^\alpha \varepsilon^\beta [v_\alpha, w_\beta]. \tag{24.3}$$

24.2. *Remark: $(T^k G)_e$ as a filtered group.* Let $G_1 := (T^k G)_e$ and $G_j := \langle \varepsilon^\alpha v_\alpha \,|\, |\alpha| > j,\, v_\alpha \in \mathfrak{g} \rangle$ be the "vertical subgroup" generated by elements from axes with total degree bigger than j. Then it follows from (24.1) that

$$\mathbf{1} = G_k \subset G_{k-1} \subset \ldots \subset G_1$$

is a *filtration* of G_1 in the sense of [Se65, LA 2.2] or [Bou72, II. 4.4]. Recall from [Se65, LA 2.3] or [Bou72] that the associated *graded group* carries a canonical structure of a Lie algebra. This Lie algebra structure is compatible with ours, but in our setting we can go further by constructing special bijections (coming from the three canonical connections) between the graded and the filtered objects.

24.3. *Right and left trivialization of $T^k G$.* Note that, according to our conventions, $T^k G$ is viewed as obtained by successive scalar extensions via $\varepsilon_1, \ldots, \varepsilon_k$, i.e. $T^{k+1}G = T_{\varepsilon_{k+1}}(T^k G)$. The iterated right and left trivialization of $T^k G$ are defined inductively: for $k = 1, 2$ they are given by (23.2) – (23.4); then the right trivialization of $T^k G$ is given by appliying Formula (23.3) to the right trivialization of $H := T^{k-1}G$, with $TH = T_{\varepsilon_k} H$. For instance, the third order right trivialization is

$$\begin{aligned} \Psi_3^R : &(\varepsilon_1\varepsilon_2\varepsilon_3 v_{123}, \varepsilon_3\varepsilon_2 v_{23}, \varepsilon_3\varepsilon_1 v_{13}, \varepsilon_3 v_3, \varepsilon_1\varepsilon_2 v_{12}, \varepsilon_2 v_2, \varepsilon_1 v_1, g) \mapsto \\ &\varepsilon_1\varepsilon_2\varepsilon_3 v_{123} \cdot \varepsilon_3\varepsilon_2 v_{23} \cdot \varepsilon_3\varepsilon_1 v_{13} \cdot \varepsilon_3 v_3 \cdot \varepsilon_1\varepsilon_2 v_{12} \cdot \varepsilon_2 v_2 \cdot \varepsilon_1 v_1 \cdot g \end{aligned}$$

(products taken in T^3G). The important feature is the order in which the product in T^3G has to be taken: here it is the "antilexographic" order (opposite order of the lexicographic order). An easy induction shows that the general formulae are

$$\begin{aligned}
\Psi_k^R &: \bigoplus_{\alpha>0} \varepsilon^\alpha \mathfrak{g} \times G \to T^kG, \quad (\varepsilon^\alpha v_\alpha)_{\alpha>0}, g) \mapsto \prod_{\alpha>0}^{\downarrow} \varepsilon^\alpha v_\alpha \cdot g, \\
\Psi_k^L &: \bigoplus_{\alpha>0} \varepsilon^\alpha \mathfrak{g} \times G \to T^kG, \quad (\varepsilon^\alpha v_\alpha)_{\alpha>0}, g) \mapsto g \cdot \prod_{\alpha>0}^{\uparrow} \varepsilon^\alpha v_\alpha,
\end{aligned} \tag{24.4}$$

where $\prod^\uparrow$ is the product of the elements labelled by $\alpha > 0$ in $(T^kG)_e$ taken with respect to the lexicographic order of the index set and $\prod^\downarrow$ is the product taken in $(T^kG)_e$ in the opposite ("antilexicographic") order.

24.4. *The case $k = 3$ and the Jacobi identity.* As an application of our formalism, let us now give an independent proof of the Jacobi identity (it is essentially equivalent to the arguments used in [Se65]): we calculate both sides of $\varepsilon_3 v_3 \cdot (\varepsilon_2 v_2 \cdot \varepsilon_1 v_1) = (\varepsilon_3 v_3 \cdot \varepsilon_2 v_2) \cdot \varepsilon_1 v_1$ in $(T^3G)_e$ and express the result in the left trivialization: on the one hand,

$$\begin{aligned}
&\varepsilon_3 v_3 \cdot (\varepsilon_2 v_2 \cdot \varepsilon_1 v_1) = \\
&\varepsilon_3 v_3 \cdot (\varepsilon_1 v_1 \cdot \varepsilon_2 v_2 \cdot \varepsilon_1\varepsilon_2 [v_2, v_1]) = \\
&\varepsilon_1 v_1 \cdot \varepsilon_3 v_3 \cdot \varepsilon_1\varepsilon_3[v_3, v_1] \cdot \varepsilon_2 v_2 \cdot \varepsilon_1\varepsilon_2[v_2, v_1] = \\
&\varepsilon_1 v_1 \cdot \varepsilon_2 v_2 \cdot \varepsilon_3 v_3 \cdot \varepsilon_1\varepsilon_3[v_3, v_1] \cdot \varepsilon_2\varepsilon_3[v_3, v_2] \cdot \varepsilon_1\varepsilon_2\varepsilon_3[[v_3, v_1], v_2] \cdot \varepsilon_1\varepsilon_2[v_2, v_1] = \\
&\varepsilon_1 v_1 \cdot \varepsilon_2 v_2 \cdot \varepsilon_1\varepsilon_2[v_2, v_1] \cdot \varepsilon_3 v_3 \cdot \varepsilon_1\varepsilon_3[v_3, v_1] \cdot \varepsilon_2\varepsilon_3[v_3, v_2] \\
&\qquad \cdot \varepsilon_1\varepsilon_2\varepsilon_3([[v_3, v_1], v_2] + [v_3, [v_2, v_1]]) = \\
&\Psi_3^L\Big(\varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2[v_2, v_1] + \varepsilon_3 v_3 + \varepsilon_1\varepsilon_3[v_3, v_1] + \varepsilon_2\varepsilon_3[v_3, v_2] \\
&\qquad + \varepsilon_1\varepsilon_2\varepsilon_3([[v_3, v_1], v_2] + [v_3, [v_2, v_1]]\Big).
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&(\varepsilon_3 v_3 \cdot \varepsilon_2 v_2) \cdot \varepsilon_1 v_1 = \varepsilon_2 v_2 \cdot \varepsilon_3 v_3 \cdot \varepsilon_2\varepsilon_3[v_3, v_2] \cdot \varepsilon_1 v_1 = \\
&\varepsilon_1 v_1 \cdot \varepsilon_2 v_2 \cdot \varepsilon_1\varepsilon_2[v_2, v_1] \cdot \varepsilon_3 v_3 \cdot \varepsilon_1\varepsilon_3[v_3, v_1] \cdot \varepsilon_2\varepsilon_3[v_3, v_2] \cdot \varepsilon_1\varepsilon_2\varepsilon_3[[v_3, v_2], v_1] = \\
&\Psi_3^L\Big(\varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2[v_2, v_1] + \varepsilon_3 v_3 + \varepsilon_1\varepsilon_3[v_3, v_1] + \varepsilon_2\varepsilon_3[v_3, v_2] \\
&\qquad + \varepsilon_1\varepsilon_2\varepsilon_3[[v_3, v_2], v_1]\Big).
\end{aligned}$$

Applying $(\Psi_3^L)^{-1}$ and comparing, we get

$$[[v_3, v_1], v_2] + [v_3, [v_2, v_1]] = [[v_3, v_2], v_1]$$

which is the Jacobi identity. Note that this proof does not make any use of the interpretation of the Lie bracket via left or right invariant vector fields, and note also that the expression of the Jacobi identity that we get here is the one defining the more general class of *Leibniz algebras*, see Remark 24.12 below. – Another, rather tricky proof of the Jacobi identity makes essential use of Theorem 6.2 by combining algebraic and differential geometric arguments: the tangent map of the adjoint action $\mathrm{Ad}_G : G \times \mathfrak{g} \to \mathfrak{g}$ is $\mathrm{Ad}_{TG} : TG \times T\mathfrak{g} \to T\mathfrak{g}$. We know that $\mathrm{Ad}(g)$, for an element $g \in G$, is a Lie algebra automorphism of $\mathfrak{g}$, and hence $\mathrm{Ad}(h)$, for $h \in TG$, is a Lie algebra automorphism of $T\mathfrak{g}$. But Theorem 6.2 implies that, moreover, $\mathrm{Ad}(h)$ is an automorphism of $T\mathfrak{g}$ over the ring $T\mathbb{K}$. Now, it is a simple exercise in linear algebra to show that $\mathrm{Aut}_{T\mathbb{K}}(T\mathfrak{g})$ is a semidirect product of $\mathrm{Aut}_{\mathbb{K}}(\mathfrak{g})$ with the vector group $\mathrm{Der}_{\mathbb{K}}(\mathfrak{g})$ of derivations. Putting things together, we see that, for all $X \in \mathfrak{g}$, $\mathrm{Ad}(\varepsilon X)$ acts like a derivation from $\mathfrak{g}$ to $\varepsilon\mathfrak{g}$, which gives again the Jacobi identity. This proof has the advantage that it is extremely close to the well-known arguments from the theory of finite-dimensional real Lie groups.

24.5. *Action of the symmetric group.* The natural action of the symmetric group Σ_k by automorphisms of $(T^k G)_e$ is described by the following formula: for $\sigma \in \Sigma_k$,

$$\sigma \cdot \prod_{\alpha}^{\uparrow} \varepsilon^{\alpha} v_{\alpha} = \prod_{\alpha}^{\uparrow} \varepsilon^{\sigma.\alpha} v_{\alpha}. \tag{24.4}$$

As a consequence, the push-forward $\Psi^{L,\sigma} = \sigma \circ \Psi^L \circ \sigma^{-1}$ of Ψ^L by σ is given by

$$\Psi^{L,\sigma}((\varepsilon^{\alpha} v_{\alpha})_{\alpha>0}, g) = g \cdot \prod_{\alpha>0}^{\uparrow} \varepsilon^{\sigma.\alpha} v_{\sigma.\alpha}. \tag{24.5}$$

In fact, Formula (24.4) is proved by exactly the same arguments as in case $k = 2$ for $\sigma = \kappa$ (cf. Equation (23.15)). Re-ordering the terms and using the commutation rules (24.1), one can calculate the curvature forms which are defined as in Chapter 18. If G is commutative, then by a change of variables the left-hand side can be re-written $\prod^{\uparrow} \varepsilon^{\alpha} v_{\sigma^{-1}\alpha}$, and this can be interpreted by saying that Ψ^L is Σ_k-invariant. Clearly, then Ψ^L and $\Psi^{L,\sigma}$ coincide. However, if G is non-commutative, then we cannot re-write the formula in the way just mentioned since the order would be disturbed. In particular, for $k > 2$, Ψ_k^L and Ψ_k^R are no longer related via a permutation since reversing the lexicographic order on I_k is not induced by an element of Σ_k acting on I_k.

24.6. *Left and right group structures.* We have defined two canonical charts Ψ_k^F, $F = R, L$, of the group $(T^k G)_e$, and we want to describe the group structure in these charts. The model space is $E = \bigoplus_{\alpha>0} \varepsilon^{\alpha}\mathfrak{g}$, which is just the tangent space $T_0((T^k G)_e)$ – being a linear space, the tangent space is canonically isomorphic to the direct sum of the axes. Via Ψ_k^F $(F = R, L)$ we transfer the group structure to E. This defines a product and an inverse,

$$\begin{aligned} *^F : E \times E \to E, &\quad \Psi_k^F v \cdot \Psi_k^F w = \Psi_k^F (v *^F w), \\ J^F : E \to E, &\quad (\Psi_k^F(v))^{-1} = \Psi_k^F (J^F v), \end{aligned} \tag{24.6}$$

called the *right*, resp. *left group structure on* E. In Theorem 23.3 we have given explicit formulae for these maps in case $k = 2$. For the general case we use the same strategy: just apply the Commutation Rules 24.1 along with $\varepsilon^{\alpha} v \cdot \varepsilon^{\alpha} w = \varepsilon^{\alpha}(v + w)$ in order to re-order the product. Before coming to the general result, we invite the reader to check, using the commutation rules, that for $k=3$ we have the following: with $v=\sum_{\substack{\alpha\in I_3\\ \alpha>0}}\varepsilon^\alpha v_\alpha$, $w=\sum_{\substack{\alpha\in I_3\\ \alpha>0}}\varepsilon^\alpha w_\alpha$, we get

$$\begin{aligned}
&\Psi_3^L(v)\cdot\Psi_3^L(w)\\
&=\varepsilon^{001}v_{001}\cdot\varepsilon^{010}v_{010}\cdot\varepsilon^{011}v_{011}\cdot\varepsilon^{100}v_{100}\cdot\varepsilon^{101}v_{101}\cdot\varepsilon^{110}v_{110}\cdot\varepsilon^{111}v_{111}\\
&\cdot\varepsilon^{001}w_{001}\cdot\varepsilon^{010}w_{010}\cdot\varepsilon^{011}w_{011}\cdot\varepsilon^{100}w_{100}\cdot\varepsilon^{101}w_{101}\cdot\varepsilon^{110}w_{110}\cdot\varepsilon^{111}w_{111}\\
&=\varepsilon^{001}(v_{001}+w_{001})\cdot\varepsilon^{010}(v_{010}+w_{010})\cdot\varepsilon^{011}(v_{011}+w_{011}+[v_{010},w_{001}])\cdot\\
&\varepsilon^{100}(v_{100}+w_{100})\cdot\varepsilon^{101}(v_{101}+w_{101}+[v_{100},w_{001}])\cdot\varepsilon^{110}(v_{110}+w_{110}+[v_{100},w_{010}])\\
&\cdot\varepsilon^{111}(v_{111}+w_{111}+[v_{110},w_{001}]+[v_{101},w_{010}]+[v_{100},w_{011}]+[[v_{100},w_{001}],w_{010}])\\
&=\Psi_3^L\Big(v+w+\varepsilon^{011}[v_{010},w_{001}]+\varepsilon^{101}[v_{100},w_{001}]+\varepsilon^{110}[v_{100},w_{010}]+\\
&\qquad\varepsilon^{111}([v_{110},w_{001}]+[v_{101},w_{010}]+[v_{100},w_{011}]+[[v_{100},w_{001}],w_{010}])\Big).
\end{aligned}$$

Similarly, we get for the inversion:

$$\begin{aligned}
&(\Psi_3^L v)^{-1}=(\varepsilon^{001}v_{001}\cdot\varepsilon^{010}v_{010}\cdot\varepsilon^{011}v_{011}\cdot\varepsilon^{100}v_{100}\cdot\varepsilon^{101}v_{101}\cdot\varepsilon^{110}v_{110}\cdot\varepsilon^{111}v_{111})^{-1}\\
&=\varepsilon^{111}(-v_{111})\cdot\varepsilon^{110}(-v_{110})\cdot\varepsilon^{101}(-v_{101})\cdot\varepsilon^{100}(-v_{100})\cdot\varepsilon^{011}(-v_{011})\cdot\\
&\quad\varepsilon^{010}(-v_{010})\cdot\varepsilon^{001}(-v_{001})\\
&=\varepsilon^{011}(-v_{011})\cdot\varepsilon^{100}(-v_{100})\cdot\varepsilon^{101}(-v_{101})\cdot\varepsilon^{110}(-v_{110})\cdot\varepsilon^{111}(-v_{111}+[v_{100},v_{011}]))\cdot\\
&\quad\varepsilon^{010}(-v_{010})\cdot\varepsilon^{001}(-v_{001})\\
&=\varepsilon^{010}(-v_{010})\cdot\varepsilon^{011}(-v_{011})\cdot\varepsilon^{100}(-v_{100})\cdot\varepsilon^{101}(-v_{101})\cdot\\
&\quad\varepsilon^{110}(-v_{110}+[v_{100},v_{010}])\cdot\varepsilon^{111}(-v_{111}+[v_{100},v_{011}]+[v_{101},v_{010}]))\cdot\varepsilon^{001}(-v_{001})\\
&=\varepsilon^{001}(-v_{001})\cdot\varepsilon^{010}(-v_{010})\cdot\varepsilon^{011}(-v_{011}+[v_{010},v_{001}])\cdot\varepsilon^{100}(-v_{100})\cdot\\
&\quad\varepsilon^{101}(-v_{101}+[v_{100},v_{001}])\cdot\varepsilon^{110}(-v_{110}+[v_{100},v_{010}])\cdot\\
&\quad\varepsilon^{111}(-v_{111}+[v_{110},v_{001}]+[v_{101},v_{010}]+[v_{100},v_{011}]-[[v_{100},v_{010}],v_{001}])\\
&=\Psi_3^L\Big(-v+\varepsilon^{011}[v_{010},v_{001}]+\varepsilon^{101}[v_{100},v_{001}]+\varepsilon^{110}[v_{100},v_{010}]+\\
&\quad\varepsilon^{111}([v_{110},v_{001}]+[v_{101},v_{010}]+[v_{100},v_{011}]-[[v_{100},v_{010}],v_{001}])\Big).
\end{aligned}$$

Theorem 24.7. (Left and right product formulae for $(T^kG)_e$.) *With respect to the left trivialization Ψ_k^L of T^kG, we have*

$$\sum_\alpha\varepsilon^\alpha v_\alpha *^L\sum_\beta\varepsilon^\beta w_\beta=\sum_\gamma\varepsilon^\gamma z_\gamma$$

with

$$z_\gamma=v_\gamma+w_\gamma+\sum_{m=2}^{|\gamma|}\sum_{\lambda\in\mathcal{P}_m(\gamma)}[\dots[[v_{\lambda^m},w_{\lambda^1}],w_{\lambda^2}],\dots,w_{\lambda^{m-1}}].$$

The left inversion formula is

$$J^L(\sum_\alpha\varepsilon^\alpha v_\alpha)=\sum_\gamma\varepsilon^\gamma(-v_\gamma+\sum_{m=2}^{|\gamma|}(-1)^m\sum_{\lambda\in\mathcal{P}_m(\gamma)}[\dots[[v_{\lambda^m},v_{\lambda^{m-1}}],v_{\lambda^{m-2}}],\dots,v_{\lambda^1}]),$$

and the right product is given by the formula

$$(v *^R w)_\gamma = v_\gamma + w_\gamma + \sum_{m=2}^{|\gamma|} \sum_{\lambda \in \mathcal{P}_m(\gamma)} [v_{\lambda^{m-1}}, \ldots, [w_{\lambda^1}, v_{\lambda^m}]].$$

In these formulae, partitions λ are always considered as ordered partitions: $\lambda^1 < \ldots < \lambda^m$.

Proof. According to (24.4), we have to show that

$$\prod_{\alpha>0}^{\uparrow} \varepsilon^\alpha v_\alpha \cdot \prod_{\beta>0}^{\uparrow} \varepsilon^\beta w_\beta = \prod_{\gamma>0}^{\uparrow} \varepsilon^\gamma z_\gamma$$

with z_γ as in the claim. Using the Commutation Rules 24.1, we will move all terms of the form $\varepsilon^\beta w_\beta$ to the left until we reach $\varepsilon^\beta v_\beta$; then both terms give rise to a new term $\varepsilon^\beta(v_\beta + w_\beta)$. But every time we exchange $\varepsilon^\beta w_\beta$ on its way to the left with an element $\varepsilon^\alpha v^\alpha$ with $\alpha > \beta$ and $\alpha \perp \beta$, applying the commutation rule we get a term $\varepsilon^{\alpha+\beta}[v_\alpha, w_\beta]$. Before doing the same thing with the next element from the second factor, we may re-order the first factor: since $\alpha > \beta$ and $\alpha \perp \beta$, the term $\varepsilon^{\alpha+\beta}[v_\alpha, w_\beta]$ commutes with all terms between position α and $\alpha + \beta$, and hence it can be pulled to position $\alpha + \beta$ without giving rise to new commutators.

We start the procedure just described with the first element $\varepsilon^{0\ldots01} w_{0\ldots01}$ of the second factor, then take the second element $\varepsilon^{0\ldots10} w_{0\ldots10}$, and so on up to the last element. Each time we get commutators of the type $\varepsilon^{\alpha+\beta}[v'_\alpha, w_\beta]$ with v'_α being the sum of v_α and all commutators produced in the preceding steps. This gives us precisely all iterated commutators of the form

$$[\ldots[v_{\mu^1}, w_{\mu^2}]\ldots, w_{\mu^\ell}]$$

with $\mu^i \perp \mu^j$ ($\forall i \neq j$) and the order conditions

$$\mu^2 < \ldots < \mu^\ell, \qquad \mu^1 > \mu^2, \quad \mu^1 + \mu^2 > \mu^3, \ldots, \mu^1 + \ldots + \mu^{\ell-1} > \mu^\ell.$$

These conditions imply $\mu^1 > \mu^j$ for all $j = 1, \ldots, \ell - 1$ because the supports of the μ^i are disjoint and hence the biggest of the μ^j is bigger than the sum of all the others. Hence $(\lambda^1, \ldots, \lambda^{\ell-1}, \lambda^\ell) := (\mu^2, \ldots, \mu^\ell, \mu^1)$ is a partition, and every partition is obtained in this way. Thus, by a change of variables $\lambda \leftrightarrow \mu$, we get the formula for the left product. Similarly for the right product.

The inversion formula is proved by the same kind of arguments: we write

$$(\Psi_k^L(\sum_\alpha \varepsilon^\alpha v_\alpha))^{-1} = (\prod^{\uparrow} \varepsilon^\alpha v_\alpha)^{-1} = \prod^{\downarrow} \varepsilon^\alpha(-v_\alpha)$$

and start to re-order this product from the left, that is: we exchange the first two terms on the left (they commute and hence this does not yet produce a commutator); then we pull the third element to the first position (in case $k = 2$ this produces already a commutator); then we do the same with the fourth element, and so on. We collect the iteraded commutators and get precisely the formula for J^L from the claim. ∎

Comparing our formulae with the "usual" Campbell-Hausdorff formula (say, for a nilpotent Lie group), one remarks that in the usual formula, inversion is just multiplication by -1, whereas here this is not the case. More generally, the powers X^n with respect to the Campbell-Hausdorff multiplication correspond to multiples nX, whereas here we have, e.g., the following *(left) squaring formula*:

$$(\Psi_k^L(v))^2 = \Psi_k^L\Big(2v + \sum_{\substack{\gamma \\ |\gamma|\geq 2}} \varepsilon^\gamma \big(\sum_{\substack{\lambda\in\mathcal{P}(\gamma) \\ \ell(\gamma)\geq 2}} [\ldots[v_{\lambda^{\ell(\lambda)}}, v_{\lambda^1}]\ldots, v_{\lambda^{\ell(\lambda)-1}}]\big)\Big).$$

Note also that our formulae for J^L and $*^F$ $(F = L, R)$ are multilinear (resp. bi-multilinear) in the sense of Section MA.5. In the following we will give a conceptual explanation of this fact by making the link with the theory of multilinear connections from Part IV. (The reader who has not gone through that part may skip the following theorem and its corollaries.)

Theorem 24.8. *Denote by $\kappa \in \Sigma_k$ the permutation that reverses the order, i.e. $\kappa(i) = k + 1 - i$, $i = 1, \ldots, k$. (For $k = 2$, this is the canonical flip, and for $k = 3$, it is the transposition (13).) Then left- and right trivialization Ψ_k^L and Ψ_k^R are multilinear connections on T^kG (in the sense of* Chapter 16*), and, up to the permutation κ, they coincide with the derived connections $D^{k-1}L^L$, resp. $D^{k-1}L^R$ defined in* Chapter 17*. Put differently, the Dombrowski Splitting Maps Φ_k^F associated to the multilinear connections $D^{k-1}L^F$, $F = L, R$, are related with the trivialization maps via*

$$\Phi_k^F = \Psi_k^{F,\kappa}$$

where $\Psi_k^{F,\kappa}$ is as in Equation (24.5)*.*

Proof. We proceed as in the proof of Theorem 23.5 where the case $k = 2$ of the claim has been proved. Starting with the left-trivialization $\Psi_1^L : G \times \mathfrak{g} \to TG$, we have two possibilities to iterate the procedure:

(a) we replace G by $H := TG$ and write the left-trivialization for TG: $TG \times T\mathfrak{g} \to TTG$, compose again with Ψ_1 and obtain $\Psi_2 : G \times \mathfrak{g} \times \mathfrak{g} \times \mathfrak{g} \to TTG$.

(b) we take the tangent map $T\Psi_1^L : TG \times T\mathfrak{g} \to TTG$, compose again with Ψ_1 and obtain another map $\widetilde{\Psi}_2 : G \times \mathfrak{g} \times \mathfrak{g} \times \mathfrak{g} \to TTG$.

Both procedures yield essentially the same result. However, the precise relation depends on the labelling of the copies of TG we use: as to convention (a) we agreed in Section 23.3 to write $H = T_{\varepsilon_1}G$ so that $TTG = TH = T_{\varepsilon_2}T_{\varepsilon_1}G$. Thus, in (b), we must take $T = T_{\varepsilon_2}$, whence the copy TG appearing in the formula really is $T_{\varepsilon_2}G$, but up to exchanging the order of ε_1 and ε_2 both constructions are the same. Now, when iterating the construction, in the k-th step, we must take the *last* tangent functor in (b) with respect to ε_1, i.e. we must use the order of scalar extensions $\varepsilon_k, \ldots, \varepsilon_1$ instead of the usual order $\varepsilon_1, \ldots, \varepsilon_k$, and then constructions (a) and (b) coincide. According to Lemma 17.7, construction (b) corresponds to the construction of the derived linear structure $D^{k-1}L^L$, where due to the change of order the permutation κ appears in the final result. ■

Corollary 24.9. *Left- and right products $*^L$ and $*^R$ can be interpreted as higher covariant differentials in the sense of* Section 16.3*: $*^F$ is the higher differential of* *the group multiplication $m : G \times G \to G$ with respect to the linear structures derived from $L^F \times L^F$ and L^F, $F = L, R$, and inversion J^F is the higher covariant differential of group inversion.*

Proof. Since left, resp. right trivialization are the linearizations associated to the connection L, the claim follows directly from the definition of the higher order covariant differentials in Section 16.3. ∎

The corollary gives an *a priori* proof of the fact that the left and right product formulae are bi-multilinear (i.e. multilinear in the sense of MA.5 in both arguments) and that J^F is multilinear.

Theorem 24.10. (Restriction to jet groups.) *Left and right trivialization of T^kG induce, by restriction, two trivializations of the group J^kG; in other words, the restrictions*

$$\Psi_k^F|_{E^{\Sigma_k}} : E^{\Sigma_k} = \oplus_{j=1}^k \delta^{(j)}\mathfrak{g} \to (J^kG)_e,$$

$F = L, R$, are well-defined.

Proof. For $k = 2$, an elementary proof has been given in the preceding chapter (Theorem 23.7). Let us give, for $k = 3$, an elementary proof in the same spirit: we get from the explicit form of the left product formula stated before Theorem 24.5, by letting $v_1 := v_{001} = v_{010} = v_{100}$, $v_2 := v_{011} = v_{101} = v_{110}$, $v_3 := v_{111}$, $v = \delta v_1 + \delta^{(2)} v_2 + \delta^{(3)} v_3$ (and similarly for w):

$$\begin{aligned}
\Psi_3^L(v) \cdot \Psi_3^L(w) &= \varepsilon^{001}(v_1 + w_1) \cdot \varepsilon^{010}(v_1 + w_1) \cdot \varepsilon^{011}(v_2 + w_2 + [v_1, w_1]) \cdot \\
&\qquad \varepsilon^{100}(v_1 + w_1) \cdot \varepsilon^{101}(v_2 + w_2 + [v_1, w_1]) \cdot \varepsilon^{110}(v_2 + w_2 + [v_1, w_1]) \cdot \\
&\qquad \varepsilon^{111}(v_3 + w_3 + [v_2, w_1] + [v_2, w_1] + [v_1, w_2] + [[v_1, w_1], w_1]) \\
&= \Psi_3^L\Big(v + w + (\varepsilon^{011} + \varepsilon^{101} + \varepsilon^{110})[v_1, w_1] \\
&\qquad + \varepsilon^{111}(2[v_2, w_1] + [v_1, w_2] + [[v_1, w_1], w_1])\Big) \\
&= \Psi_3^L\Big(v + w + \delta^{(2)}[v_1, w_1] + \delta^{(3)}(2[v_2, w_1] + [v_1, w_2] + [[v_1, w_1], w_1])\Big)
\end{aligned}$$

which shows that the group structure of $(J^3G)_e$ in the left trivialization is given by

$$\begin{aligned}
(\delta v_1 + &\delta^{(2)} v_2 + \delta^{(3)} v_3) *^L (\delta w_1 + \delta^{(2)} w_2 + \delta^{(3)} w_3) = \delta(v_1 + w_1) + \\
&\delta^{(2)}(v_2 + w_2 + [v_1, w_1]) + \delta^{(3)}(w_3 + v_3 + [v_1, w_2] + 2[v_2, w_1] + [[v_1, w_1], w_1]).
\end{aligned}$$

Similarly, we have for the inversion:

$$\Big(\Psi_3^L(\delta v_1 + \delta^{(2)} v_2 + \delta^{(3)} v_3)\Big)^{-1} = \Psi_3^L(-\delta v_1 - \delta^{(2)} v_2 + \delta^{(3)}(-v_3 + [v_2, v_1])) .$$

For general k, the combinatorial structure of the corresponding calculation becomes rather involved. For this reason, we prefer to invoke Theorem 19.2 which in a sense contains the relevant combinatorial arguments: by Theorem 24.8, Ψ_k^L and Ψ_k^R are the Dombrowski splitting maps associated to the connection $(D^{k-1}L^F)^\kappa$ ($F = L, R$). According to Theorem 19.2, such connections are weakly symmetric (jet compatible), i.e. they preserve the Σ_k-fixed subbundles. ∎

24.11. *The left and right product formulae for jets.* According to the preceding result, the product formulae from Theorem 24.5 can be restricted to spaces of Σ_k-invariants: there is a relation of the kind

$$\sum_{i=1}^{k} \delta^{(i)} v_i *^L \sum_{j=1}^{k} \delta^{(j)} w_j = \sum_{r=1}^{k} \delta^{(r)} z_r$$

with $z_r \in \mathfrak{g}$ depending on v and w. For $k = 3$ we have seen in the preceding proof that

$$\begin{aligned} z_1 &= v_1 + w_1, \\ z_2 &= v_2 + w_2 + [v_1, w_1], \\ z_3 &= v_3 + w_3 + 2[v_2, w_1] + [v_1, w_2] + [[v_1, w_1], w_1]. \end{aligned}$$

We will not dwell here on the combinatorial details of the formula in the general case, but content ourself with the case of multiplying first order jets in $J^k G$: we claim that for $v, w \in \mathfrak{g}$,

$$\delta v *^L \delta w = \delta(v + w) + \delta^{(2)}[v, w] + \ldots + \delta^{(k)}[[v, w], \ldots w]. \tag{24.8}$$

In fact, $\delta v = \sum_{j=1}^{k} \varepsilon_j v$, and with the notation of Theorem 24.5, we have $v_\alpha = 0 = w_\beta$ if $|\alpha|, |\beta| > 1$. Thus only partitions λ with $|\lambda^i| \leq 1$ contribute to z_γ, i.e. we have

$$z_\gamma = v_\gamma + w_\gamma + \sum_{\ell=2}^{k} \sum_{\lambda \in \mathcal{P}_\ell(\gamma)} [[v, w], \ldots w].$$

For $|\gamma| = 1$, there are no partitions of length at least 2, whence $z_\gamma = v_\gamma + w_\gamma$. This gives rise to the terms $\delta v + \delta w$. For $|\gamma| = j > 1$, just one partition contributes to z_γ, giving rise to a term $\delta^{(j)}[[v, w], \ldots w]$. This proves (24.8). – Note that, since $[v, v] = 0$,

$$\Psi_k^L(\delta v) = \varepsilon_1 v \cdot \ldots \cdot \varepsilon_k v = \varepsilon_k v \cdot \ldots \cdot \varepsilon_1 v = \Psi_k^R(\delta v),$$

i.e. on $\delta\mathfrak{g}$, left and right trivialization coincide; for elements of the form $\delta^{(j)} v$ with $j > 1$ this will no longer be the case.

It would be interesting to investigate further the combinatorial structures involved in the product and inversion formulae of $(J^k G)_e$. In the case of arbitrary characteristic, these formulae are in a way the best results one can get. In characteristic zero, one can use a "better" trivialization of $(T^k G)_e$ given by the *exponential map*.

24.12. *Remark on the synthetic approach.* In the recent paper [Did07], Manon Didry has shown that, starting with an arbitrary Lie algebra $\mathfrak{g}$ over an arbitrary commutative ring $\mathbb{K}$, the formula from Theorem 24.7 defines a group structure on the $\mathbb{K}$-module $\oplus_{\alpha > 0} \varepsilon^\alpha \mathfrak{g}$. Moreover, the symmetric group still acts by automorphisms, so that the jet groups $\oplus_{j=1}^{k} \delta^{(j)} \mathfrak{g}$ are defined and form a projective system. In some sense, these are the "formal" results needed for a possible generalisation of Lie's Third Theorem. A surprising feature of this "synthetic" approach is that it works not only for Lie algebras, but more generally for *Leibniz algebras* (these algebras have been introduced by J.-L. Loday; their product is not necessarily skew-symmetric, but the Jacobi-identity in its form given in Section 24.4 still holds). However, if the algebra is not a Lie algebra, then the symmetric group in general no longer acts by automorphisms, so that the groups we obtain are, in some sense, no longer "integrable".

25. Exponential map and Campbell-Hausdorff formula for jet groups

25.1. *$(T^kG)_e$ as a polynomial group.* Recall from Chapter PG, Example PG.2 (2), that $(T^kG)_e$ satisfies the properties of a *polynomial group*: there is chart (given by left or right trivialization) such that, in this chart, group multiplication is polynomial, and the iterated multiplication maps

$$m^{(j)} : (T^kG)_e \times \ldots \times (T^kG)_e \to (T^kG)_e, \quad (v_1, \ldots, v_j) \mapsto v_1 \cdots v_j$$

are polynomial maps of degree bounded by k.

Theorem 25.2. *Assume that G is a Lie group over a non-discrete topological field $\mathbb{K}$ of characteristic zero. Then there exists a unique map $\exp : (T^k\mathfrak{g})_0 \to (T^kG)_e$ such that:*

(1) *the representation of $\exp$ with respect to left trivialization, $\exp_L := (\Psi_k^L)^{-1} \circ \exp : (T^k\mathfrak{g})_0 \to (T^k\mathfrak{g})_0$, is a polynomial map,*
(2) $T_0 \exp = \mathrm{id}_{(T^k\mathfrak{g})_0}$,
(3) *for all $n \in \mathbb{Z}$ and $v \in (T^k\mathfrak{g})_0$, $\exp(nv) = (\exp(v))^n$.*

The map $\exp$ is bijective and has a polynomial inverse $\log$. Explicitly, the polynomials $\exp_L$ and $\log_L := \log \circ \Psi_k^L$ are given by

$$(\exp_L(v))_\gamma = v_\gamma + \sum_{j=2}^{k} \frac{1}{j!} \sum_{\lambda \in \mathcal{P}_j(\gamma)} [[v_{\lambda^j}, v_{\lambda^{j-1}}], \ldots, v_{\lambda^1}]$$

$$(\log_L(v))_\gamma = v_\gamma + \sum_{j=2}^{k} \frac{(-1)^{j-1}}{j} \sum_{\lambda \in \mathcal{P}_j(\gamma)} [[v_{\lambda^j}, v_{\lambda^{j-1}}], \ldots, v_{\lambda^1}].$$

Proof. Existence and uniqueness of log and exp follow from Theorem PG.6. The explicit formulae are obtained by identifying the terms $\psi_p(x)$ and $\psi_{p,p}(x)$ appearing in the explicit formulae from Theorem PG.6 : by definition, $\psi_1(v) = v$, and comparing the Left Inversion Formula from Theorem 24.7 with Formula (PG.8) for the inversion, $v^{-1} = \sum_j (-1)^j \psi_j(v)$, we get by induction, using that $\psi_{p,m} = 0$ for $m < p$ (i.e., ψ_p vanishes at 0 of order at least p),

$$(\psi_j(v))_\gamma = (\psi_{j,j}(v))_\gamma = \sum_{\lambda \in \mathcal{P}_j(\gamma)} [[v_{\lambda^j}, v_{\lambda^{j-1}}], \ldots, v_{\lambda^1}].$$

Inserting this in the formulae from Theorem PG.6, we get the claim. ∎

25.3. It is clear that exp can also be expressed by a polynomial with respect to the right trivialization. It would be interesting to have also an "absolute formula" for exp, i.e., a formula that does not refer to any trivialization. For $k = 2$, such a formula is given by

$$\begin{aligned}\exp(\varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12}) &= \varepsilon_1 v_1 \cdot \varepsilon_2 v_2 \cdot \varepsilon_1\varepsilon_2 (v_{12} + \frac{1}{2}[v_2, v_1]) \\ &= \varepsilon_2 \frac{v_2}{2} \cdot \varepsilon_1 v_1 \cdot \varepsilon_1\varepsilon_2 v_{12} \cdot \varepsilon_2 \frac{v_2}{2} \\ &= \varepsilon_2 \frac{v_2}{2} \cdot \varepsilon_1 \frac{v_1}{2} \cdot \varepsilon_1\varepsilon_2 v_{12} \cdot \varepsilon_1 \frac{v_1}{2} \cdot \varepsilon_2 \frac{v_2}{2},\end{aligned}$$

where the last expression appears to be fairly symmetric (it is a special case of the formula for the canonical connection of a symmetric space, see Chapter 26). It is not at all obvious how to generalize this formula to the case of general k.

Theorem 25.4. *The exponential map commutes with Lie group automorphisms $\varphi : (T^kG)_e \to (T^kG)_e$ in the sense that $\varphi \circ \exp = \exp \circ T_0\varphi$, and it can be extended in a G-invariant way to a trivialization of T^kG, again denoted by $\exp : G \times (\oplus_\alpha \varepsilon^\alpha \mathfrak{g}) \to T^kG$. This trivialization is a totally symmetric multilinear connection on T^kG. The restriction of $\exp$ to a map $(J^k\mathfrak{g})_0 \to (J^kG)_e$ defines an exponential map of $(J^kG)_e$.*

Proof. One easily checks that both $\exp$ and $\varphi \circ \exp \circ (T_0\varphi)^{-1}$ satisfy Properties (1), (2), (3) of an exponential map and hence agree by the uniqueness statement of Theorem 25.2. Applying this to the inner automorphisms $(T^kc_g)_e$ (where $c_g : G \to G$ is conjugation by $g \in G$), we see that $\exp$ can be transported in a well-defined way to any fiber of T^kG over G, thus defining the trivialization map of T^kG. On the fiber over e, it is multilinearly related to all chart structures (by the explicit formula from Theorem 25.2, which clearly is multilinear), and hence the same is true in any other fiber. Thus $\exp$ is a multilinear connection. This connection is totally symmetric because the symmetric group Σ_k acts by automorphisms of T^kG, and we have just seen that the exponential map is invariant under automorphisms. Therefore the restriction of $\exp$ to $(J^kG)_e$ is well-defined, and the uniqueness statement of Theorem PG.6 implies that it is then "the" exponential map of the polynomial group $(J^kG)_e$. ∎

Theorem 25.5. *With respect to the exponential connection from the preceding theorem, the group structure of $(T^kG)_e$ and the one of $(J^kG)_e$ is given by the Campbell-Hausdorff multiplication with respect to the nilpotent Lie algebras $(T^k\mathfrak{g})_0$, resp. $(J^k\mathfrak{g})_0$.*

Proof. This follows from Theorem PG.8 which states that the group structure with respect to the exponential map is given by the Campbell-Hausdorff formula. ∎

Recall that the classical Campbell-Hausdorff formula for Lie groups is given by

$$X * Y := \log(\exp(X) \cdot \exp(Y)) = X + \sum_{\substack{k,m \geq 0 \\ p_i + q_i > 0}} \frac{(-1)^k}{(k+1)(q_1 + \ldots q_k + 1)} \cdot \frac{(\operatorname{ad} X)^{p_1} (\operatorname{ad} Y)^{q_1} \cdots (\operatorname{ad} X)^{p_k} (\operatorname{ad} Y)^{q_k} (\operatorname{ad} X)^m}{p_1! q_1! \cdots p_k! q_k! m!} Y,$$

and the first four terms are:

$$X * Y = X + Y + \frac{1}{2}[X,Y] + \frac{1}{12}([X,[X,Y]] + [Y,[Y,X]]) + \frac{1}{24}[X,[Y,[X,Y]]] + \ldots$$

(cf., e.g., [T04] for the fourth term). In particular, we may apply this to the nilpotent Lie algebra

$$(J^k\mathfrak{g})_0 = \delta\mathfrak{g} \oplus \delta^{(2)}\mathfrak{g} \oplus \ldots \oplus \delta^{(k)}\mathfrak{g} \subset \mathfrak{g} \otimes_{\mathbb{K}} J^k\mathbb{K}$$

where $\mathfrak{g}$ is an arbitrary $\mathbb{K}$-Lie algebra. Elements are multiplied by the rule

$$\delta^{(i)}\delta^{(j)} = \binom{i+j}{i} \delta^{(i+j)}, \qquad [\delta^{(i)}X, \delta^{(j)}Y] = \binom{i+j}{i} \delta^{(i+j)}[X,Y].$$

We re-obtain for $k=2$ the result from Theorem 23.7:

$$(\delta v_1 + \delta^{(2)} v_2) * (\delta w_1 + \delta^{(2)} w_2) = \delta(v_1 + w_1) + \delta^{(2)}(v_2 + w_2 + [v_1, w_1]),$$

and for $k=3$ we have:

$$\begin{aligned}(\delta v_1 + \delta^{(2)} v_2 + \delta^{(3)} v_3) &* (\delta w_1 + \delta^{(2)} w_2 + \delta^{(3)} w_3) = \\ &\delta(v_1 + w_1) + \\ &\delta^{(2)}(v_2 + w_2 + [v_1, w_1]) + \\ &\delta^{(3)}(w_3 + v_3 + \frac{3}{2}[v_1, w_2] + \frac{3}{2}[v_2, w_1] + \frac{1}{2}([v_1, [v_1, w_1]] + [w_1, [w_1, v_1]])).\end{aligned}$$

Note that this formula makes sense whenever 2 is invertible in $\mathbb{K}$. Compare also with the formula from the proof of Theorem 24.10 which is valid in arbitrary characteristic.

25.6. *Example: Associative continuous inverse algebras.* An associative topological algebra $(V, \cdot)$ over $\mathbb{K}$ is called a *continuous inverse algebra (CIA)* if the set $V^\times$ of invertible elements is open in V and inversion $i : V^\times \to V$ is continuous. We assume that V has a unit element $\mathbf{1}$. It is easily proved that multiplication and inversion are actually smooth, and hence $G = V^\times$ is a Lie group. We consider V as a global chart of G and identify $(T^k G)_e$ with $\mathbf{1} + (T^k V)_0$ where $(T^k V)_0 = \oplus_{\alpha > 0} \varepsilon^\alpha V$. The group $(T^k G)$ is just the group of invertible elements in the associative algebra $T^k V = V \otimes_{\mathbb{K}} T^k \mathbb{K}$. The inverse of $\mathbf{1} + \varepsilon v$ is $\mathbf{1} - \varepsilon v$, and hence we get the group commutator

$$[\mathbf{1} + \varepsilon_1 v, \mathbf{1} + \varepsilon_2 w] = (\mathbf{1} + \varepsilon_1 v)(\mathbf{1} + \varepsilon_2 w)(\mathbf{1} - \varepsilon_1 v)(\mathbf{1} - \varepsilon_2 v) = \mathbf{1} + \varepsilon_1 \varepsilon_2 (vw - wv),$$

which implies by the commutation relations (23.1) that the Lie algebra of G is V with the usual commutator bracket.

Proposition 25.7. *Let $G = V^\times$ be the group of invertible elements in a continuous inverse algebra. Then the group $(T^k G)_e = (T^k V)_0^\times$ is a polynomial group with respect to the natural chart given by the ambient $\mathbb{K}$-module $(T^k V)_0$, and the exponential map with respect to this polynomial group structure is given by the usual exponential formula*

$$\exp(\sum_{\alpha > 0} \varepsilon^\alpha v_\alpha) = \sum_{j=0}^{k} \frac{1}{j!} (\sum_{\alpha > 0} \varepsilon^\alpha v_\alpha)^j.$$

Proof. The group $(T^k G)_e$ is polynomial in the natural chart since the algebra $(T^k V)_0$ is nilpotent of order k, and hence the degree of the iterated product maps $m^{(j)}$ is bounded by k. Clearly, there is just one homogeneous component of $m^{(j)}$ and we have $\psi_j(x) = \psi_{j,j}(x) = x^j$. Inserting this in the formulae from Theorem PG6, we get the usual formulae for the exponential map and the logarithm. ■

25.8. *Comparing the exponential map with right trivialization.* Expanding all terms in the formula for exp from the preceding proposition, we get

$$\exp(\sum_{\alpha>0}\varepsilon^\alpha v_\alpha)=\sum_\alpha \varepsilon^\alpha \sum_\ell \frac{1}{\ell!}\sum_{\Lambda\in\mathcal{P}_\ell(\alpha)}\sum_{\sigma\in\Sigma_\ell} v_{\sigma.\Lambda}$$
$$=\sum_\ell \frac{1}{\ell!}\sum_{\Lambda\in \mathrm{Part}_\ell(I_k)} \varepsilon^{\underline{\Lambda}}\sum_{\sigma\in\Sigma_\ell} v_{\sigma.\Lambda}$$
$$=\mathbf{1}+\sum_\alpha \varepsilon^\alpha v_\alpha + \sum_{\substack{\Lambda:\\ \ell(\Lambda)=2}} \varepsilon^{\underline{\Lambda}} \frac{v_{\Lambda^1}v_{\Lambda^2}+v_{\Lambda^2}v_{\Lambda^1}}{2} + \sum_{\substack{\Lambda:\\ \ell(\Lambda)=3}} \varepsilon^{\underline{\Lambda}} \frac{v_{\Lambda^1}v_{\Lambda^2}v_{\Lambda^3}+\ldots}{6}+\ldots \tag{25.3}$$

On the other hand, the trivialization map Ψ_k^R is in the chart T^kV described by

$$\Psi_k^R(\sum_{\alpha>0}\varepsilon^\alpha v_\alpha)=\prod_\alpha^{\downarrow}(\mathbf{1}+\varepsilon^\alpha v_\alpha)=\mathbf{1}+\sum_\alpha\varepsilon^\alpha\sum_{\lambda\in P(\alpha)} v_{\lambda^1}\cdot\ldots\cdot v_{\lambda^\ell}$$
$$=e+\sum\varepsilon^\alpha v_\alpha+\sum_{\ell(\Lambda)=2}\varepsilon^{\underline{\Lambda}}v_{\Lambda^1}v_{\Lambda^2}+\sum_{\ell(\Lambda)=3}\varepsilon^{\underline{\Lambda}}v_{\Lambda^1}v_{\Lambda^2}v_{\Lambda^3}+\ldots, \tag{25.4}$$

where partitions are considered as ordered partitions. Comparing (25.3) and (25.4), we see that Ψ_k^R is defined by making one particular choice among the $\ell!$ terms in the sum in exp and forgetting division by $\ell!$, and in turn exp is obtained by symmetrizing each term of Ψ_k^R. Here, "symmetrization" is to be understood with respect to the natural linear structure of T^kV. In the case of a general Lie group, it seems much more difficult to understand the passage from the right (or left) trivialization to the exponential connection in terms of a suitable symmetrization procedure.

25.9. *Relation with an exponential map of G.* We say that a Lie group G *has an exponential map* if the following holds: for every $v\in\mathfrak{g}$ there exists a unique Lie group homomorphism $\gamma_v:\mathbb{K}\to G$ such that $\gamma_v'(0)=v$ and such that the map $\exp_G$ defined by

$$\exp_G:\mathfrak{g}\to G,\quad v\mapsto\gamma_v(1)$$

is smooth. In general, a Lie group will not have an exponential map; but if it does, then the higher differentials of exp are given by the exponential map $\exp =: \exp^{(k)}$ from Theorem 25.2, i.e.

$$(T^k\exp_G)_o=\exp^{(k)}:(T^k\mathfrak{g})_0\to(T^kG)_e. \tag{25.5}$$

In fact, it follows immediately from the definitions that $(T^k\exp_G)_0$ has the properties (1), (2), (3) from Theorem 25.2 and hence agrees with $\exp^{(k)}$. It is instructive to check (25.5) directly in the case of a CIA, using the Derivation Formula (7.19): let $p^j(x)=x^j$ be the j-th power in V. Then $d^ip_j(0)=0$ if $i\neq j$ and

$$d^jp_j(0)(v_1,\ldots,v_j)=\sum_{\sigma\in\Sigma_j}v_{\sigma(1)}\cdots v_{\sigma(j)}$$

is the total polarization of p_j. Thus

$$T^k\exp(0)(v)=\sum_\alpha\varepsilon^\alpha\sum_{\ell=1}^{|\alpha|}\sum_{\Lambda\in\mathcal{P}_\ell(\alpha)}\frac{1}{\ell!}d^\ell p_\ell(0)(v_\Lambda)$$

agrees with (25.3).

26. The canonical connection of a symmetric space

26.1. *Abelian symmetric spaces.* We return to the study of symmetric spaces and start by some general (in fact, purely algebraic) remarks on *abelian* symmetric (or reflection) spaces (Chapter 5). In the real, finite-dimensional case, connected abelian symmetric spaces are always products of tori and vector spaces (cf. [Lo69]), hence are of group type (cf. Section 5.11). In our general context, it is certainly not true that all abelian symmetric spaces are of group type, but we will single out a property (the "unique square root property") which permits to define a compatible group structure. – Recall first that a symmetric (or reflection) space is called *abelian* if the group $G(M)$ generated by all $\sigma_x \circ \sigma_y$, $x, y \in M$, is abelian. In particular, if $o \in M$ is a base point, the operators $Q(x) = \sigma_x \circ \sigma_o$, $x \in M$, commute among each other. For instance, if $(V, \cdot)$ is a commutative group, then

$$\mu(u, v) = uv^{-1}u = u^2v^{-1} \tag{26.1}$$

defines on V the structure of an abelian symmetric space, and $Q(v)x = v^2x$ is translation by v^2. We say that a group M (resp. a symmetric space M with base point o) *admits unique square roots* if, for all $v \in M$ there is a unique element $\sqrt{v} \in M$ such that $(\sqrt{v})^2 = v$ (resp. $\mu(\sqrt{v}, o) = v$, which is the same as $Q(\sqrt{v})o = v$).

Proposition 26.2. *There is a bijection between*

(1) *abelian groups $(V, \cdot)$ admitting unique square roots, and*

(2) *abelian reflection spaces with base point (V, μ, o) admitting unique square roots.*

The bijection is given by associating to $(V, \cdot)$ the symmetric space structure (26.1)*, and by associating to (V, μ, o) the group structure given by*

$$u \cdot v = \mu(\sqrt{u}, \mu(e, v)) = Q(\sqrt{u})v = Q(\sqrt{u})Q(\sqrt{v})o. \tag{26.2}$$

Proof. In this proof and on the following pages, the "fundamental formula" (5.6) will be frequently used without further comments. – We have already seen that (26.1) is an (abelian) reflection space structure; the square roots are unique since the squaring operation is the same as the one for the group. Let us prove that (26.2) defines an abelian group law on V. We have $u \cdot v = v \cdot u$ since the operators $Q(\sqrt{u})$ and $Q(\sqrt{v})$ commute by our assumption. Clearly, $v \cdot o = o \cdot v = v$. In order to establish associativity, note first that

$$\sqrt{Q(x)y} = Q(\sqrt{x})\sqrt{y}. \tag{26.3}$$

In fact, we have $Q(\sqrt{x})^2 = Q(Q(\sqrt{x})o) = Q(x)$ and hence

$$Q(Q(\sqrt{x})\sqrt{y})o = Q(\sqrt{x})^2 Q(\sqrt{y})o = Q(x)y,$$

whence (26.3). Using this, associativity follows:

$$\begin{aligned} u \cdot (v \cdot w) &= Q(\sqrt{u})Q(\sqrt{v})w, \\ (u \cdot v) \cdot w &= Q(\sqrt{Q(\sqrt{u})v})w = Q(Q(\sqrt{\sqrt{u}})\sqrt{v})w \\ &= Q(\sqrt{\sqrt{u}})^2 Q(\sqrt{v})w = Q(\sqrt{u})Q(\sqrt{v})w. \end{aligned}$$

Similarly, we see that the inverse of x is $x^{-1} = Q(x)^{-1}x$, and thus $(V,\cdot,o)$ is an abelian group; it admits unique square roots since the squaring operation is the same as the one of (V,μ,o). Finally, it is clear that both constructions are inverse to each other. ∎

In the sequel, abelian groups will mainly be written additively, so that v^2 corresponds to $2v$ and $\sqrt{v}$ corresponds to $\frac{1}{2}v$. Then (26.1) and (26.2) read

$$\mu(u,v) = 2u - v, \qquad u + v = \mu(\frac{1}{2}u, \mu(0,v)) = Q(\frac{u}{2})v = Q(\frac{u}{2})Q(\frac{v}{2})o. \tag{26.4}$$

Now let (M,μ) be an arbitrary symmetric space over $\mathbb{K}$. Recall that the higher tangent bundles $(T^kM, T^k\mu)$ are again symmetric spaces. The purpose of this chapter is to prove the following theorem:

Theorem 26.3. *Assume (M,μ) is a symmetric space over $\mathbb{K}$. There exists a linear connection L on TM which is uniquely determined by one of the following two equivalent properties:*

(a) *L is invariant under all symmetries σ_x, $x \in M$.*

(b) *For all $x \in M$, the symmetric space structure on $(TTM)_x$ is the canonical flat structure associated to the $\mathbb{K}$-module $((TTM)_x, L_x)$: for all $u,v \in (TTM)_x$, $TT\mu(u,v) = 2u - v$.*

In particular, the fibers of TTM over M are abelian. The torsion of L vanishes.

Proof. Let us first prove uniqueness. Assume two linear connections L, L' with Property (a) given. Then $A := L - L'$ is a tensor field of type $(2,1)$ that is again invariant under all symmetries. From $A_x\sigma_x = -\mathrm{id}_{T_xM}$ it follows then that $-A_x(u,v) = A_x(-u,-v) = A_x(u,v)$ and $A_x(u,v) = 0$ for all $x \in M$ and $u,v \in T_xM$. Hence $A = 0$. The same argument also shows that the torsion of L must vanish.

Now assume L is a linear structure on TM satisfying (b). Then multiplication by 2 in the fiber $(TTM)_x$ is given by $2u = TT\mu(u,0_x) = Q(u).0_x$; it is a bijection whose inverse is denoted by 2^{-1}. Thus $(TTM)_x$ has unique square roots, and therefore addition in the fiber $(TTM)_x$ is necessarily given by (26.4): $u+v = TT\mu(2^{-1}u, TT\mu(0,v)) = Q(\frac{u}{2})Q(\frac{v}{2}).0_x$ (here Q is taken with respect to the base point 0_x). Thus the addition is uniquely determined by the symmetric space structure of $(TTM)_x$. But then also the bijection

$$\begin{aligned} \varepsilon_1 T_xM \times \varepsilon_2 T_xM \times \varepsilon_1\varepsilon_2 T_xM &\to (TTM)_x, \\ (\varepsilon_1 v_1, \varepsilon_2 v_2, \varepsilon_1\varepsilon_2 v_{12}) &\mapsto \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12} \end{aligned}$$

depends only on the symmetric space structure. Since the axes are $\mathbb{K}$-modules in a canonical way, it follows that also the action of scalars on $(TTM)_x$ is uniquely determined. Summing up, there is at most one linear structure on TTM satisfying (b), and moreover it follows that such a linear structure will be invariant under all automorphisms of M since it is defined by the symmetric space structure alone (more generally, it will depend functorially on (M,μ)).

Existence of a linear structure L satisfying (b) can be proved in several ways. In the following we give a proof by using charts, which shows also that the linear structure is a connection. The proof relies on the following

Lemma 26.4. *If U is a chart domain of M around x with model space V, then for all $z, w \in (TTM)_x = V \times V \times V$,*

$$TT\mu(z,w) = 2_b z -_b w,$$

(notion for bilinearly related structures being as in Section BA.1*) where*

$$b := b_x := \frac{1}{2} d^2\sigma_x(x) : V \times V \to V$$

is given by the ordinary second differential of the symmetry σ_x at x in the chart $U \subset V$.

Proof. (Cf. [Pos01, p. 64–66] for a coordinate version of the following proof in the finite-dimensional real case.) First of all, if V_1, V_2, V_3 are $\mathbb{K}$-modules and $b \in \mathrm{Bil}(V_1 \times V_2, V_3)$, then the symmetric structure with respect to the linear structure L^b is given by the "Barycenter Formula" (BA.4), with $r = -1$:

$$2_b(u_1,u_2,u_3) -_b (v_1,v_2,v_3) = (2u_1 - v_1, 2u_2 - v_2, 2u_3 - v_3 + 2b(u_1 - v_1, u_2 - v_2)). \tag{26.5}$$

Now choose a chart U around o. From the general expression of the second tangent map of f,

$$TTf(x,u_1,u_2,u_3) = (f(x), df(x)u_1, df(x)u_2, df(x)u_3 + d^2f(x)(u_1,u_2)),$$

for $f = \mu$, we get the chart formula for $TT\mu : TT(U \times U) \to TTM$ at the point (x,x):

$$\begin{aligned} &TT\mu((x,x),(u_1,v_1),(u_2,v_2),(u_3,v_3)) \\ &\quad = \Big(\mu(x,x), d\mu(x,x)(u_1,v_1), d\mu(x,x)(u_2,v_2), \\ &\qquad\qquad d\mu(x,x)(u_3,v_3) + d^2m(x,x)((u_1,v_1),(u_2,v_2))\Big) \\ &\quad = \Big(x, 2v_1 - u_1, 2v_2 - u_2, 2v_3 - u_3 + d^2\mu(x,x)((u_1,v_1),(u_2,v_2))\Big) \end{aligned}$$

Comparing with (26.5), we see that the claim will be proved if we can show that

$$d^2\mu(x,x)((u_1,v_1),(u_2,v_2)) = d^2\sigma_x(x)(u_1 - v_1, u_2 - v_2). \tag{26.6}$$

In order to prove (26.6), we assume that, in our chart, x corresponds to the origin of V and write Taylor expansions of order 2 at the origin for $\sigma_x = \sigma_0$ and for μ,

$$\begin{aligned} \sigma_0(tv) &= -tv + \frac{t^2}{2} d^2\sigma_0(0)(v,v) + t^2 O(t), \\ \mu(tu,tv) &= t(2u - v) + \frac{t^2}{2} d^2\mu(0,0)((u,v),(u,v)) + t^2 O(t) \\ &= t(2u - v) + t^2 c(u,v) + t^2 O(t) \end{aligned}$$

with $c(u,v) := \frac{1}{2}d^2\mu(0,0)((u,v),(u,v))$ (this is not bilinear as a function of (u,v) but is a quadratic form in (u,v)). The defining identity (M2) of a symmetric space, $p = \mu(x,\mu(x,p))$, implies

$$\begin{aligned} tv = \mu(0,\sigma_0(tv)) &= \mu(0, t(-v + \frac{t}{2}d^2\sigma_0(0)(v,v) + tO(t))) \\ &= t(v - \frac{t}{2}d^2\sigma_0(0)(v,v) - tO(t)) + t^2 c((0,-v+O(t))) + t^2O(t) \\ &= tv + t^2(c(0,-v) - \frac{1}{2}d^2\sigma_0(0)(v,v)) + t^2O(t). \end{aligned}$$

Comparing quadratic terms and letting $t=0$ we get

$$d^2\sigma_0(0)(v,v) = 2c((0,-v)) = d^2\mu(0,0)\,((0,-v),(0,-v)).$$

This is a special case of (26.6). The general formula now follows: from the identity $\mu(p,p) = p$ we get

$$\partial_{v,v}\mu(x,x) = \frac{\mu(x+tv,x+tv) - \mu(x,x)}{t}|_{t=0} = \frac{tv}{t}|_{t=0} = v$$

and hence

$$d^2\mu(0,0)((u,w),(v,v)) = \partial_{(u,w)}\partial_{(v,v)}\mu(0,0) = \partial_{(u,w)}v = 0.$$

We write $(u,v) = (v,v) + (u-v,0)$, and so on, and get

$$\begin{aligned} &d^2\mu(0,0)((u_1,v_1),(u_2,v_2)) \\ &\qquad = d^2\mu(0,0)((v_1,v_1)+(u_1-v_1,0),(v_2,v_2)+(u_2-v_2,0)) \\ &\qquad = d^2\mu(0,0)((u_1-v_1,0),(u_2-v_2,0)) = d^2\sigma_0(0)(u_1-v_1,u_2-v_2) \end{aligned}$$

whence (26.6), and the claim is proved. ■

Returning to the proof of the existence part of the theorem, using a chart U as in the lemma, we define a linear structure on $(TTF)_x = T_xM \times T_xM \times T_xM$ by $L_x := L^{b_x}$ with $b_x = \frac{1}{2}d^2\sigma_x(x)$. By the lemma, this linear structure satisfies Property (b) from the theorem. We have already seen that there is at most one linear structure on $(TTM)_x$ satisfying (b), and therefore L_x does not depend on the choice of the chart. Thus $(L_x)_{x\in M}$ is a well-defined linear structure on TTM, and it is a connection since, again by the lemma, it is bilinearly related to all chart structures via bilinear maps b_x depending smoothly on x. We have already remarked that this connection is invariant under all automorphisms and hence satisfies also Property (a). ■

Corollary 26.5. *If (φ_i, U_i) is a chart of M around x such that $\varphi_i(x) = 0$ and $\sigma_x(y) = -y$ for all $y \in U$, then in this chart the symmetric space structure on $(TTM)_x$ is simply the usual flat structure of the vector space $V \times V \times V$.*

Proof. We have $d^2(-\mathrm{id})(x) = 0$ and hence $b_x = 0$. ■

For example, this may applied to the exponential map of a symmetric space, if it exists (cf. Chapter 31, or see [BeNe04]), or to Jordan coordinates of a symmetric space with twist ([Be00]). (In fact, it it is well-known that in the real finite dimensional case and for any torsionfree connection one can find a chart such that the Christoffel symbols at a given point x vanish and hence the linear structure of $(TTM)_x$ is the canonical flat one from the chart. This generalizes to all contexts in which one has an inverse function theorem at disposition.)

26.6. *Dombrowski splitting and the associated spray.* The Dombrowski splitting (Theorem 10.5) for the canonical connection of a symmetric space is given by

$$\begin{aligned} \Phi_1 : \varepsilon_1 TM \times_M \varepsilon_2 TM \times_M \varepsilon_1\varepsilon_2 TM &\to TTM, \\ (x; \varepsilon_1 v_1, \varepsilon_2 v_2, \varepsilon_1\varepsilon_2 v_{12}) \mapsto \varepsilon_1 v_1 &+ \varepsilon_2 v_2 + \varepsilon_1\varepsilon_2 v_{12} \\ &= \varepsilon_2 v_2 + (\varepsilon_1 v_1 + \varepsilon_1\varepsilon_2 v_{12}) \\ &= Q_x(\varepsilon_2 \frac{v_2}{2}) Q_x(\varepsilon_1 \frac{v_1}{2}) \varepsilon_1\varepsilon_2 v_{12} \\ &= Q_x(\varepsilon_2 \frac{v_2}{2}) Q_x(\varepsilon_1 \frac{v_1}{2}) Q_x(\varepsilon_1\varepsilon_2 \frac{v_{12}}{2}).0_x, \end{aligned} \tag{26.7}$$

where $Q_x(y) = s_y s_x$ is the quadratic map in TTM with respect to the base point $x = 0_x$ (origin in $(TTM)_x$). Restricting to invariants under the canonical flip κ, we get a bijection of vector bundles over M

$$\begin{aligned} &\delta TM \times_M \delta^{(2)} TM \to J^2 M, \\ &\quad (x; \delta u, \delta^{(2)} w) \mapsto u + \kappa(u) + w = Q_x(\varepsilon_2 \frac{u}{2}) Q_x(\varepsilon_1 \frac{u}{2}) Q_x(\varepsilon_1\varepsilon_2 \frac{w}{2}).0_x \end{aligned}$$

which for $w = 0$ gives the spray

$$TM \to J^2 M, \quad (x; u) \mapsto u + \kappa(u) = Q_x(\varepsilon_2 \frac{u}{2}) \varepsilon_1 u = Q_x(\varepsilon_1 \frac{u}{2}) \varepsilon_2 u.$$

The last expression can be rewritten, with $z : M \to TM$ being the zero section,

$$u + \kappa(u) = TT\mu(\frac{u}{2}, -\kappa(u)) = T\sigma_{\frac{u}{2}}(-\kappa(u)) = T\sigma_{\frac{u}{2}} Tz(-u)$$

since $TT\mu(v, \cdot) = T\sigma_v$ and $\kappa(v) = Tz(v)$. Compare with [Ne02a, Th. 3.4].

Corollary 26.7. *The covariant derivative associated to the canonical connection of a symmetric space M can be expressed in the following way: for $x \in M$ and $v \in T_x M$ let $l_x(v) := \widetilde{v} = \frac{1}{2} \circ Q(v) \circ z \in \mathfrak{X}(M)$ be the canonical vector field extending v (see proof of* Proposition 5.9.*). Then we have, for all $X, Y \in \mathfrak{X}(M)$ and $x \in M$,*

$$(\nabla_X Y)(x) = [l_x(X(x)), Y](x).$$

Proof. Chose a chart such that x corresponds to the origin 0 and let $v := X(0)$. Then the value of

$$T\widetilde{v} = \frac{1}{2} TQ(v) \circ Tz = \frac{1}{2} T\sigma_v \circ T\sigma_0 \circ Tz$$

at the origin is given, in the chart, by

$$d\widetilde{v}(0)w = -\frac{1}{2}d^2\mu(0,0)((v,0),(w,0)) = \frac{1}{2}d^2\sigma_0(0)(v,w)$$

where for the last equality we used (26.6). Hence

$$\begin{aligned}[l_0(X(0)),Y](0) &= [\widetilde{v},Y](0) = dY(0)v - d\widetilde{v}(0)Y(0)\\ &= dY(0)v + \frac{1}{2}d^2\sigma_0(0)(v,Y(0))\\ &= dY(0)v + b_x(v,Y(0))\\ &= (\nabla_X Y)(0)\end{aligned}$$

according to Lemma 12.2. ∎

In [Be00], the canonical covariant derivative of a (real finite-dimensional) symmetric space G/H has been defined by the formula from the preceding lemma. Finally, in the recent work [BD07] it has been shown that a good deal of the theory of TM carries over to so-called *symmetric bundles* F over M, where TF then plays the role of TTM.

27. The higher order tangent structure of symmetric spaces

27.1. In this chapter we study the structure of tangent bundles $T^k M$ of a symmetric space for $k \geq 3$. In order to simplify notation, if $v \in (T^k M)_o$, we will write $Q(v)$ for the quadratic map $Q_o(v) : T^k M \to T^k M$. If the base point $o \in M$ is fixed and if we use the notation

$$Q := Q^{(M)} : M \times M \to M, \quad (x,y) \mapsto Q(x)y = \sigma_x \sigma_o(y)$$

for the quadratic map of M w.r.t. o, then the quadratic map of $T^k M$ w.r.t. 0_o is given by $Q^{(T^k M)} = T^k(Q^{(M)})$.

Theorem 27.2. *Assume (M, μ) is a symmetric space over $\mathbb{K}$ with base point $o \in M$. For elements from the axes of $(T^3 M)_o$, the following "fundamental commutation relations" hold: for all $u, v, w \in T_o M$ and $i, j \in \{1, 2, 3\}$,*

$$\begin{aligned}
[Q(\varepsilon_i u), Q(\varepsilon_i v)]\varepsilon_j w &= \varepsilon_j w, \\
[Q(\varepsilon_i u), Q(\varepsilon_j v)]\varepsilon_j w &= \varepsilon_j w, \\
[Q(\varepsilon_1 u), Q(\varepsilon_2 v)]\varepsilon_3 w &= \varepsilon_3 w + \varepsilon_1 \varepsilon_2 \varepsilon_3 [u, v, w],
\end{aligned}$$

where the commutator is taken in the group $G(T^3 M)$ and $[u, v, w]$ is the Lie triple product of (M, o).

Proof. The first two relations simply take account of the fact that the fibers of $TTM = T_{\varepsilon_i} T_{\varepsilon_j} M$ are abelian symmetric spaces (Theorem 26.3), i.e. the quadratic operators commute on the fiber.

In order to prove the third relation, recall from Theorem 6.8 that the group $\mathrm{Diff}_{T\mathbb{K}}(TM)$ is a semidirect product of $\mathrm{Diff}_{\mathbb{K}}(M)$ with the vector group $\mathfrak{X}(M)$. It follows that the symmetric space automorphism group $\mathrm{Aut}_{T\mathbb{K}}(TM)$ can be written as a semidirect product of $\mathrm{Aut}_{\mathbb{K}}(M)$ and a vector group which is precisely the subspace $\mathfrak{g} = \mathrm{Der}(M) \subset \mathfrak{X}(M)$ of *derivations of M*, introduced in Section 5.7. (Indeed, the $T\mathbb{K}$-smooth extension $\widetilde{X}$ of a derivation $X : M \to TM$ is again a homomorphism of symmetric spaces, and since it it is invertible, it is an automorphism. We call automorphisms of the form $\widetilde{X} : TM \to TM$ *infinitesimal automorphisms.* Compare with the closely related definition of infinitesimal automorphisms in [Lo69].)

Lemma 27.3. *For all $v \in T_o M$, the map $Q(v) : TM \to TM$ is an infinitesimal automorphism. More precisely, $Q(\frac{v}{2})$ is the infinitesimal automorphism induced by the derivation $\widetilde{v}$ introduced in* Equation (5.9). *In other words, we have $Q(\frac{v}{2}) = \widetilde{\varepsilon \widetilde{v}}$.*

Proof. The automorphism $Q(v)$ of TM is smooth over $T\mathbb{K}$ since it is defined in terms of $T\mu$, and it preserves fibers since for all $u \in TM$,

$$\pi(Q(v)u) = \pi(T\mu(v, T\mu(0_o, u))) = \mu(\pi(v), \mu(o, \pi(u))) = \mu(o, \mu(o, \pi(u))) = \pi(u).$$

Thus Theorem 6.8 implies that $Q(v)$ is an infinitesimal automorphism. It is induced by the vector field

$$p \mapsto Q(v)0_p = 2\widetilde{v}(p).$$

■

Using the lemma, we can give another proof (not depending on the preceding chapter) of the first two relations from the theorem: on TTM, the quadratic operator is given by taking the tangent map of the infinitesimal automorphism from the preceding lemma,

$$Q(\varepsilon_1 \frac{v}{2}) = \widetilde{\varepsilon_1 \widetilde{v}},$$

and similarly for ε_1 and ε_2 exchanged. Then Theorem 14.4 on the Lie bracket implies

$$[Q(\varepsilon_1 \frac{u}{2}), Q(\varepsilon_2 \frac{v}{2})](w) = [\widetilde{\varepsilon_1 \widetilde{u}}, \widetilde{\varepsilon_2 \widetilde{v}}](w) = \widetilde{\varepsilon_1 \varepsilon_2 [\widetilde{u}, \widetilde{v}]}(o) + w = 0_o + w = w$$

since $[\widetilde{u}, \widetilde{v}]$ vanishes at o (Prop. 5.9). In a similar way, again with Theorem 14.4, we get the third relation: the triple bracket $[[\widetilde{u}, \widetilde{v}], \widetilde{w}]$ is given by

$$\varepsilon_1 \varepsilon_2 \varepsilon_3 [[\widetilde{u}, \widetilde{v}], \widetilde{w}] = [[Q(\frac{\varepsilon_1 u}{2}), Q(\frac{\varepsilon_2 v}{2})], Q(\frac{\varepsilon_3 w}{2})]|_{\varepsilon_1 \varepsilon_2 \varepsilon_3 TM}.$$

Evaluating at the base point, using that the vector field $[\widetilde{u}, \widetilde{v}]$ vanishes at o (Prop. 5.9) and that $Q(\frac{\varepsilon_3 w}{2}).0 = \varepsilon_3 w$, we get

$$\begin{aligned} \varepsilon_1 \varepsilon_2 \varepsilon_3 [u, v, w] &= \varepsilon_1 \varepsilon_2 \varepsilon_3 [[\widetilde{u}, \widetilde{v}], \widetilde{w}].0 \\ &= \varepsilon_1 \varepsilon_2 [\widetilde{u}, \widetilde{v}] \cdot \varepsilon_3 w - \varepsilon_3 \widetilde{w}(0) = [Q(\frac{\varepsilon_1 u}{2}), Q(\frac{\varepsilon_2 v}{2})].\varepsilon_3 w - \varepsilon_3 w, \end{aligned}$$

which had to be shown. ∎

Another way of stating the preceding relations goes as follows: let us introduce (with respect to a fixed base point $o \in M$) the map

$$\begin{aligned} &\rho := \rho_M : M \times M \times M \to M, \\ &(x, y, z) \mapsto [Q(x), Q(y)]z = Q(x)Q(y)Q(x)^{-1}Q(y)^{-1}z = \sigma_x \sigma_o \sigma_y \sigma_x \sigma_o \sigma_y . z \end{aligned}$$

Then, writing $w = \sum_\alpha \varepsilon^\alpha w_\alpha$, etc., we have

$$(T^3 \rho)_{o,o,o}(u, v, w) = w + \varepsilon^{111} [u_{100}, v_{010}, w_{001}]. \tag{27.1}$$

This follows by noting that $T^3 \rho_M = \rho_{T^3 M}$.

27.4. *Fundamental commutation relations for elements in the axes of a symmetric space.* For elements $\varepsilon^\alpha u_\alpha$, $\varepsilon^\beta v_\beta$, $\varepsilon^\gamma w_\gamma$ of the axes of $(T^k M)_o$ with $k \geq 3$ we have the following "commutation relations":

$$[Q(\varepsilon^\alpha \frac{u_\alpha}{2}), Q(\varepsilon^\beta \frac{v_\beta}{2})] \varepsilon^\gamma w_\gamma = \begin{cases} \varepsilon^\gamma w_\gamma + \varepsilon^{\alpha+\beta+\gamma} [u_\alpha, v_\beta, w_\gamma] & \text{if } \alpha \perp \beta, \beta \perp \gamma, \alpha \perp \gamma, \\ \varepsilon^\gamma w_\gamma & \text{else,} \end{cases} \tag{27.2}$$

where the bracket on the left hand side denotes the commutator in the group $\mathrm{Diff}(T^k M)$. In fact, the general case is reduced to the case $k = 3$; if α, β, γ are not disjoint, then we may further reduce to the case $k = 2$ in which the fibers are abelian symmetric spaces (Theorem 26.3), whence the claim in this case; in the other case we apply Theorem 27.2.

27.5. *The derived multilinear connections on T^kM.* Starting with the canonical connection L on TM, we may define a sequence of derived multilinear connections on T^kM (Chapter 17). We will give an explicit formula for the derived linear structure on T^3M: let $L_1 := L$ be the canonical connection of the symmetric space M and $L_k := D^{k-1}L$ be the sequence of derived linear structures, and $\Phi_k : A^{k+1}M \to T^{k+1}M$ be the corresponding linearization map. The map $\Phi_1 : A^2M \to T^2M$ is given by (26.7). We are going to calculate $\Phi_2 : A^3M \to T^3M$. First of all, we have to calculate the tangent map $T\Phi_1$. It is a bundle map over the base space $TM = T_{\varepsilon_3}M$; a typical point in the base is $\varepsilon_3 v_3 \in T_xM$. In the formula for $T\Phi_1$, the Q-operators in T^2M are replaced by the corresponding ones in T^3M, but taken with respect to the base point $\varepsilon_3 v_3 \in T_xM$. The factor $\frac{1}{2}$ in (26.7) really stands for the map

$$(\frac{1}{2})_{TTM} : TTM \to TTM, \quad v \mapsto \frac{v}{2}. \tag{27.3}$$

When taking its derivative with respect to ε_3, the base point $\varepsilon_3 v_3$ is fixed. In the following, the sign $+$ without index stands for $+_x$ (addition in fiber of TTM over x), and Q without index stands for Q_{0_x}; the index $\varepsilon_3 v_3$ labels the corresponding operations in the fiber over $\varepsilon_3 v_3$. In the following calculation, we pass from one base point to another via the "isotopy formula"

$$Q_x(y) = s_y s_o s_o s_x = Q_o(y) Q_o(x)^{-1}. \tag{27.4}$$

Then we have

$$\begin{aligned}
&\Phi_2(x; \varepsilon_1 v_1, \varepsilon_2 v_2, \varepsilon_3 v_3, \varepsilon_1\varepsilon_2 v_{12}, \varepsilon_2\varepsilon_3 v_{13}, \varepsilon_2\varepsilon_3 v_{23}) \\
&= (\varepsilon_2 v_2 + \varepsilon_3 v_3 + \varepsilon_2\varepsilon_3 v_{23}) +_{\varepsilon_3 v_3} (\varepsilon_1 v_1 + \varepsilon_3 v_3 + \varepsilon_1\varepsilon_3 v_{13}) \\
&\qquad +_{\varepsilon_3 v_3} (\varepsilon_1\varepsilon_2 v_{12} + \varepsilon_3 v_3 + \varepsilon_1\varepsilon_2\varepsilon_3 v_{123}) \\
&= Q_{\varepsilon_3 v_3}(\varepsilon_2 \frac{v_2}{2} + \varepsilon_3 v_3 + \varepsilon_2\varepsilon_3 \frac{v_{23}}{2}) \circ Q_{\varepsilon_3 v_3}(\varepsilon_1 \frac{v_1}{2} + \varepsilon_3 v_3 + \varepsilon_1\varepsilon_3 \frac{v_{13}}{2}) \circ \\
&\qquad Q_{\varepsilon_3 v_3}(\varepsilon_1\varepsilon_2 \frac{v_{12}}{2} + \varepsilon_3 v_3 + \varepsilon_1\varepsilon_2\varepsilon_3 \frac{v_{123}}{2}).0_{\varepsilon_3 v_3} \\
&= Q_{\varepsilon_3 v_3}\Big(Q(\varepsilon_2 \frac{v_2}{4}) Q(\varepsilon_3 \frac{v_3}{2}) Q(\varepsilon_2\varepsilon_3 \frac{v_{23}}{4}).0_x\Big) \circ \\
&\qquad Q_{\varepsilon_3 v_3}\Big(Q(\varepsilon_1 \frac{v_1}{4}) Q(\varepsilon_3 \frac{v_3}{2}) Q(\varepsilon_1\varepsilon_3 \frac{v_{13}}{4}).0_x\Big) \circ \\
&\qquad Q_{\varepsilon_3 v_3}\Big(Q(\varepsilon_1\varepsilon_2 \frac{v_{12}}{4}) Q(\varepsilon_3 \frac{v_3}{2}) Q(\varepsilon_1\varepsilon_2\varepsilon_3 \frac{v_{123}}{4}).0_x\Big).0_{\varepsilon_3 v_3} \\
&= Q(\varepsilon_3 \frac{v_3}{2}) Q\Big(Q(\varepsilon_2 \frac{v_2}{4}) Q(\varepsilon_2\varepsilon_3 \frac{v_{23}}{4}).0_x\Big)\ Q\Big(Q(\varepsilon_1 \frac{v_1}{4}) Q(\varepsilon_1\varepsilon_3 \frac{v_{13}}{4}).0_x\Big) \\
&\qquad Q\Big(Q(\varepsilon_1\varepsilon_2 \frac{v_{12}}{4}) Q(\varepsilon_1\varepsilon_2\varepsilon_3 \frac{v_{123}}{4}).0_x\Big).0_x \\
&= Q(\varepsilon_3 \frac{v_3}{2}) Q(\varepsilon_2 \frac{v_2}{2}) Q(\varepsilon_2\varepsilon_3 \frac{v_{23}}{2}) Q(\varepsilon_1 \frac{v_1}{2}) Q(\varepsilon_1\varepsilon_3 \frac{v_{13}}{2}) Q(\varepsilon_1\varepsilon_2 \frac{v_{12}}{2}) Q(\varepsilon_1\varepsilon_2\varepsilon_3 \frac{v_{123}}{2}).0_x
\end{aligned}$$

where the simplifications in the last lines are due to the "fundamental formula" and the fact that Q-operators having an ε_i for argument in common commute. The formula for Φ_k with general k is obtained in a similar way; we state the result without proof: if, as in Theorem 24.8, κ denotes the order-reversing permutation

$\kappa(j) = k + 1 - j$, then

$$\Phi_k(x; (v_\alpha)_{\alpha>0}) = (\prod_{\alpha>0}^{\uparrow} Q(\varepsilon^{\kappa\cdot\alpha} \frac{v_{\kappa\cdot\alpha}}{2})).0_x \tag{27.5}$$

where $\prod_\alpha^{\uparrow} A_\alpha = A_{0\ldots01} \circ \ldots \circ A_{1\ldots1}$ is the composition of operators A_α in the order given by the lexicographic order of the index set.

Theorem 27.6. *The curvature of the canonical connection of a symmetric space agrees with its Lie triple product (up to a sign): for $u, v, w \in T_oM$,*

$$R_o(u, v)w = -[u, v, w].$$

Proof. We are going to use the characterization of the curvature tensor by Theorem 18.3. (See [Lo69] and [Be00, I.2.5] for different arguments in the real finite-dimensional case, rather using covariant derivatives.) Let τ be the transposition (23). The formula for the connection $\tau \cdot L_2$ and for its Dombrowski splitting Φ_s^τ is obtained from the formula for Φ_2 by exchanging $\varepsilon_2 \leftrightarrow \varepsilon_3$, $v_1 \leftrightarrow v_3$, $v_{12} \leftrightarrow v_{13}$. Using the commutation rules, one sees that as a result one gets for Φ_s^τ a formula of the same type as for Φ_s, but beginning with $A \circ B := Q(\varepsilon_2 \frac{v_2}{2}) Q(\varepsilon_2 \frac{v_3}{3})$ instead of $B \circ A = Q(\varepsilon_3 \frac{v_3}{2}) Q(\varepsilon_2 \frac{v_2}{2})$. Using the operator formula $AB = [A, B]BA$, where $[A, B]$ is described by Theorem 27.2 (the transposition denoted there by κ is now τ), we see that the difference $\Phi_s^\tau - \Phi_s$ (with respect to an arbitrary chart representation) belongs to the axis $\varepsilon_1\varepsilon_2\varepsilon_3 V$ and is given by

$$-[Q(\frac{\varepsilon_1 v_1}{2}), Q(\frac{\varepsilon_2 v_2}{2})].\varepsilon_3 v_3.$$

On the one hand, by Formula (18.7), this term is equal to $\varepsilon_1\varepsilon_2\varepsilon_3 R(v_1, v_2)v_3$; on the other hand, by Theorem 27.2, it is equal to $\varepsilon_1\varepsilon_2\varepsilon_3 [v_1, v_2, v_3]$, whence the claim. ■

27.7. *The structure of J^3M.* The preceding calculations permit to give an explicit description of the symmetric space $(J^3M)_o$ with respect to the symmetrized linear structure $\mathrm{Sym}(DL)$ from Corollary 20.3. We will see that this "third order model" is the first non-trivial one, and it contains information that is equivalent to the Lie triple system at o. In fact, one finds that the restrictions of $\mathrm{Sym}(DL)$ and of the derived linear structure DL (corresponding to the map Φ_2 calculated above) coincide on J^3M, and the result of the calculation is given by the following formula, where Φ_2 is considered as an identification,

$$\begin{aligned}\mu(\delta v_1 + \delta^{(2)} v_2 + \delta^{(3)} v_3, \delta w_1 + \delta^{(2)} w_2 + \delta^{(3)} w_3) = {} & \delta(2v_1 - w_1) + \\ & \delta^{(2)}(2v_2 - w_2) + \\ & \delta^{(3)}(2v_3 - w_3 - ([v_1, w_1, v_1] + [w_1, v_1, w_1])).\end{aligned}$$

In terms of the quadratic map, this reads

$$\begin{aligned}Q_o(\delta v_1 + \delta^{(2)} v_2 + \delta^{(3)} v_3).(\delta w_1 + \delta^{(2)} w_2 + \delta^{(3)} w_3) = {} & \delta(2v_1 + w_1) + \\ & \delta^{(2)}(2v_2 + w_2) + \\ & \delta^{(3)}(2v_3 + w_3 + ([v_1, w_1, v_1] - [w_1, v_1, w_1])).\end{aligned}$$

In particular,

$$Q(\delta v) = \tau_v + \delta^{(3)}([v,\cdot,v] - [\cdot,v,\cdot])$$

is ordinary translation by v with respect to the linear structure plus an extra term which is the sum of a linear and a homogeneous quadratic map. Note that the Lie triple system of $(J^3M, 0_o)$ is $(J^3\mathfrak{m})_0 = \delta\mathfrak{m} \oplus \delta^{(2)}\mathfrak{m} \oplus \delta^{(3)}\mathfrak{m}$ with product

$$[\delta v_1 + \delta^{(2)} v_2 + \delta^{(3)} v_3, \delta u_1 + \delta^{(2)} u_2 + \delta^{(3)} u_3, \delta w_1 + \delta^{(2)} w_2 + \delta^{(3)} w_3] =$$
$$(0,0,\delta^3[v_1,u_1,w_1]) = (0,0,6\,\delta^{(3)}[v_1,u_1,w_1])$$

where $\mathfrak{m}$ is the Lie triple system of (M,o). It is nilpotent, but in general non-abelian, and it contains (if 6 is invertible in $\mathbb{K}$) all information on the original Lts $\mathfrak{m}$. Thus the space $(J^3M)_o$ may be considered as a "faithful model" of the original symmetric space M in the sense that it contains all infinitesimal information on M.

27.8. *On exponential mappings.* We say that a symmetric space M *has an exponential map* if, for every $x \in M$ and $v \in T_xM$, there exists a unique homomorphism of symmetric spaces $\gamma_v : \mathbb{K} \to M$ such that $\gamma(0) = x$, $\gamma_v'(0) = v$, and such that the map $\mathrm{Exp}_x : T_xM \to M$, $v \mapsto \gamma_v(1)$ is smooth (cf. [BeNe05, Ch. 2]). It is known that every real finite-dimensional symmetric space has an exponential map, and moreover this exponential map agrees with the exponential map of the canonical connection (see [Lo69]). In our general setting, not every symmetric space will have an exponential map, but we will show in Chapter 30 that, if $\mathbb{K}$ is a field of characteristic zero, there is a canonical "exponential jet"

$$\mathrm{Exp}_x^{(k)} : T_0((T^kM)_x) \to (T^kM)_x$$

having the property that, if M has an exponential map, then $\mathrm{Exp}_x^{(k)} = T^k(\mathrm{Exp}_x)_0$.

VI. Diffeomorphism groups and the exponential jet

In this final part of the main text we bring together ideas from Part V (Lie theory) and Part IV (Higher order geometry) by considering the group $G = \mathrm{Diff}_{\mathbb{K}}(M)$ as a sort of Lie group and relating its higher order "tangent groups" $T^kG = \mathrm{Diff}_{T^k\mathbb{K}}(T^kM)$ to classical notions such as flows and geodesics.

28. Group structure on the space of sections of T^kM

28.1. *Groups of diffeomorphisms.* We denote by $\mathrm{Diff}_{T^k\mathbb{K}}(T^kM)$ the group of diffeomorphisms $f : T^kM \to T^kM$ which are smooth over the ring $T^k\mathbb{K}$. By Theorem 7.2, there is a natural injection

$$\mathrm{Diff}_{\mathbb{K}}(M) \to \mathrm{Diff}_{T^k\mathbb{K}}(T^kM), \quad f \mapsto T^kf.$$

28.2. *The space of sections of T^kM.* We denote by $\mathfrak{X}^k(M) := \Gamma(M, T^kM)$ the space of smooth sections of the bundle $T^kM \to M$. A chart $\varphi : U \to V$ induces a bundle chart $T^k\varphi : T^kU \to T^kV$, and with respect to such a chart, a section $X : M \to T^kM$ over U is represented by

$$X(x) = x + \sum_{\alpha > 0} \varepsilon^\alpha X_\alpha(x) \tag{28.1}$$

with (chart-dependent) vector fields $X_\alpha : U \to V$. The inclusions of axes $\iota_\alpha : TM \to T^kM$ induce inclusions

$$\mathfrak{X}(M) \to \mathfrak{X}^k(M), \quad X \mapsto \iota_\alpha \circ X = \varepsilon^\alpha X.$$

We call such sections *(purely) vectorial.*

Theorem 28.3.

(1) *Every section $X : M \to T^kM$ admits a unique extension to a diffeomorphism $\widetilde{X} : T^kM \to T^kM$ which is smooth over the ring $T^k\mathbb{K}$. Here, the term "extension" means that $\widetilde{X} \circ z = X$, where $z : M \to T^kM$ is the zero section. In a chart representation as above, this extension is given by*

$$\widetilde{X}(x + \sum_\alpha \varepsilon^\alpha v_\alpha) = x + \sum_\alpha \varepsilon^\alpha \Big(v_\alpha + X_\alpha(x) + \sum_{\ell=2}^{|\alpha|} \sum_{\Lambda \in \mathcal{P}_\ell(\alpha)} \sum_{j=1,\ldots,\ell} (d^{\ell-1} X_{\Lambda^j})(x)(v_{\Lambda^1}, \ldots, \widehat{v}_{\Lambda^j}, \ldots, v_{\Lambda^\ell}) \Big)$$

(where "$\widehat{v}$" means that the corresponding term is to be omitted). If X is a section of J^kM over M, then the diffeomorphism $\widetilde{X}$ preserves J^kM.

(2) *There is a natural group structure on* $\mathfrak{X}^k(M)$, *defined by the formula*

$$X \cdot Y := \widetilde{X} \circ Y.$$

In a chart representation, the group structure of $\mathfrak{X}^k(M)$ *is given by*

$$(X \cdot Y)(p) = p + \sum_{\alpha} \varepsilon^{\alpha} \Big(X_{\alpha}(p) + Y_{\alpha}(p) + \sum_{\ell=2}^{|\alpha|} \sum_{\Lambda \in \mathcal{P}_{\ell}(\alpha)} \sum_{j=1}^{\ell} (d^{\ell-1} X_{\Lambda^j})(p) \big(Y_{\Lambda^1}(p), \ldots, \widehat{Y_{\Lambda^j}(p)}, \ldots, Y_{\Lambda^{\ell}}(p) \big) \Big).$$

The sections of $J^k M$ *form a subgroup of* $\mathfrak{X}^k(M)$, *and the imbeddings* $\mathfrak{X}(M) \to \mathfrak{X}^k(M)$, $X \mapsto \varepsilon^{\alpha} X$ *are group homomorphisms.*

(3) *The map* $X \mapsto \widetilde{X}$ *is an injective group homomorphism from* $\mathfrak{X}^k(M)$ *into* $\mathrm{Diff}_{T^k\mathbb{K}}(T^k M)$. *Elements of* $\mathrm{Diff}_{T^k\mathbb{K}}(T^k M)$ *permute fibers over* M, *and hence there is a well-defined projection*

$$\mathrm{Diff}_{T^k\mathbb{K}}(T^k M) \to \mathrm{Diff}_{\mathbb{K}}(M), \quad F \mapsto f := \pi \circ F \circ z.$$

The injection and projection thus defined give rise to a splitting exact sequence of groups

$$0 \;\to\; \mathfrak{X}^k(M) \;\to\; \mathrm{Diff}_{T^k\mathbb{K}}(T^k M) \;\to\; \mathrm{Diff}_{\mathbb{K}}(M) \;\to\; 1,$$

and hence $\mathrm{Diff}_{T^k\mathbb{K}}(T^k M)$ *is a semidirect product of* $\mathrm{Diff}_{\mathbb{K}}(M)$ *and* $\mathfrak{X}^k(M)$.

Proof. (1) We prove uniqueness and existence of the extension of $X : M \to T^k M$ by using a chart representation $X : U \to T^k V$ of X. The statement to be proved is:

(A) *Let* V *and* W *be* $\mathbb{K}$*-modules and* $f : U \to T^k W$ *a smooth map (over* $\mathbb{K}$*) defined on an open set* $U \subset V$. *Then there exists a unique extension of* f *to a map* $\widetilde{f} : T^k U \to T^k W$ *which is smooth over the ring* $T^k\mathbb{K}$.

We first prove uniqueness. In fact, here the arguments from the proof of Theorem 7.5 apply word by word – there the uniqueness statement was proved under the assumption that $f(U) \subset W$ (leading to the uniqueness statement in Theorem 7.6). But all arguments (i.e., essentially, the second order Taylor expansion) work in the same way for a $T^k\mathbb{K}$-smooth map taking values in $T^k W$, not only in W.

Next we prove the existence part of (A). By Theorem 7.2, $T^k f : T^k V \to T^k(T^k W)$ is smooth over the ring $T^k\mathbb{K}$. We will construct a canonical smooth map $\mu_k : T^k(T^k W) \to T^k W$ such that $\widetilde{f} := \mu_k \circ T^k f$ has the desired properties. Recall that $T^k\mathbb{K}$ is free over $\mathbb{K}$ with basis ε^{α}, $\alpha \in I_k$. The algebraic tensor product $T^k\mathbb{K} \otimes_{\mathbb{K}} T^{\ell}\mathbb{K}$ is again a commutative algebra over $\mathbb{K}$, isomorphic to $T^{k+\ell}\mathbb{K}$ after choice of a partition $\mathbb{N}_{k+\ell} = \mathbb{N}_k \dot\cup \mathbb{N}_{\ell}$. For any commutative algebra A, the product map $A \otimes A \to A$ is an algebra homomorphism, so we have, since $T^k\mathbb{K}$ is commutative, a homomorphism of $\mathbb{K}$-algebras

$$T^k m : T^{2k}\mathbb{K} \cong T^k\mathbb{K} \otimes_{\mathbb{K}} T^k\mathbb{K} \to T^k\mathbb{K}.$$

Explicitly, with respect to the $\mathbb{K}$-basis $\varepsilon^\alpha \otimes \nu^\beta$, $\alpha, \beta \in I_k$ (where the ν_i obey the same relations as the ε_i), it is given by

$$T^k m(\sum_{\alpha,\beta \in I_k} t_{\alpha,\beta}\, \varepsilon^\alpha \otimes \nu^\beta) = \sum_{\gamma \in I_k} \varepsilon^\gamma \sum_{\substack{\alpha,\beta \in I_k \\ \alpha+\beta=\gamma}} t_{\alpha,\beta}.$$

For any $\mathbb{K}$-module V, the ring homomorphism $T^k m : T^{2k}\mathbb{K} \to T^k\mathbb{K}$ induces a homomorphism of modules

$$\mu_k : T^{2k}V = V \otimes_{\mathbb{K}} T^{2k}\mathbb{K} \to T^k V = V \otimes_{\mathbb{K}} T^k\mathbb{K},$$
$$x + \sum_{\substack{\alpha,\beta \in I_k \\ (\alpha,\beta)\neq 0}} \varepsilon^\alpha \nu^\beta v_{\alpha,\beta} \mapsto x + \sum_{\gamma \in I_k} \varepsilon^\gamma \sum_{\substack{\alpha,\beta \in I_k \\ \alpha+\beta=\gamma}} v_{\alpha,\beta}.$$

This map is $T^{2k}\mathbb{K}$-linear (recall that any ring homomorphism $R \to S$ induces an R-linear map $V_R \to V_S$ of scalar extensions; here $R = T^{2k}\mathbb{K}$, $S = T^k\mathbb{K}$). It is of class C^0 since it is a finite composition of certain sums and projections which are all C^0, and hence μ_k is of class C^∞ over $T^{2k}\mathbb{K}$. There are two imbeddings of $T^k\mathbb{K}$ as a subring of $T^{2k}\mathbb{K} \cong T^k\mathbb{K} \otimes T^k\mathbb{K}$ having image $1 \otimes T^k\mathbb{K}$, resp. $T^k\mathbb{K} \otimes 1$; thus μ_k is also smooth over these rings. Therefore, if we let

$$\widetilde{f} := \mu_k \circ T^k f : T^k U \to T^{2k} W \to T^k W,$$

$\widetilde{f}$ is smooth over $T^k\mathbb{K}$ (which we may identify with $T^k\mathbb{K} \otimes 1 \subset T^{2k}\mathbb{K}$). Let us prove that it is an extension: for all $x \in U$,

$$\widetilde{f}(x) = \mu_k(T^k f(x)) = \mu_k(f(x)) = f(x)$$

where the second equality holds since $T^k f(x) = T^k f(x+0) = f(x)$ and the third equality holds since $\mu_k(x + \sum_\alpha \varepsilon^\alpha v_{\alpha,0}) = x + \sum_\alpha \varepsilon^\alpha v_{\alpha,0}$, i.e. $\mu_k \circ \iota = \mathrm{id}_{T^k M}$, where $\iota : T^k V \to T^{2k} V$ is the imbedding induced by $T^k\mathbb{K} \to T^{2k}\mathbb{K}$, $r \mapsto r \otimes 1$. Thus claim (A) is completely proved.

Now we prove part (1) of the theorem. The extension $\widetilde{X}$ of X is defined, locally, by using (A) in a chart representation; by the uniqueness statement of (A), this extension does not depend on the chart and hence is uniquely and globally defined. (See also Remark 28.4 for variants of this proof.) The chart formula given in the claim is simply Formula (28.2), where $f = X = \mathrm{pr}_1 + \sum_\alpha \varepsilon^\alpha X_\alpha$ is decomposed into its (chart-dependent) components according to Formula (28.1). All that remains to be proved is that $\widetilde{X}$ is bijective with smooth inverse. This will be done in the context of the proof of part (2).

(2) Clearly, $X \cdot Y$ is again a smooth section of $T^k M$, and hence the product is well-defined. The uniqueness part of (1) shows that $\widetilde{X \cdot Y} = \widetilde{X} \circ \widetilde{Y}$ since both sides are $T^k\mathbb{K}$-smooth extensions of $\widetilde{X} \circ Y$. This property implies associativity:

$$(X \cdot Y) \cdot Z = \widetilde{X \cdot Y} \circ Z = \widetilde{X} \circ \widetilde{Y} \circ Z = \widetilde{X} \circ (Y \cdot Z) = X \cdot (Y \cdot Z).$$

Moreover, for the zero section $z : M \to T^k M$ we have $\widetilde{z} = \mathrm{id}_{T^k M}$ and hence z is a neutral element.

All that remains to be proved is that X has an inverse in $\mathfrak{X}^k(M)$. Equivalently, we have to show that $\widetilde{X}$ is bijective and has an inverse of the form $\widetilde{Z}$ for some $Z \in \mathfrak{X}^k(M)$. This is proved by induction on k. In case $k = 1$, $\widetilde{X}$ is fiberwise translation in tangent spaces and hence is bijective with $(\widetilde{X})^{-1} = \widetilde{-X}$. For the case $k = 2$, see the explicit calculation of $(\widetilde{X})^{-1}$ in the proof of Theorem 14.3. For the proof of the induction step for general k, we give two slightly different versions:

(a) First version. Note that the chart formula for $X \cdot Y$ from the claim is a simple consequence of the chart formula for $\widetilde{X}(v)$ from Part (1) (take $v = Y(x)$). This formula shows that left multiplication by X, $l_X : \mathfrak{X}^k(M) \to \mathfrak{X}^k(M)$, $Y \mapsto X \cdot Y$, is an affine-multilinear self-map of the multilinear space $\mathfrak{X}^k(M)$ in the sense explained in Section MA.19. Moreover, l_X is unipotent, hence regular, and by the affine-multilinear version of Theorem MA.6 (cf. MA.19) it is thus bijective. Now let $Z := (l_X)^{-1}(0)$ where $0 := z$ is the zero section. Then $X \cdot Z = l_X \circ (l_X)^{-1}(0) = 0$. Similarly, $Y \cdot X = 0$, where $Y := (r_X)^{-1}(0)$ is gotten from the bijective right multiplication by X. It follows that X is invertible and $X^{-1} = Z = Y$.

(b) Second version. Note first that, if $\varepsilon^\alpha X$, $\varepsilon^\alpha Y$ with $X, Y \in \mathfrak{X}(M)$ are purely vectorial, then $\varepsilon^\alpha X \cdot \varepsilon^\alpha Y = \varepsilon^\alpha (X + Y)$, as follows from the explicit formula in Part (2). Thus $\varepsilon^\alpha X$ is invertible with inverse $\varepsilon^\alpha(-X)$. One proves by induction that every element $X \in \mathfrak{X}^k(M)$ can be written as a product of purely vectorial sections (Theorem 29.2): $X = \prod_\alpha^\uparrow \varepsilon^\alpha X_\alpha$, and hence X is invertible.

(3) We have already remarked that $X \mapsto \widetilde{X}$ is a group homomorphism, and injectivity follows from the extension property. Next we show that the projection is well-defined. More generally, we show that a $T^k\mathbb{K}$-smooth map $F : T^k M \to T^k N$ maps fibers to fibers. In fact, letting $f := F \circ z : M \to T^k N$, we are in the situation of Statement (A) proved above. It follows that $F = \widetilde{f}$ is, over a chart domain, given by the formula from Theorem 7.5, which clearly implies that $F(x + \sum_{\alpha > 0} \varepsilon^\alpha v_\alpha) \in F(x) + \sum_{\alpha > 0} \varepsilon^\alpha W$, and hence F maps fibers to fibers. In particular, the projection $\mathrm{Diff}_{T^k\mathbb{K}}(T^k M) \to \mathrm{Diff}_{\mathbb{K}}(M)$ is well-defined. Let us prove that the imbedded image $\mathfrak{X}^k(M)$ is the kernel of the natural projection $\mathrm{Diff}_{T^k\mathbb{K}}(T^k M) \to \mathrm{Diff}_{\mathbb{K}}(M)$, $F \mapsto f$. In fact, $X := F \circ z : M \to T^k M$ is a section of $T^k M$ iff $f = \mathrm{id}_M$. But then, by the uniqueness statement in Part (1), $F = \widetilde{X}$; conversely, for any $X \in \mathfrak{X}^k(M)$, $\widetilde{X}$ clearly belongs to the kernel of the projection. Finally, it is clear that $f \mapsto T^k f$ is a splitting of the projection $\mathrm{Diff}_{T^k\mathbb{K}}(T^k M) \to \mathrm{Diff}_{\mathbb{K}}(M)$. ∎

28.4. *Comments on Theorem 28.3 and its proof.* For the real finite-dimensional case of Theorem 28.3 (2), see [KM87, Theorem 4.6] and [KMS93, Theorem 37.7], where the proof is carried out in the "dual" picture, corresponding to function-algebras. The proof given there has the advantage that the problem of bijectivity (i.e., the existence of inverses) is easily settled by using the Neumann series in a nilpotent associative algebra, but it does not carry over to our general situation. Moreover, in a purely real theory it is less visible why the extension of a section X to a diffeomorphism $\widetilde{X}$ is so natural.

Comparing with the proof of [KMS93, Theorem 37.7], the proof of the existence part in (1) may be reformulated as follows. We define $\widetilde{X} = \mu_{2k,k} \circ T^k X$ where $\mu_{2k,k} : T^{2k} M \to T^k M$ is the natural map which, in a chart representation, is given by $\mu_k : T^k(T^k V) \to T^k V$. Note that, for $k = 1$, $\mu_{2,1} : TTM \to TM$ can directly be defined by

$$\mu_{2,1} = a \circ (p_1 \times_M p_2) : TTM \to TM \times_M TM \to TM,$$

where $p_i : TTM \to TM$, $i = 1, 2$, are the canonical projections and $a : TM \times_M TM \to TM$ the addition map of the vector bundle $TM \to M$. In a chart,

$$\mu_{2,1}(x + \varepsilon_1 v_1 + \varepsilon_2 v_2 + \varepsilon_1 \varepsilon_2 v_{12}) = x + \varepsilon(v_1 + v_2)$$

which corresponds to the formula for $\mu_1 : TTV \to TV$.

29. The exponential jet for vector fields

29.1. *The "Lie algebra of $\mathfrak{X}^k(M)$".* As explained in the introduction, heuristically we may think of $G := \mathrm{Diff}_{\mathbb{K}}(M)$ as a Lie group with Lie algebra $\mathfrak{g} = \mathfrak{X}(M)$. Then the group $\mathrm{Diff}_{T^k\mathbb{K}}(T^kM)$ takes the rôle of T^kG and the group $\mathfrak{X}^k(M)$ the one of $(T^kG)_e$. The Lie algebra of T^kG is then $T^k\mathfrak{g} = \mathfrak{g} \otimes_{\mathbb{K}} T^k\mathbb{K}$ and therefore the one of $\mathfrak{X}^k(M)$ has to be $(T^k\mathfrak{g})_0 = \oplus_{\alpha>0}\varepsilon^\alpha\mathfrak{g}$. It is the space of sections of the *axes-bundle* $A^kM = \bigoplus_{\alpha\in I_k,\alpha>0}\varepsilon^\alpha TM$. The space of sections of A^kM will be denoted by

$$\mathcal{A}^k(M) := \Gamma(M, A^kM) = \bigoplus_{\alpha\in I_k,\alpha>0} \varepsilon^\alpha \mathfrak{X}(M). \tag{29.1}$$

We will construct canonical bijections $\mathcal{A}^k(M) \to \mathfrak{X}^k(M)$, taking the rôle of left-trivialization, resp. of an exponential map.

Theorem 29.2. *There is a bijective "left trivialization" of the group $\mathfrak{X}^k(M)$, given by*

$$\mathcal{A}^k(M) \to \mathfrak{X}^k(M), \quad (X_\alpha)_\alpha \mapsto \prod^{\uparrow}_{\substack{\alpha\in I_k\\ \alpha>0}} \varepsilon^\alpha X_\alpha,$$

where the product is taken in the group $\mathfrak{X}^k(M)$, with order corresponding to the lexicographic order of the index set I_k. Then the product in $\mathfrak{X}^k(M)$ is given by

$$\prod^{\uparrow}_{\alpha} \varepsilon^\alpha X_\alpha \cdot \prod^{\uparrow}_{\beta} \varepsilon^\beta Y_\beta = \prod^{\uparrow}_{\gamma} \varepsilon^\gamma Z_\gamma$$

with Z_γ given by the "left product formula" from Theorem 24.7.

Proof. First of all, note that Theorem 14.4 implies the following commutation relations for purely vectorial elements in the group $\mathfrak{X}^k(M)$: for all vector fields X_α, $Y_\beta \in \mathfrak{X}(M)$,

$$[\varepsilon^\alpha X_\alpha, \varepsilon^\beta Y_\beta] = \varepsilon^{\alpha+\beta}[X_\alpha, Y_\beta], \tag{29.2}$$

where the last bracket is taken in the Lie algebra $\mathfrak{X}(M)$. Now we define the map $\Psi : \mathcal{A}^k(M) \to \mathfrak{X}^k(M)$, $(X_\alpha)_\alpha \mapsto \prod^{\uparrow}_{\substack{\alpha\in I_k\\ \alpha>0}} \varepsilon^\alpha X_\alpha$ as in the claim and show that it is a bijection. We claim that the image of Ψ is a subgroup of $\mathfrak{X}^k(M)$. In fact, this follows from the relations (29.2): re-ordering the product of two ordered products is carried out by the procedure described in the proof of Theorem 24.7 and leads to the left-product formula mentioned in the claim. Thus the image of Ψ is the subgroup of $\mathfrak{X}^k(M)$ generated by the purely vectorial sections, with group structure as in the claim. But this group is in fact the whole group $\mathfrak{X}^k(M)$: using the explicit chart formula for the product in $\mathfrak{X}^k(M)$ (Theorem 28.3 (2)), it is seen that Ψ is a unipotent multilinear map between multilinear spaces over $\mathbb{K}$, whence is bijective (Theorem MA.6). ∎

In the same way as in the theorem, a "right-trivialization map" $\mathcal{A}^k(M) \to \mathfrak{X}^k(M)$ is defined. Next we prove the analog of Theorem 25.2 on the exponential map:

Theorem 29.3. *Assume that $\mathbb{K}$ is a non-discrete topological field of characteristic zero and M is a manifold over $\mathbb{K}$. Then there is a unique map $\exp_k : \mathcal{A}^k(M) \to \mathfrak{X}^k(M)$ such that:*

(1) *In every chart representation, $\exp_k$ is a polynomial map (of degree at most k); equivalently, $\exp_k^L := \Psi^{-1} \circ \exp_k : \mathcal{A}^k(M) \to \mathcal{A}^k(M)$, with Ψ as in the preceding theorem, is a polynomial map,*

(2) *for all $n \in \mathbb{Z}$ and $X \in \Gamma(M, A^k M)$, $\exp(nX) = (\exp X)^n$,*

(3) *on the axes, $\exp_k$ agrees with the inclusion maps, i.e., if $X : M \to TM$ is a vector field, then $\exp(\varepsilon^\alpha X) = \varepsilon^\alpha X$ is the purely vectorial section corresponding to $\varepsilon^\alpha X$.*

Proof. As for Theorem 25.2, this in an immediate consequence of Theorem PG.6, applied to the polynomial group $\mathfrak{X}^k(M)$. ■

Theorem 29.4. *The map $\exp_k$ from the preceding theorem commutes with the canonical action of the symmetric group Σ_k and hence restricts to a bijection*

$$\exp_k : (\mathcal{A}^k(M))^{\Sigma_k} = \oplus_{j=1}^k \delta^{(j)} \mathfrak{X}(M) = \Gamma(\oplus_{j=1}^k \delta^{(j)} TM) \to (\mathfrak{X}^k(M))^{\Sigma_k} = \Gamma(J^k M).$$

In particular, we get an injection

$$\mathfrak{X}(M) \to G^k(M), \quad X \mapsto \widetilde{\exp_k(\delta X)}$$

such that $\widetilde{\exp_k(\delta n X)} = (\widetilde{\exp_k(\delta X)})^n$ for all $n \in \mathbb{Z}$.

Proof. This is the analog of Theorem 25.4, and it is proved in the same way. ■

29.5. *Smooth actions of Lie groups.* Assume G is a Lie group over $\mathbb{K}$, acting smoothly on M via $a : G \times M \to M$. Deriving, we get an action $T^k a : T^k G \times T^k M \to T^k M$ of $T^k G$ on $T^k M$. Then $T^k a(g, \cdot)$ is smooth over $T^k \mathbb{K}$ and hence belongs to the group $\mathrm{Diff}_{T^k \mathbb{K}}(T^k M)$. Therefore we may consider $T^k a$ and $(T^k a)_e$ as group homomorphisms

$$T^k a : T^k G \to \mathrm{Diff}_{T^k \mathbb{K}}(T^k M), \quad \text{resp.} \quad (T^k a)_e : (T^k G)_e \to \mathfrak{X}^k(M).$$

In particular, for $k = 1$ we get a homomorphism of abelian groups $\mathfrak{g} = T_e G \to \mathfrak{X}(M)$, associating to $X \in \mathfrak{g}$ the "induced vector field on M", classically often denoted by $X_* \in \mathfrak{X}(M)$. This map is a Lie algebra homomorphism since $T^2 a$ is a group homomorphism and the Lie bracket on $\mathfrak{g}$, resp. on $\mathfrak{X}(M)$, can be defined in similar ways via group commutators (Theorem 14.5 and Theorem 23.2). Taking a direct sum of such maps, we get a Lie algebra homomorphism $(T^k \mathfrak{g})_0 \to \mathcal{A}^k(M)$, and from the uniqueness property of the exponential map we see that the following diagram commutes:

$$\begin{array}{ccc} (T^k G)_e & \overset{T^k a}{\to} & \mathfrak{X}^k(M) \\ \exp \uparrow & & \uparrow \exp_k \\ (T^k \mathfrak{g})_0 & \to & \mathcal{A}^k(M) \end{array}$$

Summing up, everything behaves as if $a : G \to \mathrm{Diff}(M)$ were a honest Lie group homomorphism.

29.6. *Relation with the flow of a vector field.* The diffeomorphism $\widetilde{\exp_k(\delta X)}$ of $T^k M$ (Theorem 29.4) can be related to the k-jet of the flow of X, if flows exist. Let us explain this. We assume that $\mathbb{K} = \mathbb{R}$ and M is finite-dimensional (or a Banach manifold). Then any vector field $X \in \mathfrak{X}(M)$ admits a flow map $(t,x) \mapsto \mathrm{Fl}_t^X(x)$, defined on some neighborhood of $\{0\} \times M$ in $\mathbb{R} \times M$ (cf. [La99, Ch. IV]). In general, Fl_1^X is not defined as a diffeomorphism on the whole of M, and this obstacle prevents us from defining the exponential map of the group $G = \mathrm{Diff}(M)$ by $\exp(X) := \mathrm{Fl}_1^X$. Nevertheless, this is what one would like to do, and one may give some sense to this definition by fixing some point $p \in M$ and by considering the subalgebra

$$\mathfrak{X}(M)_p := \{X \in \mathfrak{X}(M) |\, X(p) = 0, \forall Y \in \mathfrak{X}(M) : [X,Y](p) = 0\}$$

of vector fields vanishing at p of order at least 2. Then it is easily seen from the proof of the existence and uniqueness theorem for local flows that, for all $X \in \mathfrak{X}(M)_p$, $\exp(X) := \mathrm{Fl}_1^X$ is defined on some open neighborhood of p, and moreover that $T_p(\exp(X)) = \mathrm{id}_{T_p M}$ (cf., e.g., [Be96, Appendix (A2)]). Thus the group $\mathrm{Diff}(M)_p$ of germs of local diffeomorphisms that are defined at p and have trivial first order jet there, admits an exponential map given by $\exp = \mathrm{Fl}_1 : \mathfrak{X}(M)_p \to \mathrm{Diff}(M)_p$. We claim that, under these assumptions, the k-jet of Fl_1^X at p is given by the map from Theorem 29.4:

$$(\widetilde{\exp_k(\delta X)})_p = (T^k(\mathrm{Fl}_1^X))_p,$$

or, in other words, that the following diagram commutes:

$$\begin{array}{ccccc}
\mathrm{Diff}(M)_p & \overset{T^k}{\to} & G^k(M) & \overset{\mathrm{ev}_p}{\to} & \mathrm{Am}^{k,1}(T^k M)_p \\
\mathrm{Fl}_1 \uparrow & & & & \uparrow \mathrm{ev}_p \\
\mathfrak{X}(M)_p & \overset{X \mapsto \delta X}{\to} & \mathcal{A}^k(M) & \overset{\exp_k}{\to} & G^k(M)
\end{array}$$

where $\mathrm{Am}^{k,1}(T^k M)_p$ denotes the affine-multilinear group of the fiber $(T^k M)_p$. The proof is by using the uniqueness of the map exp from Theorem PG.6: we contend that the map $X \mapsto (T^k(\mathrm{Fl}_1^X))_p$ also has the properties (1) and (2) from Theorem PG.6 and hence coincides with the exponential map constructed above by using Theorem PG.6. Now, Property (1) follows by observing that

$$T^k(\mathrm{Fl}_1^{nX}) = T^k((\mathrm{Fl}_1^X)^n) = (T^k(\mathrm{Fl}_1^X))^n,$$

and Property (2) follows from the fact that $T_p(\mathrm{Fl}_1^X) = \mathrm{id}_{T_p M}$ (here we need the fact that X vanishes at p of order at least two, cf. [Be96, Appendix (A.2)]) and hence our candidate for the exponential mapping acts trivially on axes, as it should. Summing up, for any $X \in \mathfrak{X}(M)_p$ the k-jet of $\exp(X)$ at p may be constructed in a purely "synthetic way", i.e., without using the integration of differential equations. If X does not belong to $\mathfrak{X}(M)_p$, then there still is a relation, though less canonical, with the k-jet of the flow of X, see the next subsection.

29.7. *Relation with "Chapoton's formula".* If X is a vector field on $V = \mathbb{R}^n$, identified with the corresponding map $V \to V$, then the local flow of X can be written

$$\mathrm{Fl}_t^X(v) = v + \sum_{j \geq 1} \frac{t^j}{j!} X^{\star j}(v)$$

where

$$(X \star Y)(x) = (\nabla_X Y)(x) = dY(x) \cdot X(x)$$

denotes the product on $\mathfrak{X}(V)$ given by the canonical flat connection of V, and $X^{\star j+1} = X \star X^{\star j}$ are the powers with respect to this product – see [Ch01, Prop. 4]. Thus, letting $t = 1$, the exponential map associated to the Lie algebra of vector fields on V can be written as a formal series

$$\exp(X) = \mathrm{id}_V + \sum_{j \geq 1} \frac{t^j}{j!} X^{\star j}.$$

This formal series may be seen as the projective limit of our $\exp_k(\delta X) : V \to J^k V$ for $k \to \infty$ (cf. Chapter 32). In fact, according to Theorem PG.6, we have the explicit expression of the exponential map of the polynomial group $\mathfrak{X}^k(V)$ by $\exp(\delta X) = \sum_{j=1}^k \frac{1}{j!} \psi_{j,j}(\delta X)$ where $\psi_{j,j}(X)$ is gotten from the homogeneous term of degree j in the iterated multiplication map $m^{(j)}$. The chart formula in the global chart V for the product (Theorem 28.3 (2)) permits to calculate these terms explicitly, leading to the formula

$$\exp_k(\delta X) = \mathrm{id}_V + \sum_{j \geq 1} \frac{\delta^{(j)}}{j!} X^{\star j}.$$

A variant of these arguments is implicitly contained in [KMS93, Prop. 13.2] where the exponential mapping of the "jet groups" is determined: for a vector field X vanishing at the origin, $\exp(j_0^k X) = j_0^k \mathrm{Fl}_1^X$.

30. The exponential jet of a symmetric space

30.1. As an application of the preceding results, we will construct a canonical isomorphism $\mathrm{Exp}_k : A^kM \to T^kM$ (called the *exponential jet*, cf. Introduction) for symmetric spaces over $\mathbb{K}$. In order to motivate our construction, we start by considering the group case, where such a construction has already been given in Chapter 25 (cf. also Section 29.5). We relate those results to the context explained in the preceding chapters.

30.2. *Case of a Lie group.* We continue to assume that $\mathbb{K}$ is a non-discrete topological field of characteristic zero. We construct the exponential jet $\mathrm{Exp}_k : A^kG \to T^kG$ as follows: for $X \in \mathfrak{g}$, let $X^R \in \mathfrak{X}(G)$ be the right-invariant vector field such that $X^R(e) = X$. The "extension map"

$$l_R : \mathfrak{g} \to \mathfrak{X}(G), \quad X \mapsto l_R(X) := X^R$$

is injective. Denote by

$$(T^k l_R)_0 : (T^k\mathfrak{g})_0 \to (T^k\mathfrak{X}(G))_0 = \mathcal{A}^k(G)$$

its scalar extension (simply a direct sum of copies of l_R). Then the exponential map $\mathrm{Exp} = \mathrm{Exp}_k : (T^k\mathfrak{g})_0 \to (T^kG)_e$ of the polynomial group $(T^kG)_e$ (Theorem 25.2) is recovered from these data via

$$\mathrm{Exp}_k(X) = (\exp((T^k l_R)X))(e) : \qquad \begin{array}{ccc} (T^kG)_e & \stackrel{\mathrm{ev}_e}{\leftarrow} & \mathfrak{X}^k(G) \\ \mathrm{Exp}\uparrow & & \uparrow \exp_k \\ (T^k\mathfrak{g})_0 & \stackrel{(T^k l_R)_0}{\rightarrow} & \mathcal{A}^k(G) \end{array}$$

In fact, this is seen by the arguments given in Section 29.5, applied to the action of G on itself by left translations.

In the case of finite dimension over $\mathbb{K} = \mathbb{R}$ (or for general real Banach Lie-groups), the group G does admit an exponential map $\mathrm{Exp} : \mathfrak{g} \to G$ in the sense explained in Section 25.9, and it is well-known that then the exponential map of G is given by $\exp(X) = \mathrm{Fl}_1^{X^R}(e)$, i.e., by the commutative diagram

$$\begin{array}{ccc} G & \stackrel{\mathrm{ev}_e}{\leftarrow} & \mathrm{Diff}(G) \\ \mathrm{Exp}\uparrow & & \uparrow \mathrm{Fl}_1 \\ \mathfrak{g} & \stackrel{l_R}{\rightarrow} & \mathfrak{X}(G)_c \end{array}$$

(where $\mathfrak{X}(G)_c$ is the space of complete vector fields on G). The preceding definition is obtained by applying, formally (i.e., as if $\mathrm{Diff}(G)$ were a Lie group), the functor $(T^k)_0$ to this diagram.

30.3. *Exponential jet of a symmetric space.* Assume M is a symmetric space over $\mathbb{K}$ with base point $o \in M$. Recall from the proof of Prop. 5.9 that every tangent vector $v \in T_oM$ admits a unique extension to a vector field $l(v)$ (denoted by $\widetilde{v}$ in Chapter 5) such that $l(v)$ is a derivation of M and $\sigma_o \cdot l(v) = -l(v)$. The "extension map" $l := l_o : T_oM \to \mathfrak{X}(M)$ is linear, and its higher order tangent map $(T^k l)_0 : (A^kM)_o \to \mathcal{A}^k(M)$ is simply a direct sum $\oplus_{\alpha>0}\varepsilon^\alpha l$ of copies of this map.

Theorem 30.4. *Let (M, o) be a symmetric space over $\mathbb{K}$ and assume that $\mathbb{K}$ is non-discrete topological field of characteristic zero. Then the map* $\mathrm{Exp} := \mathrm{Exp}_k : (A^k M)_o \to (T^k M)_o$ *defined by*

$$\mathrm{Exp}_k(v) = (\exp_k(T^k l(v)))(o) : \qquad \begin{array}{ccc} (T^k M)_o & \overset{\mathrm{ev}_o}{\leftarrow} & \mathfrak{X}^k(M) \\ \mathrm{Exp} \uparrow & & \uparrow \exp_k \\ (A^k M)_o & \overset{(T^k l)_0}{\rightarrow} & \mathcal{A}^k(M) \end{array}$$

is a unipotent multilinear isomorphism of fibers over $o \in M$ commuting with the action of the symmetric group Σ_k. Moreover, Exp_k *satisfies the "one-parameter subspace property" from* Section 27.8*: for all $u \in (A^k M)_o$, the map*

$$\gamma_u : \mathbb{K} \to (T^k M)_o, \quad t \mapsto \mathrm{Exp}_k(tu)$$

is a homomorphism of symmetric spaces such that $\gamma_u'(0) = u$. In particular, the property $\mathrm{Exp}(nu) = (\mathrm{Exp}(u))^n$ *holds for all $n \in \mathbb{Z}$ (where the powers in the symmetric space $(T^k M)_o$ are defined as in* Equation (5.5)*).*

Proof. By its definition, Exp_k is a composition of multilinear maps and hence is itself multilinear. In order to prove that Exp_k is bijective, it suffices, according to Theorem MA.6, to show that the restriction of Exp_k to axes is the identity map. But this is an immediate consequence of the corresponding property of the map $\exp_k$ (Theorem 29.3, (3)): let $v = \varepsilon^\alpha v_\alpha$ be an element of the axis $\varepsilon^\alpha V$; then $X := T^k l(\varepsilon^\alpha v_\alpha) = \varepsilon^\alpha l(v_\alpha)$ is a purely vectorial section, hence $\exp_k$ applied to this section is simply given by inclusion of axes. Evaluating at the base point, we get

$$(\exp_k(\varepsilon^\alpha l(v_\alpha)))(o) = \varepsilon^\alpha l(v_\alpha)(o) = \varepsilon^\alpha v_\alpha$$

since $\mathrm{ev}_o \circ l = \mathrm{id}$. The map Exp commutes with the Σ_k-action since so do all maps of which it is composed.

Finally, we prove the one-parameter subspace property. In a first step, we show that, for all $n \in \mathbb{Z}$, $\mathrm{Exp}(nu) = \mathrm{Exp}(u)^n$. We write $u = x + \sum_\alpha \varepsilon^\alpha v_\alpha$; then, using that $\exp_k(nX) = (\exp_k(X))^n$,

$$\begin{aligned} \mathrm{Exp}(nu) &= \exp_k(n \sum_\alpha \varepsilon^\alpha l(v_\alpha)).o = (\exp(\sum_\alpha \varepsilon^\alpha l(v_\alpha)))^n.o \\ &= (\exp(\sum_\alpha \varepsilon^\alpha l(v_\alpha)).o)^n = (\mathrm{Exp}\, u)^n, \end{aligned}$$

where the third equality is due to a general property of symmetric spaces: if (N, p) is a symmetric space with base point and $g \in \mathrm{Aut}(N)$ such that $\sigma_p \circ g \circ \sigma_p = g^{-1}$, then $(g.p)^n = g^n.p$. (This is proved by an easy induction, using the definition of the powers in a symmetric space given in Equation (5.5).) This remark is applied to $(N, p) = ((T^k M)_o, 0_o)$ and $g = \exp_k(\sum_\alpha \varepsilon^\alpha l(v_\alpha))$.

Now let $\gamma(t) := \mathrm{Exp}(tu)$. Then, in any chart representation, $\gamma(t) = tu$ modulo higher order terms in t (since Exp is unipotent, as already proved), and hence $\gamma'(0) = u$. We have to show that $\gamma : \mathbb{K} \to T^k M$ is a homomorphism of symmetric spaces. As we have just seen, $\gamma(n) = (\gamma(1))^n$. By power associativity in symmetric spaces (cf. Equation (5.7)), and denoting by μ the binary product in the symmetric spaces $\mathbb{K}$, resp. $T^k M$, this implies $\gamma(\mu(n, m)) = \gamma(2n - m) = (\gamma(1))^{2n-m} = \mu(\gamma(m), \gamma(n))$, and thus the restriction of γ to $\mathbb{Z}$ is an "algebraic symmetric space homomorphism". But $\gamma : \mathbb{K} \to (T^k M)_o$ is a polynomial map between two symmetric spaces with polynomial multiplication maps, and hence by the "polynomial density argument" used several times in Section PG, it follows that also γ must be a homomorphism. ■

As in Step 1 of the proof of Theorem PG.6, it is seen that γ_u is the only *polynomial* one-parameter subspace with $\gamma_u'(0) = u$. We do not know wether this already implies that γ_u is the only *smooth* one-parameter subspace with this property.

In the finite-dimensional case over $\mathbb{K} = \mathbb{R}$, it is known that the exponential map of a symmetric space can be defined by

$$\mathrm{Exp}_o(v) = \mathrm{Fl}_1^{l(v)}(o) = \mathrm{ev}_o \circ \mathrm{Fl}_1 \circ l : \qquad \begin{array}{ccc} M & \stackrel{\mathrm{ev}_o}{\leftarrow} & \mathrm{Diff}(M) \\ \mathrm{Exp} \uparrow & & \uparrow \mathrm{Fl}_1 \\ T_o M & \stackrel{l}{\rightarrow} & \mathfrak{X}(M)_c \end{array}$$

(see [Lo69], [Be00, I.5.7]). As in the case of Lie groups, the definition of Exp_k from the preceding theorem is formally obtained by applying the functor $(T^k)_0$ to this diagram.

31. Remarks on the exponential jet of a general connection

31.1. *The exponential jet.* As explained in the Introduction (Section 0.16), for real finite dimensional or Banach manifolds M equipped with a (torsionfree) linear connection on TM, we can define a canonically associated bundle isomorphism $T^k \operatorname{Exp} : A^k M \to T^k M$, called the *exponential jet (of order k).* For a general base field or -ring $\mathbb{K}$, there is no exponential map Exp_x, but still it is possible to construct a bundle isomorphism $A^k M \to T^k M$ having all the good properties of the exponential jet and coinciding with $T^k \operatorname{Exp}$ in the real finite-dimensional or Banach case. In the following, we explain the basic ideas of two different approaches to such a construction; details will be given elsewhere.

31.2. *Vector field extensions.* The definition of the exponential jet for Lie groups and symmetric spaces in the preceding chapter suggests the following approach: in the real finite-dimensional case, every connection L gives rise to a, at least locally, on a star-shaped neighborhood U of x defined "vector field extension map" $l := l_x : T_x M \to \mathfrak{X}(U)$, assigning to $v \in T_x M$ the *adapted vector field* $l(v)$ which is defined by taking as value $(l(v))(p) \in T_p M$ the parallel transport of v along the unique geodesic segment joining x and p in U. Then clearly l is linear, and $l(v)(x) = v$. Moreover, one may check that the covariant derivative of L is recovered by $(\nabla_X Y)(x) = [l_x(X(x)), Y](x)$. The exponential map Exp_x is given, as in the preceding chapter, by $\operatorname{Exp}_x(v) = (\operatorname{Fl}_1^{l(v)})(x)$.

For general base fields and rings $\mathbb{K}$, one may formalize the properties of such "vector field extension maps". One does not really need the whole vector field $l(v) \in \mathfrak{X}(U)$, but only its k-jet at x. Then one may try to define the exponential jet in a "synthetic way" by using the exponential maps of the polynomial groups $\operatorname{Gm}^{1,k}(T^k M)_x$ and $\operatorname{Am}^{1,k}(T^k M)_x$ acting on the fiber over x.

31.3. *The geodesic flow.* Another "synthetic" approach to the exponential jet is motivated as follows: in the classical case, if the connection is complete, the geodesic flow $\Phi_t : TM \to TM$ is the flow of a vector field $S : TM \to T(TM)$ which is called the *spray of the connection* and which is equivalent to the (torsionfree part of) the connection itself (cf. Chapter 11). Then the exponential map $\operatorname{Exp}_x : T_x M \to M$ at a point $x \in M$ is recovered from the flow Fl_1^S via $\operatorname{Exp}_x(v) = \pi(\operatorname{Fl}_1^S(v))$, where $\pi : TM \to M$ is the canonical projection:

$$\begin{array}{ccc} TM & \overset{\operatorname{Fl}_1^S}{\to} & TM \\ \uparrow & & \downarrow \\ T_x M & \overset{\operatorname{Exp}_x}{\to} & M \end{array}$$

Applying the k-th order tangent functor T^k to this diagram and restricting to the fiber over 0_x, we express the exponential jet via the k-jet of Fl_1^S and canonical maps:

$$\begin{array}{ccc} T^{k+1} M & \overset{T^k \operatorname{Fl}_1^S}{\to} & T^{k+1} M \\ \uparrow & & \downarrow \\ T^k(T_x M) & \overset{T^k(\operatorname{Exp}_x)}{\to} & T^k M \end{array}$$

As explained in Chapter 29, under some conditions, the map $T^k \mathrm{Fl}_1^S$ can be expressed in our general framework via $\exp(\delta S)$, and thus we should get an expression of the exponential jet in terms of maps that all can be defined in a "synthetic" way. However, the problem in the preceding diagram is that we cannot simply restrict everything to the fiber over $x \in M$, because Exp_x does not take values in some "fiber over x". One may try to overcome this problem by deriving once more, or, in other words, by introducing one more scalar extension, thus imbedding $A^k M$ into $T^{k+2}M$ and not into $T^{k+1}M$ and defining

$$\mathrm{Exp}_k := \mathrm{pr} \circ \widetilde{\exp(\delta S)} \circ \iota : \qquad \begin{array}{ccc} T^{k+1}(TM) & \overset{\widetilde{\exp(\delta S)}}{\rightarrow} & T^{k+1}(TM) \\ \iota \uparrow & & \downarrow \mathrm{pr} \\ A^k M & \overset{\mathrm{Exp}_k}{\rightarrow} & T^k M \end{array}$$

where the imbedding ι and the projection pr are defined by suitable choices of k infinitesimal units among the $k+2$ infinitesimal units $\varepsilon_0, \ldots, \varepsilon_{k+1}$, and the diffeomorphism $\widetilde{\exp(\delta S)} : T^{k+2}M \to T^{k+2}M$ is associated to the spray $S : T_{\varepsilon_0}M \to T_{\varepsilon_1}T_{\varepsilon_0}M$ of the connection L in the way described by Theorem 29.4.

31.4. *Jet of the holonomy group.* A major interest of the construction of the exponential jet, either in the way indicated in 31.2 or in the one from 31.3, should be the possibility to describe the k-jet of the *holonomy group* of (M, L) at x. It should be given by parallel transport along "infinitesimal rectangles", which can be developed at arbitrary order into its "Taylor series". All these are interesting topics for further work.

32. From germs to jets and from jets to germs

32.1. *Infinity jets (the functor J^∞).* The jet projections $\pi_{\ell,k} : J^kM \to J^\ell M$ ($k \geq \ell$) form a projective system, i.e., if $k_1 \geq k_2 \geq k_3$, then $\pi_{k_3,k_1} = \pi_{k_3,k_2} \circ \pi_{k_2,k_1}$ (cf. Equation (8.6)). Therefore we can form the projective limit $J^\infty M := \lim_{\leftarrow} J^kM$. We do not define a topology on $J^\infty M$. As usual for projective limits, we have projections

$$\pi_k : J^\infty M \to J^kM, \quad u = (u_0, u_1, u_2, \ldots) \mapsto u_k$$

(where $u_j \in J^jM$ with $u_j = \pi_{j,i}(u_i)$). In particular,

$$\pi_0 : J^\infty M \to M, \quad u = (u_0, u_1, u_2, \ldots) \mapsto u_0$$

is the projection of the "algebraic bundle $J^\infty M$" onto its base M. Clearly, smooth maps $f : M \to N$ induce maps $J^\infty f : J^\infty M \to J^\infty N$, and the functorial rules hold. For instance, for every Lie group G we get an abstract group $J^\infty G$. Note that it is not possible to define a functor T^∞ in the same way since the projections associated to T^k form a cube and not an ordered sequence of projections.

32.2. *Formal jets.* There are several possible ways to define homomorphisms between fibers $(J^kM)_x \to (J^kN)_y$, $x \in M$, $y \in N$, $k \in \mathbb{N} \cup \{\infty\}$. Our definition will be such that, for $k = \infty$, homomorphisms in a fixed fiber essentially correspond to "formal power series". First, for $k \in \mathbb{N}$ and $x \in M$ and $y \in N$, we let

$$J^k_x(M,N)_y := \mathrm{Hom}((J^kM)_x, (J^kN)_y)),$$

where "Hom" means the space of all algebraic homomorphisms $f : (J^kM)_x \to (J^kN)_y$ in the strong sense that f comes from a totally symmetric and shift-invariant homomorphism of multilinear spaces $\widetilde{f} : (T^kM)_x \to (T^kN)_x$. (One would also get a category if one only requires that $\widetilde{f}$ be weakly symmetric in the sense of Section SA.2. But the condition of total symmetry and shift-invariance is necessary if one wants to integrate jets to germs, see below.) A *formal jet (from the fiber at $x \in M$ to the fiber at $y \in N$)* is the projective limit $f_\infty : (J^\infty M)_x \to (J^\infty N)_y$ given by a family $f_k \in J^k_x(M,N)_y$, $k \in \mathbb{N}$, such that $\pi_{\ell,k} \circ f_k = f_\ell \circ \pi_{\ell,k}$. We denote by $J^\infty_x(M,N)_y$ the set of formal jets from x to y. For instance, if $h : M \to N$ is a smooth map such that $h(x) = y$, then $(J^\infty h)_x : (J^\infty M)_x \to (J^\infty N)_y$ is a formal jet. It is not required that all formal jets are obtained in this way – see Section 32.5 below.

32.3. *Germs.* For manifolds M, N over $\mathbb{K}$ and $x \in M$, $y \in N$, we denote by $\mathcal{G}_x(M,N)_y$ the space of germs of locally defined smooth maps f from M to N such that $f(x) = y$, defined as usual by the equivalence relation $f \sim g$ iff the restrictions of f and g agree on some neighborhood of x. Then the k-jet $(J^kf)_x$ of a representative f of $[f] = f/\sim$ does not depend on the choice of f because, in a chart, $f|_U = g|_U$ implies, by the "Determination Principle" (Section 1.2), $f^{[k]}|_{U^{[k]}} = g^{[k]}|_{U^{[k]}}$ (notation from Chapter 1) and hence $d^kf(x) = d^kg(x)$ for all $x \in U$. Therefore, for all $k \in \mathbb{N} \cup \{\infty\}$, we have a well-defined *k-th order evaluation map*

$$\mathrm{ev}^k_x : \mathcal{G}_x(M,N)_y \to J^k_x(M,N)_y, \quad [f] \mapsto (J^kf)_x$$

assigning to a germ at x its k-th order jet there. The kernel of ev_x^k is the space of germs vanishing of order $k+1$ at x; if $k=\infty$ and $\mathrm{ev}_x^\infty([f])=0$, we say that f *vanishes of infinite order at* x.

Proposition 32.4. *Assume the integers are invertible in* $\mathbb{K}$. *Then, for all* $k \in \mathbb{N}$ *and* $x \in M$, $y \in N$, *the* k*-th order evaluation map* ev_x^k *is surjective.*

Proof. Let $f \in J_x^k(M,N)_y$, i.e., f is given by a totally symmetric and shift-invariant homomorphism $\widetilde{f} : (T^k M)_x \to (T^k N)_y$. We choose bundle charts at x and at y and represent $\widetilde{f}$ in the usual form of a homomorphism of multilinear maps

$$\widetilde{f}(\sum_\alpha \varepsilon^\alpha v_\alpha) = \sum_\alpha \varepsilon^\alpha \sum_{\ell=1}^{|\alpha|} \sum_{\Lambda \in \mathcal{P}_\ell(\alpha)} b^\Lambda(v_{\Lambda^1}, \ldots, v_{\Lambda^\ell}).$$

As shown in Theorem SA.16, total symmetry and shift-invariance imply that b^Λ depends essentially only on $\ell = \ell(\Lambda)$: there exist multilinear maps $d^\ell : V^\ell \to V$ (in the usual sense) such that

$$\widetilde{f}(\sum_\alpha \varepsilon^\alpha v_\alpha) = \sum_\alpha \varepsilon^\alpha \sum_{\ell=1}^{|\alpha|} \sum_{\Lambda \in \mathcal{P}_\ell(\alpha)} d^\ell(v_{\Lambda^1}, \ldots, v_{\Lambda^\ell}).$$

Then we define, still with respect to the fixed charts, on a neighborhood of x, a map

$$F(x+h) := y + \sum_{\ell=1}^k \frac{1}{\ell!} d^\ell(h, \ldots, h).$$

Then $d^\ell F(x) = d^\ell$ and hence, by Theorem 7.5., $(T^k F)_x = \widetilde{f}$, whence $(J^k F)_x = f$. Therefore, the germ $[F] \in \mathcal{G}_x(M,N)_y$ satisfies $\mathrm{ev}_x^k([F]) = f$. ■

In particular, for $k \in \mathbb{N}$, the evaluation maps from germs of curves to $J^k M$,

$$\mathrm{ev}_0^k : \mathcal{G}_0(\mathbb{K}, M)_x \to J_0^k(\mathbb{K}, M)_x \cong (J^k M)_x,$$
$$[\alpha] \mapsto (J^k \alpha)_0 \cong (J^k \alpha)_0(\delta 1)$$

are surjective.

32.5. *From jets to germs.* The problem of "integrating jets to germs" amounts to construct some sort of inverse of the "infinite order evaluation map" $\mathrm{ev}_x^\infty : \mathcal{G}_x(M,N)_y \to J_x^\infty(M,N)_y$. However, in general this map is not invertible – there are obstructions both to surjectivity and to injectivity.

(a) *Failure of surjectivitiy. Convergence.* Let us call the image of ev_x^∞ in $J_x^\infty(M,N)_y$ the *space of integrable jets*. For a general topological field or ring $\mathbb{K}$ it seems impossible to give some explicit description of this space. But if, e.g., $\mathbb{K}$ is a complete valued field, then we have the notion of *convergent power series* (see, e.g., [Se65]), and we may define the notion of *convergent jets*. Then the convergent jets are integrable. For instance, if G is a Lie group, then the sequence of exponential maps $\exp_k : (J^k \mathfrak{g})_0 \to (J^k G)_e$ defines a formal jet $\exp_\infty \in J_0^\infty(\mathfrak{g}, G)_e$ with inverse $\log_\infty$. It is of basic interest to know wether these jets are integrable or not. Our construction of $\exp_k$ and $\log_k$ is explicit enough (Theorem 25.2) to provide estimates assuring convergence in suitable situations.

(b) *Failure of injectivity. Analyticity.* The kernel of the evaluation map ev_x^∞ in $J_x^\infty(M,N)_y$ is the space of germs of smooth maps vanishing of infinite order at x and hence measures, in some sense, the "flabbiness" of our manifolds. In order to control this kernel one needs some sort of rigidity, e.g., the assumption that our manifolds are analytic (where again we need additional structures on $\mathbb{K}$ such as above).

Summing up, under the above mentioned assumptions, the construction of an inverse of ev_x^∞ would look as follows: given an infinite jet $f \in J_x^\infty(M,N)_y$, we pick a chart around x and write $f = \sum_{j=0}^\infty \delta^{(j)} d^j$ (i.e., $f = \lim_{\leftarrow} f_k$, $f_k = \sum_{j=0}^k \delta^{(j)} d^j$), then, if this jet is convergent, we let in a chart

$$F(x+h) = \sum_{j=0}^\infty \frac{1}{j!} d^j(h,\ldots,h),$$

which is "well-defined up to terms vanishing of infinite order at x".

32.6. *Analytic structures on Lie groups and symmetric spaces.* It is a classical result that finite-dimensional real Lie groups admit a compatible analytic Lie group structure. Recently, this result has been extended to the case of Lie groups over ultrametric fields by H. Glöckner [Gl03c]. However, the proof uses quite different methods than in the real case. It seems possible that, with the methods presented in this work, a common proof for both cases could be given, including also the case of symmetric spaces. The first step should be to prove that, in finite dimension over a complete valued field, the exponential jet $\exp_\infty$ of a Lie group or a symmetric space (Chapter 30) is in fact convergent and defines the germ of a diffeomorphism (with inverse coming from $\log_\infty$); then to define, with respect to some fixed chart, an exponential map at the origin by this convergent jet; to use left translations to define an atlas of G and finally to show that this atlas is analytic since transition functions are given by the Campbell-Hausdorff series (limit of Theorem 25.5).

Appendix L. Limitations

The prize for generalizing a theory like differential geometry is of course a limitation of the number of properties and results that will survive in the more general framework. Thus, in our general category of manifolds over topological fields or rings $\mathbb{K}$, some constructions which the reader will be used to from the theory of real manifolds (finite-dimensional, Banach, locally convex or other) are not possible. In particular, this concerns *dual* and *tensor product constructions* which would permit to construct new vector bundles from old ones. In this appendix we give a short overview over the kinds of limitations that we have to accept when working in the most general category.

L.1. *Cotangent bundles and dual bundles.* Formally, we can define a cotangent bundle T^*M following the pattern described in Chapter 3, just by taking the topological dual of the tangent space T_xM as fiber. However, in general there is no reasonable topology on topological duals which could serve to turn T^*M into a smooth manifold, and similarly for general dual bundles of vector bundles over M. Moreover, in order to define a manifold structure on T^*M it is not enough just to have the structure of a topological vector space on the dual of the model space – see [Gl03b] for a counterexample. It seems that the most natural framework where dual bundles always exist is the category of (real, complex or ultrametric) *Banach manifolds*, i.e. all model spaces are Banach spaces (cf. [La99], [Bou67]).

L.2. *Tensor products.* In our approach, the use of multilinear bundles allows us to develop the theory of connections and related notions without having to introduce tensor products.[1] However, in specific situations one may wish to have a tensor product of vector bundles at disposition. In fact, it is possible to define *topological tensor products* in the category of topological $\mathbb{K}$-modules be inquiring the usual universal property for *continuous* bilinear maps (generalizing the so-called *projective topology* from the real locally convex case, cf. [Tr67, Ch. 43]), and doing this construction fiberwise, we get an abstract bundle over the base M whose fibers are topological vector spaces. But in general it is not possible to turn this bundle into a manifold (see [Gl03b] for a discussion of the real locally convex case). Moreover, the topological tensor product of topological modules does not have the good properties known from the algebraic tensor product: tensor products in this sense are in general *not* associative ([Gl04b]); this fails already for (non-locally convex) real topological vector spaces. For all of these reasons, we will never use topological tensor products in the present work. As for dual bundles, the most natural category where tensor products of bundles make no problem seems to be the category of Banach manifolds (cf. [La99]).

L.3. *The main categories.* Here is a short (and by no means complete) list of the main categories of manifolds, resp. of their model spaces. We indicate what

[1] We agree with the remark in the introduction of [La99] (loc. cit. p. vi): "An abuse of multilinear algebra in standard treatises arises from an unnecessary double dualization and an abusive use of the tensor product."

constructions and results are still available, starting from general categories and ending with most specific ones. Details can be found in the quoted references. For more information on related topics we refer, in particular, to work of H. Glöckner (cf. References).

(1) *The case where $\mathbb{K}$ is a field.* Locality – see Section 2.4.

(2) *Metrizable topological vector spaces over $\mathbb{K} = \mathbb{R}$ or an ultrametric field.* Smoothness can be tested by composition with smooth maps $\mathbb{K}^k \to V$ from finite dimensional spaces into the model space [BGN04, Theorem 12.4].

(3) *Locally convex vector spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$.* Manifolds in our sense are the same as those in the sense of Michal-Bastiani [BGN04, Prop. 7.4].
   - Abundance of smooth functions (at least locally): the representation of the space of vector fields by differential operators,

$$\mathfrak{X}(U) \to \mathrm{End}_{\mathbb{K}}(\mathcal{C}^\infty(U)), \quad X \mapsto L_X,$$

     is injective for chart domains $U \subset M$; thus the Lie bracket can be defined via $L_{[X,Y]} = -[L_X, L_Y]$ (or taking the opposite sign).
   - Solutions of the differential equation $df = 0$ are locally constant: uniqueness results for flows (if they exist), etc.
   - Fundamental theorem of differential and integral calculus for curves: Poincaré Lemma (cf. [Ne02b, Lemma II.3.5]).

(4) *Complete locally convex vector spaces over $\mathbb{K} = \mathbb{R}, \mathbb{C}$.* One can speak about analytic maps; analytic manifolds can be defined, and they are smooth in our sense [BGN04, Prop. 7.20]. For general complete ultrametric fields this holds e.g. in the category of Banach manifolds [Bou67].

(5) *Banach manifolds over $\mathbb{K} = \mathbb{R}, \mathbb{C}$ or a complete ultrametric field.*
   - Cotangent, dual and hom-bundles exist in the given category.
   - Tensor products of vector bundles exist and have the usual properties.
   - The group $\mathrm{Gl}(V)$ of continuous linear bijections of the model space is a Lie group; hence the frame bundle of a manifold exists, and natural bundles are associated bundles of the frame bundle (in the sense of [KMS93]).
   - Inverse function and implicit function theorems (cf. [La99] in the real case and [Gl03a] for the ultrametric case, including much more general results); their consequences for the theory of submanifolds

(6) *Banach manifolds over $\mathbb{K} = \mathbb{R}$.* Existence and uniquenes theorem for ordinary differential equations and all of its consequences: existence of flows, geodesics, parallel transports; Frobenius theorem; exponential map: existence and local diffeomorphism. (See [La99] for all this.) Connected symmetric spaces are homogeneous (cf. [BeNe05]).

(7) *Finite dimensional manifolds over locally compact fields.*
   - Existence of compact objects.

– Existence of volume forms (cf. [Bou67]), Haar measure on Lie groups.

(8) *Finite dimensional real manifolds.*
- Existence of partitions of unity: global existence of metrics and connections and all consequences.
- Equivalence of all definitions of linear differential operators mentioned in Chapter 21: "algebraic definition" of tangent spaces and of vector fields; tensor fields are the same as differential operators of degree zero.

Appendix G. Generalizations

Generalizations of the present work are possible in two directions: one could treat more general classes of manifolds, or one could leave the framework of manifolds and use more formal concepts. In either case, it seems that the main features of our approach – use of scalar extensions and use of "geometric multilinear algebra" – could carry over and play a useful rôle.

G.1. *General C^0-concepts.* As to generalizations in the framework of manifolds, we just discuss here the manifolds associated to a general C^0*-concept* in the sense of [BGN04]. Differential calculus, as explained in Chapter 1, can be generalized as follows: $\mathbb{K}$ is a field or unital ring with topology and $\mathcal{M}$ a class of $\mathbb{K}$-modules with topology; for each open set U and each $W \in \mathcal{M}$, a set $C^0(U,W)$ of continuous mappings is given, such that several natural conditions are satisfied, among them the *determination principle* (DP) from Section 1.1 which makes a differential calculus possible. We require also that there is a rule defining a topology on the direct product of spaces from the class $\mathcal{M}$ (and making them a member of M) such that the ring operations (addition and multiplication) are of class C^0. A typical example is the class of finite dimensional vector spaces over an infinite field $\mathbb{K}$, equipped with their Zariski-topology, and C^0 meaning "rational": in this case, $\mathbb{K}$ is not a topological field in the usual sense, and the members of $\mathcal{M}$ are not topological vector spaces in the usual sense. Nevertheless, differential calculus, the theory of manifolds and differential geometry can be developed almost word by word as in the present work. This is true, in particular, for Theorem 6.2 and its proof. The only difference is that direct products of manifolds now carry a topology which, in general, is finer than the product topology; this is true also for local product situations: bundles are locally direct products and hence their topology, even locally, is in general finer then the local product topologies. We leave the details to the interested reader.

G.2. *Synthetic differential geometry.* Among the more formal approaches to differential geometry we would like to mention the so-called "synthetic differential geometry" (cf. [Ko81] or [Lav87]) and the theory of its models such as the so-called "smooth toposes" (over $\mathbb{R}$) (cf. [MR91]). In [MR91, p. 1–3], the authors give three main reasons for generalizing the theory of manifolds by the theory of smooth toposes:

(1) the category of smooth manifolds is not cartesian closed (spaces of mappings between manifolds are not always manifolds),
(2) the lack of finite inverse limits in the category of manifolds (in particular, manifolds can not have "singularities"),
(3) the absence of a convenient language to deal explicitly and directly with structures in the "infinitely small".

The "naive" approach by R. Lavendhomme [Lav87] focuses mainly on Item (3), whereas A. Kock in [Ko81] emphasizes Item (1), quoting F. W. Lawvere (loc. cit. p. 1): "...it is necessary that the mathematical world picture involve a cartesian closed category $\mathcal{E}$ of smooth morphisms between smooth spaces." Of course, this statement is more philosophical than mathematical in nature; but we feel that our approach proves at least that Items (1) and (3) are not as strongly coupled among each other as one used to believe up to now – the infinitely small has a very natural and useful place in our approach, and this works in a category based on conventional logic and usual set-theory and not needing any sophisticated model theory.[1] Nevertheless, it is certainly still interesting to generalize results to categories that behave well with respect to (1) and (2) – here it seems possible, in principle, to develop a theory of "smooth toposes over $\mathbb{K}$" (based on differential calculus over $\mathbb{K}$ as used here) and to transpose much of the material presented here to this context which then would unify differential geometry and algebraic geometry over topological base fields and -rings.

G.3. *Other categorial approaches.* Just as the concept of a polynomial is a formal generalization of the more primitive concept of a polynomial mapping (see [Ro63] for the general case; cf. also [BGN04, Appendix A]), it seems possible to generalize the differential calculus from [BGN04] in a formal way such that it becomes meaningful for any commutative ring or field, including finite ones. A "categorial" generalisation of ordinary differential calculus has already been given by L.D. Nel [N88]; however, instead of taking Condition (0.2) as starting point, he works with the condition $f(y) - f(x) = \Phi(x, y) \cdot (y - x)$ which is well-adapted for differential calculus in *one* variable, but causes much trouble in the further development of differential calculus in several variables. Since the arguments proving our Theorem 6.2 are of purely categorial nature, we believe that the scalar extension functor should exist and play a central rôle in a categorial theory based on (0.2).

[1] cf. [Lav87, p. 2]: "...nous nous sommes strictement limités au point de vue "naïf" ... sans tenter d'en décrire des "modèles" qui eux vivent dans des topos hors d'atteinte à ce niveau élémentaire."

Appendix: Multilinear Geometry

This appendix is completely independent of the main text and contains all algebraic definitions and results that are used there. Chapter BA on bilinear algebra is the special case $k = 2$ of Chapter MA on multilinear algebra, and the reader who is interested in general results may directly start by looking at Chapter MA. However, the case $k = 2$ is of independent interest since it corresponds to *linear* connections (Chapter 10). – In the following, $\mathbb{K}$ is any commutative ring with unit 1.

BA. Bilinear algebra

BA.1. *Bilinearly related structures.* Assume V_1, V_2, V_3 are three $\mathbb{K}$-modules together with a bilinear map $b : V_1 \times V_2 \to V_3$. We let $E := V_1 \times V_2 \times V_3$. Then there is a structure of a $\mathbb{K}$-module on E, depending on b, and given by

$$\begin{aligned}(u,v,w) +_b (u',v',w') &:= (u+u', v+v', w+w'+b(u,v')+b(u',v)),\\ \lambda_b(u,v,w) &:= (\lambda u, \lambda v, \lambda w + \lambda(\lambda-1)b(u,v)).\end{aligned} \tag{BA.1}$$

One may either check directly the axioms of a $\mathbb{K}$-module, or use the following observation: the map

$$f_b : E \to E, \quad (u,v,w) \mapsto (u,v,w+b(u,v)) \tag{BA.2}$$

is a bijection whose inverse is given by $f_b^{-1}(u,v,z) = (u,v,z-b(u,v))$. In fact, it is immediately seen that

$$f_b \circ f_c = f_{b+c}, \quad f_0 = \mathrm{id}_E,$$

i.e. the additive group $\mathrm{Bil}(V_1 \times V_2, V_3)$ of bilinear maps from $V_1 \times V_2$ to V_3 acts on E. This action is not linear, and a straighforward calculation shows that the linear structure

$$L^b := (E, +_b, \cdot_b) \tag{BA.3}$$

is simply the push-forward of the original structure (corresponding to $b = 0$) via f_b, i.e.

$$\lambda_b x = f_b(\lambda\, f_b^{-1}(x)), \quad x +_b y = f_b(f_b^{-1}(x) + f_b^{-1}(y)),$$

where $+$ means $+_0$. We call two linear structures L^b, L^c with $b, c \in \mathrm{Bil}(V_1 \times V_2, V_3)$ *bilinearly related*, and we denote by $\mathrm{brs}(V_1, V_2; V_3)$ the space of all $\mathbb{K}$-module structures on E that are bilinearly related to the original structure. From Formula (BA.1) we get the following expressions for the negative, the difference and the barycenters with respect to L^b:

$$\begin{aligned}&-_b\,(u,v,w) = (-u,-v,-w+2b(u,v)),\\ &(u,v,w) -_b (u',v',w') = (u-u', v-v', w-w'+2b(u',v')-b(u,v')-b(u',v)),\\ &(1-r)_b(x,y,z) +_b r_b(x',y',z') =\\ &\qquad ((1-r)x+rx', (1-r)y+ry', (1-r)z+rz'+r(r-1)b(x-x',y-y')).\end{aligned} \tag{BA.4}$$

Translations and Dilations now depend on b. We will take account of this by adding b as a subscript in the notation: for $x = (x_1, x_2, x_3) \in E$ and $b \in \mathrm{Bil}(V_1 \times V_2, V_3)$, the *translation by x with respect to b* is

$$\tau_{b,x} : E \to E, \quad (u, v, w) \mapsto (x_1, x_2, x_3) +_b (u, v, w), \tag{BA.5}$$

and the *dilation with ratio $r \in \mathbb{K}$ with respect to b* is

$$r_b : E \to E, \quad x \mapsto r_b x = (rx_1, rx_2, rx_3 + (r^2 - r)b(x_1, x_2)). \tag{BA.6}$$

BA.2. *Intrinsic objects of bilinear geometry.* We have remarked above that the set of transformations of E given by

$$\mathrm{Gm}^{1,2}(E) := \{f_b |\ b \in \mathrm{Bil}(V_1 \times V_2, V_3)\} \tag{BA.7}$$

is a vector group, isomorphic to the $\mathbb{K}$-module $\mathrm{Bil}(V_1 \times V_2, V_3)$. It will be called the *special bilinear group.* As already remarked, it acts simply transitively on the set $\mathrm{brs}(V_1, V_2; V_3)$ of bilinearly related structures which in this way is turned into an affine space over $\mathbb{K}$. The "original" linear structure (which corresponds to $b = 0$) shall not be preferred to the others. By *bilinear geometry* we mean the study of all geometric and algebraic properties of E that are common for all bilinearly related linear structures, or, equivalently, that are invariant under the special bilinear group. Slightly more generally, we may consider also all properties that are invariant under the *general bilinear group*

$$\mathrm{Gm}^{0,2}(E) := \{f = (h_1, h_2, h_3) \circ f_b |\ h_i \in \mathrm{Gl}(V_i),\ i = 1, 2, 3,\ f_b \in \mathrm{Gm}^{1,2}(E)\}, \tag{BA.8}$$

i.e. under all bijections of E having the form

$$f(u, v, w) = (h_1(u), h_2(v), h_3(w + b(u, v)))$$

with b as above and invertible linear maps h_i. In fact, the group $\mathrm{Gm}^{0,2}(E)$ also acts by push-forward on the space of bilinearly related structures. The stabilizer of the base structure L^0 is the group $H := \mathrm{Gl}(V_1) \times \mathrm{Gl}(V_2) \times \mathrm{Gl}(V_3)$. The group H acts affinely on the affine space $\mathrm{brs}(V_1, V_2; V_3)$ and hence acts linearly with respect to zero vector L^0 in this space. The situation is the same as for usual affine groups, and hence we have an exact splitting sequence of groups

$$0 \quad \to \quad \mathrm{Gm}^{1,2}(E) \quad \to \quad \mathrm{Gm}^{0,2}(E) \quad \to \quad \mathrm{Gl}(V_1) \times \mathrm{Gl}(V_2) \times \mathrm{Gl}(V_3) \quad \to \quad 1\,. \tag{BA.9}$$

Now we say that an object on E is *intrinsic* if its definition is independent of b. Here are some examples:

(A) Injections, axes. The origin $(0, 0, 0) \in E$ is invariant under $\mathrm{Gm}^{1,2}(E)$, and so are the three "axes" $V_1 \times 0 \times 0$, $0 \times V_2 \times 0$, $0 \times 0 \times V_3$ (referred to as "first, second, third axis"; they are submodules isomorphic to, and often identified with, V_1, V_2, resp. V_3). Also, the submodules $V_1 \times 0 \times V_3$ and $0 \times V_2 \times V_3$ together with their inherited module structure are intrinsic, but the "submodule $V_1 \times V_2$" (which should correspond to $V_1 \oplus_b V_2 \oplus_b 0$) is not intrinsic: in fact,

$$f_b(V_1 \times V_2 \times 0) = V_1 \oplus_b V_2 \oplus_b 0 = \{(u, v, b(u, v)) |\ u \in V_1, v \in V_2\} \tag{BA.10}$$

clearly depends on b (it is the graph of b). Thus we have the following intrinsic diagram of inclusion of "axes" and "planes":

$$\begin{array}{ccccc} V_1 & \to & E^{13} := V_1 \times V_3 & & \\ & \nearrow & & \searrow & \\ V_3 & & & & E = V_1 \times V_2 \times V_3 \\ & \searrow & & \nearrow & \\ V_2 & \to & E^{23} := V_2 \times V_3 & & \end{array}$$

(B) Projections, fibrations. It follows immediately from (BA.1) that the two projections

$$\mathrm{pr}_j : V_1 \times V_2 \times V_3 \to V_j \tag{BA.11}$$

for $j = 1, 2$ are invariant under f_b in the sense that $\mathrm{pr}_j \circ f_b = \mathrm{pr}_j$, and hence they are linear for every choice of linear structure L^b. (This is not true for the third projection!) Thus also

$$\mathrm{pr}_{12} := \mathrm{pr}_1 \oplus \mathrm{pr}_2 : V_1 \times V_2 \times V_3 \to V_1 \times V_2, \quad (u, v, w) \mapsto (u, v) \tag{BA.12}$$

is an intrinsic linear map, and the following three exact sequences are intrinsically defined:

$$\begin{array}{ccccccccc} 0 & \to & V_3 & \to & V_1 \times V_2 \times V_3 & \xrightarrow{\mathrm{pr}_{12}} & V_1 \times V_2 & \to 0 \\ 0 & \to & V_2 \times V_3 & \to & V_1 \times V_2 \times V_3 & \xrightarrow{\mathrm{pr}_1} & V_1 & \to 0 \\ 0 & \to & V_1 \times V_3 & \to & V_1 \times V_2 \times V_3 & \xrightarrow{\mathrm{pr}_2} & V_2 & \to 0 \end{array} \tag{BA.13}$$

The fibers of pr_j are invariant under $\mathrm{Gm}^{1,2}(E)$, and moreover they are affine subspaces of E, for any choice of linear structure L^b. But more is true: the induced affine structure on the fibers does not depend on L^b at all.

Proposition BA.3. *For all $z \in E$, the spaces*

$$E_z^{13} := z +_b E^{13}, \quad E_z^{23} := z +_b E^{23}$$

are intrinsic affine subspaces in the sense that both the sets and the structures of an affine space over $\mathbb{K}$ induced from the ambient space (E, L^b) are independent of b.

Proof. Let b be fixed. The map f_b preserves the subset

$$z +_0 E^{13} = \{(z_1 + v, z_2, z_3 + w) | \, v \in V_1, w \in V_3\}$$

on which it acts *linearly*: with respect to the origin $(0, z_2, 0)$ it is described, in the obvious matrix notation, by

$$\begin{pmatrix} \mathbf{1} & b(\cdot, z_2) \\ 0 & \mathbf{1} \end{pmatrix}.$$

Therefore the set and the linear structure on this set are independent of b. The claim for E_z^{13} now follows since E_z^{13} and $z +_0 E^{13}$ are the same sets, and similarly for E_z^{23}. ■

The proposition shows that E has an *intrinsic double fibration* by affine spaces. Every fiber contains one distinguished element (intersection with an axis), and taking this element as origin, the fiber has a canonical linear structure (e.g., the origin in $V^{13}_{(u,v,w)}$ is $(0,v,0)$, etc.) This implies, for instance, that the maps $E \to E$, $(u,v,w) \mapsto (\lambda u, v, \lambda w)$ and $(u,v,w) \mapsto (u, \lambda v, \lambda w)$ have an intrinsic interpretation, and so has their composition $(u,v,w) \mapsto (\lambda u, \lambda v, \lambda^2 w)$.

(C) Parallelism. The notion of being "parallel to an axis" is not always intrinsic: it is so for the parallels to the third axis, that is, fibers of pr_{12}, which we call *vertical spaces*, written, for $z = (z_1, z_2, z_3) \in E$,

$$E^3_z := E^{13}_z \cap E^{23}_z = z +_b V_3 = \{(z_1, z_2, z_3 + w) | \, w \in V_3\}. \tag{BA.14}$$

But the parallel to, say, $0 \times V_2 \times 0$ through $(u,0,0)$ is the "horizontal space"

$$H_{u,b} := (u,0,0) +_b V_2 = \{(u,0,0) +_b (0,v,0) = (u,v,b(u,v)) | \, v \in V_2\} \tag{BA.15}$$

which clearly depends on b.

Special cases

BA.4. *The special case $V_2 = V_3$ and shifts.* In the setup of Section BA.1, consider the special case $V_2 = V_3$. Then a direct computation shows that the *shift*

$$S = \begin{pmatrix} 1 & & \\ & 0 & 0 \\ & 1 & 0 \end{pmatrix} : V_1 \times V_2 \times V_3 \to V_1 \times V_2 \times V_3, \quad (u,v,w) \mapsto (u,0,v) \tag{BA.16}$$

commutes with all $f_b \in \mathrm{Gm}^{1,2}(E)$, and hence S is an "intrinsic linear map", i.e. it is linear with respect to all L^b. It stabilizes the fibers E^{23}_z, and the restriction to E^{23}_z, with respect to the origin $(\mathrm{pr}_1(z), 0, 0)$, is a two-step nilpotent linear map given in matrix form by the lower right corner of the matrix S. It defines on E^{23}_z an intrinsic structure of a module over the dual numbers $\mathbb{K}[\varepsilon]$. If $V_1 = V_3$, we get in a similar way a shift S'.

BA.5. *The special case $V_1 = V_2$ and torsion.* In case $V_1 = V_2$, the symmetric group $\Sigma_2 = \{\mathrm{id}, (12)\}$ acts on $E = V_1 \times V_1 \times V_3$, on $\mathrm{Bil}(V_1 \times V_1, V_3)$ and on the space of bilinearly related linear structures, where the transposition (12) acts by the *exchange map* or *flip*

$$\kappa : V_1 \times V_1 \times V_3 \to V_1 \times V_1 \times V_3, \quad (u,v,w) \mapsto (v,u,w).$$

We say that L^b is *torsionfree* if b is symmetric; this means that the exchange map is a linear automorphism with respect to L^b. In general, the push-forward of L^b by κ is given by $\kappa \cdot L^b = L^{\kappa \cdot b}$, where

$$\kappa \cdot b(u,v) = b(v,u).$$

The difference $L^b - L^c$ in the affine space $\mathrm{brs}(V_1, V_2; V_3)$ corresponds to the difference $b - c$; thus the difference $L^b - \kappa \cdot L^b$ corresponds to the skew-symmetric bilinear map $t(u,v) := b(u,v) - b(v,u)$, called the *torsion of the linear structure* L^b.

BA.6. *Quadratic spaces.* Assume $V_1 = V_2$ and b is symmetric. Then the κ-fixed space $\{(v_1,v_1,v_3) | v_1 \in V_1, v_3 \in V_3\}$ is a submodule of (E, L^b), isomorphic to $V_1 \times V_3$ with $\mathbb{K}$-module structure

$$\begin{aligned}(v_1,v_3) +_b (v_1',v_3') &= (v_1 + v_1', v_3 + v_3' + b(v_1,v_1') + b(v_1',v_1)) \\ \lambda(v_1,v_3) &= (\lambda v_1, \lambda v_3 + \lambda(\lambda-1)b(v_1,v_1)).\end{aligned}$$

Note that this structure is already determined by the quadratic map $q(v) = b(v,v)$, even in case 2 is not invertible in $\mathbb{K}$, and thus for any quadratic map $q : V_1 \to V_3$ we may define a linear structure on $V_1 \times V_3$ given by

$$\begin{aligned}(v_1,v_3) +_q (v_1',v_3') &= (v_1 + v_1', v_2 + v_2' + q(v_1 + v_1') - q(v_1') - q(v_1)) \\ \lambda(v_1,v_3) &= (\lambda v_1, \lambda v_3 + \lambda(\lambda-1)q(v_1)),\end{aligned}$$

leading to the notion of a *quadratic space* (we will not give here an axiomatic definition). The third axis still corresponds to an intrinsic subspace in $V_1 \times V_3$, but the first and second axis do not give rise to intrinsic subspaces: they are replaced by the *diagonal imbedding*

$$V_1 \to V_1 \times V_1 \times V_3, \quad v \mapsto (v,0,0) +_b (0,v,0) = (v,v,q(v))$$

which clearly depends on q and hence on b.

Intrinsically bilinear maps

Proposition BA.7. *Let E, V_i, $i = 1,2,3$ be as above and W be a linear space, i.e. a $\mathbb{K}$-module. For a map $\omega : E \to W$ the following properties are equivalent:*

(1) *There exists a linear map $\lambda : V_3 \to W$ and a $\mathbb{K}$-bilinear map $\nu : V_1 \times V_2 \to W$ such that*
$$\omega(u,v,w) = \nu(u,v) + \lambda(w).$$

(2) *For all $z \in E$, the restrictions*
$$\omega_z^1 := \omega|_{E_z^{13}} : E_z^{13} \to W, \quad \omega_z^2 := \omega|_{E_z^{23}} : E_z^{23} \to W$$
to the fibers of the canonical double fibration are linear.

Proof. Assume (1) holds. Let $z = (0,v,0)$; then $\omega_z^1(u,w) = \nu(u,v) + \lambda(w)$ is linear in (u,w). Conversely, assume ω satisfies (2). We let $\lambda := \omega|_{V_3}$ (restriction to third axis) and $\nu := \omega \circ \iota_{12}$, where $\iota_{12} : V_1 \times V_2 \to E$ depends on the linear structure L^0. Then, by (2), $\nu(u,v) = \omega(u,v,0)$ is linear in u and in v, i.e. it is $\mathbb{K}$-bilinear. We decompose

$$(u,v,w) = ((0,0,w) +_{(0,0,0)} (0,v,0)) +_{(0,v,0)} (u,v,0),$$

where the first sum is taken in the fiber E_0^{13} and the second sum in the fiber E_z^{23} with $z = (0,v,0)$. Using Property (2), we get

$$\begin{aligned}\omega(u,v,w) &= \omega((0,0,w) +_{(0,0,0)} (0,v,0)) + \omega(u,v,0) \\ &= \omega(0,0,w) + \omega(0,v,0) + \omega(u,v,0) \\ &= \lambda(w) + \nu(v,0) + \nu(u,v) = \lambda(w) + \nu(u,v)\end{aligned}$$

as had to be shown. ∎

If ω satisfies the equivalent conditions of the preceding proposition, we say that ω is *(intrinsically) bilinear.* For instance, the third projection with respect to a fixed linear structure, $\mathrm{pr}_3^b : E \to V_3$, is intrinsically bilinear (it corresponds to $W = V_3$, $\nu = 0$, $\lambda = \mathrm{id}_{V_3}$).

Corollary BA.8. *In the situation of the preceding proposition are equivalent:*

(1) $\lambda = 0$.

(2) *For all $z \in E$, the restriction of ω to the third axis is constant,*

$$\forall w \in V_3 : \quad \omega(z_1, z_2, z_3 + w) = \omega(z_1, z_2, z_3). \qquad \blacksquare$$

If ω satisfies the properties of the corollary, then we say that ω is *homogeneous bilinear*, and we let $\mathcal{M}_h(E, W)$ the space of homogeneous bilinear maps $E \to W$. The preceding statements show that

$$\mathrm{Bil}(V_1 \times V_2; W) \to \mathcal{M}_h(E, W), \quad \nu \mapsto \omega := \nu \circ \mathrm{pr}_{12}$$

is a bijection. Its inverse $\omega \mapsto \omega \circ \iota_{12}$ is not canonical since the injection ι_{12} depends on the choice of a linear structure on E.

BA.9. *The skew-symmetrization operator.* Assume we are in the special case $V_1 = V_2$. Then to every intrinsically bilinear map $\omega : E \to W$ we can associate a *homogeneous* intrinsically bilinear map

$$\mathrm{alt}\,\omega := \omega - \omega \circ \kappa$$

In fact, ω and $\omega \circ \kappa$ are intrinsically bilinear, hence so is $\omega - \omega \circ \kappa$, and $\mathrm{alt}\,\omega$ is homogeneous: writing ω in the form $\omega(u, v, w) = \nu(u, v) + \lambda(w)$ as in the preceding paragraph, we get

$$\mathrm{alt}\,\omega(u, v, w) = \omega(u, v, w) - \omega(v, u, w) = \nu(u, v) - \nu(v, u).$$

which clearly is homogeneous bilinear. The bilinear map $V_1 \times V_1 \to V_3$ corresponding to $\mathrm{alt}\,\omega$ is

$$\nu_{\mathrm{alt}\,\omega}(u, v) = \omega(u, v, w) - \omega(v, u, w).$$

For instance, if $\omega = \mathrm{pr}_3^b$, then $\mathrm{alt}\,\omega$ is the torsion of L^b.

Relation with the tensor product

BA.10. *The tensor space.* It is possible to translate everything we have said so far into usual *linear* algebra by using *tensor products.* For arbitrary $\mathbb{K}$-modules V_1, V_2, V_3, we define the *tensor space* $Z := (V_1 \otimes V_2) \oplus V_3$ and the *big tensor space* $\widehat{Z} := V_1 \oplus V_2 \oplus V_3 \oplus (V_1 \otimes V_2)$ together with the canonical maps

$$\begin{aligned} &V_1 \times V_2 \times V_3 \to Z, \quad (v_1, v_2, v_3) \mapsto v_1 \otimes v_2 + v_3, \\ &V_1 \times V_2 \times V_3 \to \widehat{Z}, \quad (v_1, v_2, v_3) \mapsto v_1 + v_2 + v_3 + v_1 \otimes v_2. \end{aligned} \tag{BA.17}$$

We claim that the linear structure on Z, resp. on $\widehat{Z}$ is intrinsic, i.e. independent of the linear structure L^b on $V_1 \times V_2 \times V_3$: in fact, let $B : V_1 \otimes V_2 \to V_3$ be the linear map corresponding to $b : V_1 \times V_2 \to V_3$ and let

$$F_b := \begin{pmatrix} \mathbf{1} & 0 \\ B & \mathbf{1} \end{pmatrix} : Z \to Z.$$

Then the diagram

$$\begin{array}{ccc} (V_1 \times V_2) \times V_3 & \rightarrow & Z \\ f_b \downarrow & & \downarrow F_b \\ (V_1 \times V_2) \times V_3 & \rightarrow & Z \end{array} \tag{BA.18}$$

commutes. Since F_b is a linear automorphism of Z, it follows that the linear structure on Z does not depend on b, and similarly for $\widehat{Z}$. Thus Z and $\widehat{Z}$ are a sort of universal linear space for all bilinearly equivalent structures. (The space $\widehat{Z}$ is faithful in the sense that different elements of $\mathrm{Gm}^{0,2}(E)$ induce different automorphisms of $\widehat{Z}$, whereas for Z this property holds only with respect to the smaller group $\mathrm{Gm}^{1,2}(E)$.) Note that the splitting of Z into a direct sum is not intrinsic, but the subspace $0 \oplus V_3$ is intrinsic: the sequence

$$0 \to V_3 \to Z \to V_1 \otimes V_2 \to 0 \tag{BA.19}$$

is intrinsic. (If $V_1 = V_2$, then, with respect to symmetric b's, then we may define $Z' := S^2V_1 \oplus V_3$ and we get an intrinsic sequence $V_3 \to Z' \to S^2V_1$.) As we have seen, in presence of a fixed linear structure L^b, this sequence splits. Conversely, a splitting $\beta : V_1 \otimes V_2 \to Z$ of this sequence induces a linear map $B := \mathrm{pr}_{V_3} \circ \beta : V_1 \otimes V_2 \to V_3$ (where the projection is taken w.r.t. the splitting given by $b = 0$); then, if b is the bilinear map corresponding to B, it is seen that the splitting β is the one induced by L^b. Summing up, splittings of (BA.19) are in one-to-one correspondence with bilinearly related structures. Once again this shows that the space of bilinearly related structures carries the structure of an affine space over $\mathbb{K}$. Moreover, intrinsically bilinear maps $f : V_1 \times V_2 \times V_3 \to W$ correspond to linear maps $F : Z \to W$; homogeneous ones correspond to linear maps that vanish on W. In particular, the "third projection" corresponds to the linear map $Z \to V_3$, and the projection $Z \to V_1 \otimes V_2$ corresponds to f_ω, where $\omega : V_1 \times V_2 \to V_1 \otimes V_2$ is the tensor product map.

MA. Multilinear algebra

For a correct generalization of the bilinear theory from the preceding chapter to the multilinear case we have to change notation: the $\mathbb{K}$-modules V_1, V_2 from BA.1 play a symmetric rôle and should be denoted by V_{01}, V_{10}, whereas V_3 belongs to a "higher" level and shall be denoted by V_{11}. The multi-indices $01, 10, 11$ in turn correspond to non-empty subsets of the 2-element set $\{1, 2\}$. In the general k-multilinear case, our $\mathbb{K}$-modules will be indexed by the lattice of non-empty subsets of the k-element set $\{1, \ldots, k\}$.

MA.1. *The index set.* For a finite set M, we denote by 2^M its power set, and taking for M the k-element set

$$\mathbb{N}_k := \{1, \ldots, k\}, \tag{MA.1}$$

we define our *index set* $I := I_k := 2^{\mathbb{N}_k}$. Often we will identify I_k with $\{0,1\}^k$ via the bijection

$$2^{\mathbb{N}_k} \to \{0,1\}^k, \quad A \mapsto \alpha := \mathbf{1}_A, \tag{MA.2}$$

where $\mathbf{1}_A$ is the characteristic function of a set $A \subset \mathbb{N}_k$. The inverse mapping assigns to $\alpha \in \{0,1\}^k$ its *support*

$$A := \mathrm{supp}(\alpha) := \{i \in \mathbb{N}_k | \, \alpha_i = 1\}.$$

The natural total order on $\mathbb{N}_k$ induces a natural total order "$\leq$" on I_k which is the lexicographic order. For instance, I_3 is ordered as

$$\begin{aligned} I_3 &= (\emptyset, \{1\}, \{2\}, \{1,2\}, \{3\}, \{3,1\}, \{3,2\}, \{3,2,1\}) \\ &\cong ((000), (001), (010), (011), (100), (101), (110), (111)), \end{aligned} \tag{MA.3}$$

where, for convenience, we prefer to write elements $\alpha \in \{0,1\}^3$ in the form $(\alpha_3, \alpha_2, \alpha_1)$ and not $(\alpha_1, \alpha_2, \alpha_3)$. There are other useful total orderings of I_k, but the natural order has the advantage that it is compatible with the natural inclusion $\mathbb{N}_k \subset \mathbb{N}_{k+1}$ and hence is best adapted to induction procedures. In terms of multi-indices $\alpha \in \{0,1\}^k$, the basic set-theoretic operations are interpreted as follows:

(a) the cardinality of A corresponds to $|\alpha| := \sum_i \alpha_i$;

(b) the singletons $A = \{i\}$ correspond to the "canonical basis vectors" e_i, $i = 1, \ldots, k$;

(c) inclusion $\alpha \subseteq \beta$ is defined by $\mathrm{supp}(\alpha) \subseteq \mathrm{supp}(\beta)$ (note that this partial order is compatible with the total order: $\alpha \subseteq \beta$ implies $\alpha \leq \beta$);

(d) disjointness will be denoted by

$$\alpha \perp \beta \quad :\Leftrightarrow \quad \mathrm{supp}(\alpha) \cap \mathrm{supp}(\beta) = \emptyset \quad \Leftrightarrow \quad \sum_i \alpha_i \beta_i = 0, \tag{MA.4}$$

which is equivalent to saying that the componentwise sum $\alpha + \beta$ is again an element of $\{0,1\}^k$.

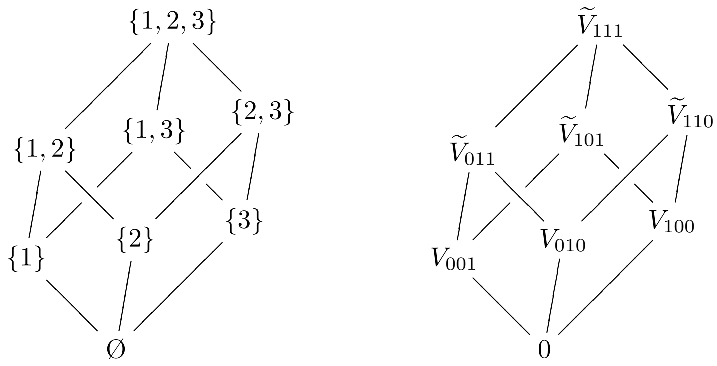
MA.2. *Cubes of $\mathbb{K}$-modules.* A *k-dimensional cube of $\mathbb{K}$-modules* is given by a family V_α, $\alpha \in I_k \setminus 0$, of $\mathbb{K}$-modules. We define the *total space*

$$E := \bigoplus_{\alpha \in I_k \setminus 0} V_\alpha .$$

Elements of E will be denoted by $v = \sum_{\alpha > 0} v_\alpha$ or by $v = (v_\alpha)_{\alpha > 0}$ with $v_\alpha \in V_\alpha$. The subspaces $V_\alpha \subset E$ are called the *axes of E*. For every $\alpha \in I$, $\alpha > 0$, we let

$$\widetilde{V}_\alpha := \bigoplus_{\beta \subseteq \alpha} V_\beta , \qquad H_\alpha := \bigoplus_{\alpha \subseteq \beta} V_\beta . \tag{MA.5}$$

If $\alpha \subseteq \alpha'$, then $\widetilde{V}_\alpha \subseteq \widetilde{V}_{\alpha'}$, and the lattice of subsets of I_k corresponds thus to a commutative cube of inclusions. For instance, for $k = 3$,



(MA.6)

MA.3. *Partitions.* A *partition* of a finite set A is given by a subset $\Lambda = \{\Lambda^1, \dots, \Lambda^\ell\}$ of the power set of A such that all Λ^i are non-empty and A is the disjoint union of the Λ^i: $A = \dot\cup_{i=1,\dots,\ell} \Lambda^i$. We then say that $\ell := \ell(\Lambda)$ is the *length* of the partition. We denote by $\mathcal{P}(A)$ the set of all partitions of A and by $\mathcal{P}_\ell(A)$ (resp. by $\mathcal{P}_{\geq \ell}(\alpha)$) the set of all partitions of α of length ℓ (resp. at least ℓ):

$$\mathcal{P}_\ell(A) = \{\Lambda = \{\Lambda^1, \dots, \Lambda^\ell\} \in 2^{(2^A)} | \quad \dot\cup_{i=1,\dots,\ell} \Lambda^i = A; \emptyset \notin \Lambda\}. \tag{MA.7}$$

Finally, for a finite set M, we denote by

$$\mathrm{Part}(M) := \bigcup_{\substack{A \in 2^M \\ A \neq \emptyset}} \mathcal{P}(A) \subset 2^{2^M} \tag{MA.8}$$

the set of all partitions of non-empty subsets of M, and for a partition Λ of some set $A \subset M$, the set A is called the *total set of the partition* Λ, denoted by

$$\underline{\Lambda} := A = \bigcup_{\nu \in \Lambda} \nu . \tag{MA.9}$$

If we use multi-index notation, this corresponds to $\underline{\lambda} := \alpha = \sum_i \lambda^i$.

Now let $M = \mathbb{N}_k$, $A \subset M$. Using the total order on $I_k = 2^{\mathbb{N}_k}$, every partition of length ℓ can also be written as an ordered ℓ-tupel:

$$\Lambda = (\Lambda^1, \ldots, \Lambda^\ell), \quad \emptyset < \Lambda^1 < \ldots < \Lambda^\ell \leq A.$$

In terms of multi-indices, a *partition of a multi-index* $\alpha \in \{0,1\}^k$ is described as $\lambda = \{\lambda^1, \ldots, \lambda^\ell\} \subset I_k$ such that $\sum_{i=1}^{\ell} \lambda^i = \alpha$ and all $\lambda^i > 0$, or, using ordered sequences:

$$\lambda = (\lambda^1, \ldots, \lambda^\ell) \in (I_k)^\ell, \quad 0 < \lambda^1 < \ldots < \lambda^\ell, \quad \sum_{i=1}^{\ell} \lambda^i = \alpha. \tag{MA.10}$$

The *trivial partition* $\{A\}$ of a set A, i.e. the one of length one, is identified with A itself. Sometimes it is useful to represent a partition by an $\ell \times k$-matrix with rows $\lambda^1, \ldots, \lambda^\ell$. So, for instance,

$$\mathcal{P}(\{1,2,3\}) = \Big\{\{1,2,3\}, \{\{1\},\{2,3\}\}, \{\{2\},\{1,3\}\}, \{\{1,2\},\{3\}\}, \{\{1\},\{2\},\{3\}\}\Big\}$$

corresponds to

$$\mathcal{P}((111)) = \left\{\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\right\}. \tag{MA.11}$$

The set $\mathcal{P}(A)$ can again be totally ordered, but we will not use this order.

The symmetric group Σ_k acts on $\mathbb{N}_k$, on I_k and on $\mathrm{Part}(\mathbb{N}_k)$ in the usual way. Then $\alpha, \beta \in I_k$ are conjugate under this action iff $|\alpha| = |\beta|$, and two partitions $\lambda, \mu \in \mathcal{P}(I_k)$ are conjugate under Σ_k iff

$$\ell(\mu) = \ell(\lambda) =: \ell \quad \text{and} \quad \forall i = 1, \ldots, \ell : \, |\lambda^i| = |\mu^i|. \tag{MA.12}$$

For instance, in (MA.11) the second, third and fourth partition are equivalent under Σ_3. Note that, when writing partitions as ordered sequences, the action of Σ_k does not preserve the order. In (MA.11), the lines of the fourth matrix had to be exchanged after letting act the transposition (23) on the third matrix.

MA.4. *Refinements of partitions.* We say that a partition Λ is a *refinement* of another partition Ω, or that Ω is *coarser than* Λ, and we write $\Omega \preceq \Lambda$, if Λ and Ω are partitions of the same set A and every set $L \in \Lambda$ is contained in some set $O \in \Omega$:

$$\Omega \preceq \Lambda \quad :\Leftrightarrow \quad \underline{\Omega} = \underline{\Lambda}, \; \forall L \in \Lambda : \exists O \in \Omega : \, L \subseteq O. \tag{MA.13}$$

This means that the equivalence relation on $\underline{\Lambda}$ induced by Λ is finer than the one induced by Ω. For instance, the last partition in Equation (MA.11) is a refinement of all the preceding ones, and the first partition is coarser than all the others. The relation $\preceq$ defines a partial order on $\mathcal{P}(\alpha)$; for instance, the three middle partitions in (MA.11) cannot be compared among each other. In any case, if Λ is a partition of a set A with $|A| = j$, we have a finest and a coarsest partition of this set:

$$\underline{\Lambda} = A = \{a_1, \ldots, a_j\} \preceq \Lambda \preceq \overline{\Lambda} := \Big\{\{a_1\}, \ldots, \{a_j\}\Big\} = \Big\{\{a\} | \, a \in A\Big\}. \tag{MA.14}$$

If Λ is finer than Ω, we define, for $O \in \Omega$, the Λ-*induced partition of* O, by

$$O|\Lambda := \{L \in \Lambda | \, L \subset O\} \in \mathcal{P}(O). \tag{MA.15}$$

Another way of viewing refinements is via *partitions of partitions*: a partition Λ is a set and hence can again be partitioned. In fact, every partition Ω that is coarser than Λ defines a partition of the partition Λ via

$$\Lambda = \dot{\cup}_{O \in \Omega} \, O|\Lambda,$$

and in this way we get a canonical bijection between the set of all partitions Ω that are coarser than Λ and the set $\mathcal{P}(\mathcal{P}(\Lambda))$ of all partitions of the partition Λ.

MA.5. *Multilinear maps between cubes of* $\mathbb{K}$*-modules.* Given two cubes of $\mathbb{K}$-modules $(V_\alpha), (V'_\alpha)$ with total spaces E, E', for every partition Λ of an element $\alpha \in I_k$, we assume a $\mathbb{K}$-multilinear map

$$b^\Lambda : V_{\Lambda^1} \times \ldots \times V_{\Lambda^{\ell(\Lambda)}} \to V'_\alpha \tag{MA.16}$$

be given. Then a *multilinear map* or a *homomorphism of cubes of* $\mathbb{K}$*-modules* is a map of the form

$$f := f_b : E \to E', \quad v = \sum_{\alpha > 0} v_\alpha \mapsto \sum_{\alpha > 0} \sum_{\Lambda \in \mathcal{P}(\alpha)} b^\Lambda(v_{\Lambda^1}, \ldots, v_{\Lambda^{\ell(\Lambda)}}). \tag{MA.17}$$

Written out more explicitly,

$$f_b(v) = \sum_{\substack{\alpha \in I_k \\ \alpha > 0}} \sum_{\ell=1}^{|\alpha|} \sum_{\Lambda \in \mathcal{P}_\ell(\alpha)} b^\Lambda(v_{\Lambda^1}, \ldots, v_{\Lambda^\ell}).$$

To get a more concise notation, we let

$$V_\Lambda := V_{\Lambda^1} \times \ldots \times V_{\Lambda^\ell} \tag{MA.18}$$

and write $b^\Lambda : V_\Lambda \to V'_{\underline{\Lambda}}$. Note that $\Lambda = \{\Lambda^1, \ldots, \Lambda^\ell\}$ is a set, hence we will not distinguish between the various versions of V_Λ or b^Λ obtained by permuting the factors. If one whishes, one may single out the "ordered version" $\Lambda^1 < \ldots < \Lambda^\ell$, but as long as we distinguish the various spaces V_α by their index (even if they happen to be isomorphic), this will not be necessary. We write $\mathrm{Hom}(E, E')$ for the set of all multilinear maps $f = f_b : E \to E'$ and $\mathrm{End}(E) := \mathrm{Hom}(E, E)$. The family $b = (b^\lambda)_{\lambda \in \mathrm{Part}(I)}$ of multilinear maps is called a *multilinear family*, and we denote by $\mathrm{Mult}(E, E')$ the set of all multilinear families on E with values in E'. We introduce the following conditions on multilinear maps and on the corresponding multilinear families:

(reg) We say that $f_b : E \to E'$ is *regular* if, whenever $\ell(\Lambda) = 1$ (i.e. $\Lambda = \{\alpha\} = \{\Lambda^1\}$ is a trivial partition), the (linear) map $b^\Lambda : V_\alpha \to V'_\alpha$ is bijective. We denote by

$$\mathrm{Gm}^{0,k}(E) := \{f_b \in \mathrm{End}(E) | \, b \in \mathrm{Mult}(E), \, \ell(\Lambda) = 1 \Rightarrow b^\Lambda \in \mathrm{Gl}_{\mathbb{K}}(V_{\lambda^1})\}$$

the set of all regular endomorphisms of E.

(up) We say that an element $f_b \in \mathrm{Gm}^{0,k}(E)$ is *unipotent* if, whenever $\Lambda = \{\alpha\}$ is a trivial partition, we have $b^\Lambda = \mathrm{id}_{V_\alpha}$. We denote by

$$\mathrm{Gm}^{1,k}(E) := \{f_b \in \mathrm{Gm}^{0,k}(E) \mid b \in \mathrm{Mult}(E),\, \ell(\Lambda) = 1 \Rightarrow b^\Lambda = \mathrm{id}_{V_{\lambda^1}}\}$$

the set of all unipotent endomorphisms of E.

(si) Let $1 < j \leq k$. We say that f_b *is singular of order* j *(at the origin)* or just: f_b *is of order* j if $b^\Lambda = 0$ whenever $\ell(\Lambda) \leq j$, and we denote by $\mathrm{Mult}_{>j}(E)$ the set of all multilinear families on E that are singular of order j.

(ups) For $1 < j \leq k$, we denote by

$$\mathrm{Gm}^{j,k}(E) := \{f_b \in \mathrm{Gm}^{1,k}(E) \mid \ell(\Lambda) = 2, \ldots, j \Rightarrow b^\Lambda = 0\}$$

the set of all unipotent endomorphisms that are of the form $f = \mathrm{id}_E + h_b$ with h_b singular of order j.

Theorem MA.6.

(1) *Cubes of* $\mathbb{K}$*-modules together with their homomorphisms form a category (which we call the* category of k-multilinear spaces*).*

(2) *A homomorphism* $f_b \in \mathrm{Hom}(E, E')$ *is invertible if and only if it is regular, and then* $(f_b)^{-1}$ *is again a homomorphism. In particular,* $\mathrm{Gm}^{0,k}(E)$ *is a group, namely it is the automorphism group of* E *in the category of* k*-multilinear spaces.*

(3) *We have a chain of subgroups*

$$\mathrm{Gm}^{0,k}(E) \supset \mathrm{Gm}^{1,k}(E) \supset \ldots \supset \mathrm{Gm}^{k-1,k}(E) \supset \mathrm{Gm}^{k,k}(E) = \{\mathrm{id}_E\}.$$

Proof. (1) It is fairly obvious that the composition of homomorphisms is again a homomorphism. However, let us, for later use, spell out the composition rule in more detail: assume $g : E' \to E''$ and $f : E \to E'$ are multilinear maps; for simplicity we will denote the multilinear families corresponding to g and to f by $(g^\Lambda)_\Lambda$ and $(f^\Lambda)_\Lambda$. Then the composition rule is obtained simply by using multilinearity and ordering terms according to the space in which we end up: let $w_\alpha = \sum_{\nu \in \mathcal{P}(\alpha)} f^\nu(v_{\nu_1}, \ldots, v_{\nu_\ell})$, then

$$\begin{aligned}
g(f(v)) &= g(\sum_\alpha w_\alpha) \\
&= \sum_\alpha \sum_{\Omega \in \mathcal{P}(\alpha)} g^\Omega(w_{\Omega^1}, \ldots, w_{\Omega^r}) \\
&= \sum_\alpha \sum_{\Omega \in \mathcal{P}(\alpha)} \sum_{\substack{\Lambda_1 \in \mathcal{P}(\Omega^1) \ldots \\ \Lambda_r \in \mathcal{P}(\Omega^r)}} g^\Omega\Big(f^{\Lambda_1}(v_{\Omega^1_1}, \ldots, v_{\Omega^1_{\ell_1}}), \ldots, f^{\Lambda_r}(v_{\Omega^r_1}, \ldots, v_{\Omega^r_{\ell_r}})\Big) \\
&= \sum_\alpha \sum_{\Lambda \in \mathcal{P}(\alpha)} \sum_{\Omega \preceq \Lambda} g^\Omega\Big(f^{\Omega^1|\Lambda}(v_{\Omega^1_1}, \ldots, v_{\Omega^1_{\ell_1}}), \ldots, f^{\Omega^r|\Lambda}(v_{\Omega^r_1}, \ldots, v_{\Omega^r_{\ell_r}})\Big)
\end{aligned}$$

where $r = \ell(\Omega)$ and $\ell_j = \ell(\Lambda_j)$ and $\Lambda_j = \Omega^j|\Lambda$ (cf. Eqn. (MA.15); this implies $\Lambda = \cup_{j=1}^r \Lambda_j$). It follows that $g \circ f$ corresponds to the multilinear family

$$(g \circ f)^\Lambda(v_\Lambda) = \sum_{\Omega \preceq \Lambda} g^\Omega\Big(f^{\Omega^1|\Lambda}(v_{\Omega^1_1}, \ldots, v_{\Omega^1_{\ell_1}}), \ldots, f^{\Omega^r|\Lambda}(v_{\Omega^r_1}, \ldots, v_{\Omega^r_{\ell_r}})\Big). \quad \text{(MA.19)}$$

Finally, the identity map clearly is a homomorphism, and hence we have defined a category.

(2) If $f_b : E \to E'$ is a homomorphism, then the definition of f_b shows that f_b maps the α-axis into the α-axis: $f_b(V_\alpha) \subset V'_\alpha$, and the induced map $V_\alpha \to V'_\alpha$ agrees with $b^\alpha := b^\Lambda$ for $\Lambda = \{\alpha\}$. Therefore, if f_b is bijective, then so is b^Λ whenever $l(\Lambda) = 1$, and thus f_b is regular. In order to prove the converse, we define, for every homomorphism $f = f^b : E \to E'$, the map

$$T_0 f := \times_\alpha b^\alpha : E \to E', \quad \sum_\alpha v_\alpha \mapsto \sum_\alpha b^\alpha(v_\alpha), \qquad \text{(MA.20)}$$

where as above $b^\alpha := b^{\{\alpha\}}$. This is again a homomorphism, acting on each axis in the same way as f. From the definition of multilinear maps one immediately gets the functorial rules $T_0(g \circ f) = T_0 g \circ T_0 f$ and $T_0(T_0 f) = T_0 f$. Now assume that $f := f_b$ is regular. This is equivalent to saying that $T_0 f$ is bijective. Then clearly $(T_0 f)^{-1}$ is again a homomorphism and $T_0(T_0 f)^{-1} = (T_0 f)^{-1}$. Let $g := (T_0 f)^{-1} \circ f : E \to E$; then g is an endomorphism with $T_0 g = \mathrm{id}_E$, i.e. g is regular and moreover unipotent, i.e. $g \in \mathrm{Gm}^{1,k}(E)$. Thus we are done if we can show that $\mathrm{Gm}^{1,k}(E)$ is a group.

Let us, more generally, prove by a descending induction on $j = k, k-1, \ldots, 1$ that all $\mathrm{Gm}^{j,k}(E)$ are groups: it is clear that $\mathrm{Gm}^{k,k}(E) = \{\mathrm{id}_E\}$ is a group. For the induction step with $j \geq 1$, note first that $\mathrm{Gm}^{j,k}(E)$ is stable under composition. We need the lowest term of the composition of elements $f_a, f_b \in \mathrm{Gm}^{j,k}(E)$: define c by $f_c = f_b \circ f_a$; then we have $c^\lambda = a^\lambda + b^\lambda$ whenever $l(\lambda) = j+1$. It follows that $g := f_{-b} \circ f_b \in \mathrm{Gm}^{j+1,k}(E)$ for all $f_b \in \mathrm{Gm}^{j,k}(E)$. But by induction $\mathrm{Gm}^{j+1,k}(E)$ is a group, and hence f_b is invertible with inverse $g^{-1} \circ f_{-b} \in \mathrm{Gm}^{j,k}(E)$. (In particular, this argument shows that $\mathrm{Gm}^{k-1,k}(E)$ is a vector group.) Finally, (3) has just been proved. ■

As seen in the proof, (2) implies that $f : E \to E'$ is invertible if and only if $T_0 f$ is invertible. If we interpret $T_0 f$ as the "total differential of f at the origin", then this may be seen as a kind of *inverse function theorem for homomorphisms of multilinear spaces.*

MA.7. *The general and the special multilinear group.* The automorphism group $\mathrm{Gm}^{0,k}(E)$ is called the *general multilinear group of* E, and the group $\mathrm{Gm}^{1,k}(E)$ is called the *special multilinear group of* E. The latter acts "simply non-linearly" on E in the sense that the only group element that acts linearly on E is the identity. Therefore it plays a more important rôle in our theory than the full automorphism group. The preceding proof shows that we have a (splitting) exact sequence of groups

$$1 \;\to\; \mathrm{Gm}^{1,k}(E) \;\to\; \mathrm{Gm}^{0,k}(E) \;\overset{T_0}{\to}\; \times_{\alpha>0}\,\mathrm{Gl}(V_\alpha) \;\to\; 1 \qquad \text{(MA.21)}$$

showing that $\mathrm{Gm}^{1,k}(E)$ is normal in the automorphism group, and the quotient is a product of ordinary linear groups. For more information on the structure of $\mathrm{Gm}^{1,k}(E)$, see Sections MA.12 – MA.15. and Chapter PG.

MA.8. *Example: the case* $k=3$. For $k=3$, elements of $\mathrm{Gm}^{1,3}(E)$ are described by six bilinear maps and one trilinear map:

$$b = (b^{001,010}, b^{001,100}, b^{010,100}, b^{001,110}, b^{010,101}, b^{011,100}, b^{001,010,100}),$$

and thus

$$\begin{aligned} f_b(v) = v + b^{001,010}(v_{001}, v_{010}) + b^{001,100}(v_{001}, v_{100}) + b^{010,100}(v_{010}, v_{100}) + \\ b^{001,110}(v_{001}, v_{110}) + b^{010,101}(v_{010}, v_{010}) + b^{011,100}(v_{011}, v_{100}) + \\ b^{001,010,100}(v_{001}, v_{010}, v_{100}), \end{aligned}$$

where the term in the second and third line describes the component in V_{111}. Composition $f^b \circ f^a = f^c$ of such maps is described by the following formulae: if $\ell(\lambda) = 2$, then we have simply $c^\lambda = b^\lambda + a^\lambda$. If $\ell(\lambda) = 3$, i.e., $\lambda = ((001),(010),(100))$, then there are three non-trivial partitions Ω that are strictly coarser than Λ, and we have:

$$\begin{aligned} c^\lambda(u,v,w) = a^\lambda(u,v,w) + b^\lambda(u,v,w) + \\ b^{001,110}(u, a^{010,100}(v,w)) + b^{010,101}(v, a^{001,100}(u,w)) + b^{011,100}(a^{001,010}(u,v), w). \end{aligned}$$

Inversion of a unipotent $f = f_b$ is described as follows: if $\ell(\lambda) = 2$, then $(f^{-1})^\lambda = -f^\lambda$, and the coefficient belonging to $\lambda = ((100)(010)(001))$ is

$$\begin{aligned} (f^{-1})_\lambda(u,v,w) = -b^\lambda(u,v,w) + \\ b^{001,110}(u, b^{010,100}(v,w)) + b^{010,101}(v, b^{001,100}(u,w)) + b^{011,100}(b^{001,010}(u,v), w). \end{aligned}$$

MA.9. *Multilinear connections and the category of multilinear spaces.* Since $\mathrm{Gm}^{1,k}(E)$ acts non-linearly on E, a multilinear space E does not have any distinguished linear structure. We say that two linear structures on E are *k-multilinearly related* if they are conjugate to each other under the action of $\mathrm{Gm}^{1,k}(E)$, and we say that a linear structure on E is a *(multilinear) connection on E* if it is k-multilinearly related to the initial $\mathbb{K}$-module structure L^0 defined by the bijection $E \cong \oplus_{\alpha>0} V_\alpha$. As in Section BA.1, we denote by $L^b = (E, +_b, \cdot_b)$ the push-forward of the original linear structure on E via f_b; then $\mathrm{Gm}^{1,k}(E)$ acts simply transitively on the space of all multilinear connections via $(f^c, L^b) \mapsto (f^c \circ f^b) \cdot L^0$. Recall that, for $k=2$, $\mathrm{Gm}^{1,k}(E)$ is a vector group, and hence in this case the space of connections is an affine space over $\mathbb{K}$. For general k, the explicit formulae for addition $+_b$ and multiplication by scalars $\cdot_b$ in the $\mathbb{K}$-module (E, L^b) are much more complicated than the formulae given in Section BA.2 for the case $k=2$, and we will not try to write them out in full detail. Let us just make the following remarks. It is easy to write the sum of elements taken from axes: since f_b acts trivially on each axis, this is the "b-sum"

$$(b\sum)_\alpha v_\alpha := f_b(\sum_\alpha v_\alpha). \tag{MA.22}$$

Moreover, in the sum $v +_b v'$ for arbitrary elements $v, v' \in E$, components with $|\alpha| = 1$ are simply of the form $v_\alpha + v'_\alpha$; components with $|\alpha| = 2$ are of the form $v_\alpha + v'_\alpha + b^{\beta,\gamma}(v_\beta, v'_\gamma) + b^{\beta,\gamma}(v'_\beta, v_\gamma)$ where $\alpha = \beta + \gamma$ is the unique decomposition with $|\beta| = |\gamma| = 1$ and $\beta < \gamma$.

Intrinsic geometric structures

We say that a structure on E is *intrinsic* if it is invariant under the action of the general multilinear group $\mathrm{Gm}^{0,k}(E)$.

MA.10. *Intrinsic structures: subgeometries.* The group $\mathrm{Gm}^{0,k}(E)$ preserves axes and acts linearly on them; hence the axes are intrinsic subspaces, not only as sets, but also as sets with their induced $\mathbb{K}$-module structure. For any $\alpha \in I_k$, the set $\widetilde{V}_\alpha := \bigoplus_{\beta \subseteq \alpha} V_\alpha$ (cf. Eqn. (MA.5)) is invariant under the action of $\mathrm{Gm}^{1,k}(E)$ and hence is an intrinsic subspace of E; but the action of $\mathrm{Gm}^{1,k}(E)$ on these subspaces is in general not trivial. For $k = 3$, diagram (MA.6) describes the cube of these intrinsic subspaces. Note that, if $|\alpha| = j$, then $\widetilde{V}_\alpha$ is a j-linear space in its own right which is represented by the j-dimensional subcube having bottom point 0 and top point $\widetilde{V}_\alpha$. For instance, if $k = 3$, the three bottom faces of the cube represent three bilinear subspaces of E. In a similar way, if $\alpha, \beta, \alpha + \beta \in I$, then the direct sum $V_\alpha \oplus V_\beta \oplus V_{\alpha+\beta}$ is an invariant bilinear subspace of E. More generally, for any partition Λ, we can define an invariant subspace E_Λ which is a cube of $\mathbb{K}$-modules with bottom $\Lambda^1, \ldots, \Lambda^\ell$.

MA.11. *Intrinsic structures: projections, hyperplanes.* Let

$$p_\beta : E \to \widetilde{V}_\beta, \quad \sum_\alpha v_\alpha \mapsto \sum_{\alpha \subset \beta} v_\alpha \tag{MA.23}$$

be the projection onto the invariant subspace $\widetilde{V}_\beta$. Then a straightforward check shows that p_β is intrinsic in the sense that $p_\beta \circ f_b = f_b \circ p_\beta$ for all $f_b \in \mathrm{Gm}^{1,k}(E)$:

$$\begin{array}{ccc} E & \xrightarrow{f_b} & E \\ p_\beta \downarrow & & p_\beta \downarrow \\ \widetilde{V}_\beta & \xrightarrow{f_b} & \widetilde{V}_\beta \end{array} \tag{MA.24}$$

In particular, the factors of the decomposition

$$E = \widetilde{V}_\beta \oplus K_\beta, \qquad K_\beta := \ker(p_\beta) = \bigoplus_{\alpha \not\subseteq \beta} V_\alpha = \sum_{\alpha \perp \beta} H_\alpha$$

are invariant under $\mathrm{Gm}^{1,k}(E)$; but of course $\mathrm{Gm}^{1,k}(E)$ does in general not act trivially on the factors, nor are the "parallels" of $\widetilde{V}_\beta$ and K_β invariant under $\mathrm{Gm}^{1,k}(E)$. We say that the subgeometry $\widetilde{V}_\beta$ is *maximal* if $|\beta| = k-1$ and *minimal* if $|\beta| = 1$. (For $k = 2$ every proper subgeometry is both maximal and minimal.)

(1) Maximal subgeometries: fix $j \in \{1, \ldots, k\}$ and let $\beta := e_j^* := \sum_{i \neq j} e_i$ correspond to the complement $\mathbb{N}_k \setminus \{j\}$. The corresponding projection is

$$p_{e_j^*} : E \to \widetilde{V}_{(1\ldots101\ldots1)} = \bigoplus_{\substack{\alpha \in I \\ \alpha_j = 0}} V_\alpha, \quad \sum_{\alpha > 0} v_\alpha \mapsto \sum_{\alpha_j = 0} v_\alpha. \tag{MA.25}$$

The kernel is called the j-*th hyperplane*:

$$H_j := K_{e_j^*} = \ker(p_j) = \bigoplus_{\substack{\alpha\in I\\ \alpha_j=1}} V_\alpha = H_{e_j} \qquad \text{(MA.26)}$$

We claim that $\mathrm{Gm}^{1,k}(E)$ acts trivially on H_j. Indeed, let $v = \sum_{\alpha_j=1} v_\alpha \in H_{e_j}$. Then $f_b(v) = v + \sum_\lambda b^\lambda(v_{\lambda^1}, \ldots, v_{\lambda^\ell}) = v$ since $\lambda^i_j = 1$ for $i = 1, \ldots, \ell$ and hence the λ^i cannot form a non-trivial partition and the sum over λ has to be empty. Hence H_{e_j} together with its linear structure is invariant under f_b and thus it is an intrinsic subspace of E. In the same way we see that all fibers $y + H_{e_j}$, for any $y \in E$, are intrinsic affine subspaces of E. There is a distinguished origin in $y + H_{e_j}$, namely $\sum_{\alpha_j=0} y_\alpha$, and hence $y + H_{e_j}$ carries an intrinsic linear structure. (In the context of higher order tangent bundles T^kM, these hyperplanes correspond to tangent spaces $T_u(T^{k-1}M)$ contained in $E = (T^kM)_x$.)

(2) Minimal subgeometries: $\beta = e_j$. Then the projection onto $\widetilde{V}_{e_j} = V_{e_j} =: V_j$ is

$$\mathrm{pr}_j := p_{e_j} : E \to V_j, \quad v = \sum v_\alpha \mapsto v_{e_j}. \qquad \text{(MA.27)}$$

The group $\mathrm{Gm}^{1,k}(E)$ acts trivially on the axes V_j and hence $p_{e_j} \circ f_b = p_{e_j}$; thus not only the kernel, but all fibers of p_{e_j} are $\mathrm{Gm}^{1,k}(E)$-invariant subspaces.

(3) Intersections. Clearly, intersections of intrinsic subspaces are again intrinsic subspaces. For instance, if $i \neq j$,

$$H_{ij} := H_i \cap H_j = \ker(p_{e_i^*} \times p_{e_j^*}) = \bigoplus_{\substack{\alpha\in I\\ \alpha_i=1=\alpha_j}} V_\alpha = H_{e_i+e_j} \qquad \text{(MA.28)}$$

is an intrinsic subspace, and so is the "vertical subspace" $\bigoplus_{\alpha:|\alpha|\geq 2} V_\alpha$, kernel of the projection

$$\mathrm{pr} := \mathrm{pr}_1 \times \ldots \times \mathrm{pr}_k : E \to V_1 \times \ldots \times V_k, \quad v \mapsto (\mathrm{pr}_1(v), \ldots, \mathrm{pr}_k(v)). \quad \text{(MA.29)}$$

Summing up, multilinear spaces have a non-trivial "intrinsic incidence geometry", and it seems that the complementation map of the lattice I_k corresponds to some duality of this incidence geometry. It should be interesting to develop this topic more systematically, in particular in relation with Lie groups and symmetric spaces.

Complements on the structure of the general multilinear group

MA.12. *"Matrix coefficients" and "matrix multiplication".* The choice of a multilinear connection in a multilinear space should be seen as an analog of choosing a base in a free (say, n-dimensional) $\mathbb{K}$-module V, and the analog of the map $\mathbb{K}^n \to V$ induced by a basis is the map

$$\Phi^b : \bigoplus_{\alpha>0} V_\alpha \to E, \quad v = (v_\alpha)_\alpha \mapsto f_b(v) \qquad \text{(MA.30)}$$

(in differential geometry, Φ^b is the *Dombrowski splitting* or *linearization map*, cf. Theorem 16.3). Then the "matrix" $D^b_c(f)$ of a homomorphism $f : E \to E'$ with respect to linear structures L^b and L^c is defined by

$$\begin{array}{ccc} \bigoplus_{\alpha>0} V_\alpha & \overset{\Phi^b}{\to} & E \\ D^b_c(f)\downarrow & & \downarrow f \\ \bigoplus_{\alpha>0} V'_\alpha & \overset{\Phi^c}{\to} & E' \end{array} \qquad \text{(MA.31)}$$

We have the usual "matrix multiplication rules" $D^a_c(g \circ f) = D^b_c(g) \circ D^a_b(f)$, $D^a_b(\mathrm{id}) = \mathrm{id}$. Just as the individual matrix coefficients of a linear map f with respect to a basis are defined by $a_{ij} := \mathrm{pr}_i \circ f \circ \iota_j$, where pr_i and ι_j are the projections, resp. injections associated with the basis vectors, we define for a homomorphism $f : E \to E'$, with respect to fixed connections in E and in E',

$$f^{\Omega|\Lambda} := \mathrm{pr}_\Omega \circ f \circ \iota_\Lambda : V_\Lambda \to V'_\Omega, \qquad \begin{array}{ccc} E & \overset{f}{\to} & E' \\ \iota_\Lambda \uparrow & & \downarrow \mathrm{pr}_\Omega \\ V_\Lambda & \overset{f^{\Omega|\Lambda}}{\to} & V'_\Omega \end{array} \qquad \text{(MA.32)}$$

where $\iota_\Lambda : V_\Lambda \to E$ is inclusion and $\mathrm{pr}_\Omega : E' \to V'_\Omega$ is projection with respect to the fixed connections. Not all matrix coefficients are interesting – we are mainly interested in the following:

(1) We say that a matrix coefficient $f^{\Omega|\Lambda}$ is *effective* if $\Omega \preceq \Lambda$. In this case let us write $\Omega = \{\Omega^1, \ldots, \Omega^r\}$ if $\ell(\Omega) = r$, $r \leq \ell(\Lambda)$, $\Omega^i = \{\Omega^i_1, \ldots, \Omega^i_{\ell_i}\}$. Then the effective matrix coefficient is explicitly given by

$$\begin{aligned} f^{\Omega|\Lambda} &: V_\Lambda \to V_\Omega = V_{\Omega^1} \times \ldots \times V_{\Omega^r}, \\ (v_{\Lambda^1}, \ldots, v_{\Lambda^\ell}) &\mapsto (b^{\Omega^1|\Lambda}(v_{\Omega^1_1}, \ldots, v_{\Omega^1_{\ell_1}}), \ldots, b^{\Omega^r|\Lambda}(v_{\Omega^r_1}, \ldots, v_{\Omega^r_{\ell_r}})) \end{aligned} \qquad \text{(MA.33)}$$

(2) We say that a matrix coefficient $f^{\Omega|\Lambda}$ is *elementary* if $\ell(\Omega) = 1$ and $\Omega = \underline{\Lambda}$. (Since $\underline{\Lambda} \preceq \Lambda$, elementary coefficients are effective.) With respect to the linear structures $L^0, (L^0)'$, the elementary matrix coefficients are just the multilinear maps $b^\Lambda : V_\Lambda \to V'_{\underline{\Lambda}}$ defining $f = f^b$:

$$b^\Lambda = f^{\underline{\Lambda}|\Lambda} : V_\Lambda \to V'_{\underline{\Lambda}} \qquad \begin{array}{ccc} E & \overset{f_b}{\to} & E' \\ \iota_\Lambda \uparrow & & \downarrow \mathrm{pr}_\alpha \\ V_\lambda & \overset{b^\Lambda}{\to} & V'_\alpha \end{array} \qquad \text{(MA.34)}$$

where $\alpha = \underline{\Lambda}$. In this situation, we may also use the notation $f^{\Omega|\Lambda} =: b^{\Omega|\Lambda}$, $f^\Lambda := b^\Lambda$.

Proposition MA.13. ("Matrix multiplication.") *Assume E, E', E'' are multilinear spaces with fixed multilinear connections and $g : E' \to E''$ and $f : E \to E'$ are homomorphisms. Let $\Omega \preceq \Lambda$. Then the effective matrix coefficient $(g \circ f)^{\Omega|\Lambda}$ is given by*

$$(g \circ f)^{\Omega|\Lambda} = \sum_{\Omega \preceq \nu \preceq \Lambda} g^{\Omega|\nu} \circ f^{\nu|\Lambda}$$

In particular, for elements of the group $\mathrm{Gm}^{1,k}(E)$ *we have the composition rule*

$$(g \circ f)^\Lambda = g^\Lambda + f^\Lambda + \sum_{r=2}^{\ell(\Lambda)-1} \sum_{\substack{\Omega \prec \Lambda \\ \ell(\Omega)=r}} g^\Omega \circ f^{\Omega|\Lambda}.$$

Proof. Using the terminology of matrix coefficients, Formula (MA.19) from the proof of Theorem MA.6 can be re-written

$$(g \circ f)^\Lambda = \sum_{\Omega \preceq \Lambda} g^\Omega \circ f^{\Omega|\Lambda} = \sum_{r=1}^{\ell(\Lambda)} \sum_{\substack{\Omega \preceq \Lambda \\ \ell(\Omega)=r}} g^\Omega \circ f^{\Omega|\Lambda}. \qquad \text{(MA.35)}$$

The proposition follows from this formula. ∎

Corollary MA.14. *The subgroup* $\mathrm{Gm}^{k-1,k}(E) \subset \mathrm{Gm}^{1,k}(E)$ *is a central vector group. More precisely, if* $f \in \mathrm{Gm}^{k-1,k}(E)$, $g \in \mathrm{Gm}^{1,k}(E)$, *then*

$$(f \circ g)^\Lambda = f^\Lambda + g^\Lambda = (g \circ f)^\Lambda.$$

Proof. Note that f belongs to $\mathrm{Gm}^{k-1,k}(E)$ if and only if: $f^\Lambda = 0$ whenever $\ell(\Lambda) \in \{2, \ldots, k-1\}$. Thus the sum in the composition rule reduces to two terms, and we easily get the claim. ∎

Under suitable assumptions, one can show that $\mathrm{Gm}^{k-1,k}(E)$ is precisely the center of $\mathrm{Gm}^{1,k}(E)$. The cosets of $\mathrm{Gm}^{k-1,k}(E)$ in $\mathrm{Gm}^{1,k}(E)$ are easily described: $f_c = f_a \circ f_b$ with $f_b \in \mathrm{Gm}^{k-1,k}(E)$ iff

$$\ell(\Lambda) < k \Rightarrow c^\Lambda = a^\Lambda, \qquad \ell(\Lambda) = k \Rightarrow c^\Lambda = a^\Lambda + h$$

with some k-multilinear map $h : V^k \to V$.

Intrinsic multilinear maps

MA.15. *Intrinsic multilinear maps.* Let V' be an arbitrary $\mathbb{K}$-module and E a k-linear space. A map $f : E \to V'$ is called *(intrinsically) multilinear* if, for all $y \in E$ and $j = 1, \ldots, k$, the restrictions to the intrinsic affine spaces (hyperplanes) $f|_{y+H_j} : y + H_j \to V'$ are $\mathbb{K}$-linear. We say that f is *homogeneous multilinear* if, for all $i \neq j$, the restriction of f to the affine spaces $y + H_{ij}$ is constant (for notation, cf. Eqns. (MA.26), (MA.28)). We denote by $\mathcal{M}(E, V')$, resp. by $\mathcal{M}_h(E, V')$ the space of (homogeneous) intrinsically multilinear maps from E to $\mathbb{K}$ and by $\mathrm{Hom}(V_1, \ldots, V_k; V')$ the space of multilinear maps $V_1 \times \ldots \times V_k \to V'$ in the usual sense (over $\mathbb{K}$).

Proposition MA.16.

(1) *A map $f : E \to V'$ is intrinsically multilinear if and only if it can be written, with respect to a fixed linear structure $L = L^0$, in the form*

$$f(\sum_\alpha v_\alpha) = \sum_{\ell=1}^{k} \sum_{\Lambda \in \mathcal{P}_\ell(\mathbb{N}_k)} b^\Lambda(v_{\Lambda^1}, \ldots, v_{\Lambda^\ell}).$$

with multilinear maps (in the ordinary sense) $b^\Lambda : V_\Lambda \to V'$. Then f is homogeneous if and only if the only non-vanishing term belongs to $\ell = k$, i.e.,

$$f(\sum_\alpha v_\alpha) = b(v_{e_1}, \ldots, v_{e_k})$$

with multinear $b : V_1 \times \ldots \times V_k \to V'$.

(2) *The map*

$$\mathrm{Hom}(V_1, \ldots, V_k; V') \to \mathcal{M}_h(E, V'), \quad f \mapsto \overline{f} := f \circ \mathrm{pr}$$

is a well-defined bijection.

Proof. (1) Assume first that f is of the form given in the claim. We decompose $v = \sum_{\alpha_j=1} v_\alpha + \sum_{\alpha_j=0} v_\alpha = u + w$ so that the first term u belongs to H_j. Keeping the second term constant, the dependence on the first term is linear: in fact, since Λ is a partition of the full set $\mathbb{N}_k$, each term $b^\Lambda(v_{\Lambda^1}, \ldots, v_{\Lambda^\ell})$ involves exactly one argument $v_{\Lambda^i} \in H_j$, and hence depends linearly on this argument.

Conversely, assume $f : E \to V'$ is intrinsically multilinear. Then we prove by induction on k that f is of the form given in the claim. For $k = 1$ the claim is trivial, and for $k = 2$ see Prop. BA.7. By induction, the restriction of f to all maximal subgeometries $\widetilde{V}_{(1\ldots101\ldots1)}$ (which are $k-1$-cubes of $\mathbb{K}$-modules, see Section MA.11) is given by the formula from the claim. Thus, decomposing $v = u + w \in E$ as above, we write $f(v) = f_u(w)$, apply the induction hypothesis to f_u with coefficients depending on u. But the dependence on u must be linear, and this yields the claim for f.

(2) This is a direct consequence of (1) since $b = f \circ \mathrm{pr}$. ■

In order to define an inverse of the map $f \mapsto \underline{f}$, we fix a linear structure L^b and let

$$\iota^b : V_1 \times \ldots \times V_k \to E, \quad (v_{e_1}, \ldots, v_{e_k}) \mapsto (b \sum) v_{e_i}$$

be the inclusion map for the longest partition (sum with respect to the linear structure L^b). Then the inverse map is given by

$$\mathcal{M}_h(E, V') \to \mathrm{Hom}(V_1, \ldots, V_k; V'), \quad f \mapsto \underline{f} := f \circ \iota^b.$$

Summing up, the following diagram encodes two different ways of seeing the same object f:

$$\begin{array}{ccc} E & & \\ \mathrm{pr} \downarrow\uparrow \iota^b & \overset{\overline{f}}{\searrow} & \\ V_1 \times \ldots \times V_k & \overset{\underline{f}}{\to} & V' \end{array} \tag{MA.36}$$

Note that the map $f \mapsto \underline{f}$ can be extended to a map

$$\mathcal{M}(E,V') \to \mathrm{Hom}(V_1,\ldots,V_k;V'), \quad f \mapsto \underline{f} := f \circ \iota^b.$$

which does depend on L^b. Finally, as in Section BA.9, in case that all V_α are canonically isomorphic (see next chapter), one can define a skew-symetrization operator from multilinear maps to homogeneous ones.

Tensor space and tensor functor

MA.17. *The tensor space.* For a partition Λ we let

$$\widehat{V}_\Lambda := V_{\Lambda^1} \otimes \ldots \otimes V_{\Lambda^\ell}$$

and we define the *(big) tensor space of* E

$$\widehat{E} := \bigoplus_{\Lambda \in \mathrm{Part}(I)} \widehat{V}_\Lambda = \bigoplus_{\alpha>0} \bigoplus_{\ell=1}^{|\alpha|} \bigoplus_{\Lambda \in \mathcal{P}_\ell(\alpha)} \widehat{V}_\Lambda.$$

There is a canonical map

$$E \to \widehat{E}, \quad (v_\alpha)_\alpha \mapsto \sum_\Lambda v_{\Lambda^1} \otimes \ldots \otimes v_{\Lambda^\ell}.$$

Next, to a homomorphism $f : E \to E'$ we associate a linear map $\widehat{f} : \widehat{E} \to \widehat{E}'$ in the following way: to every multilinear map $b^\Lambda : V_\Lambda \to V_\alpha$, using the universal property of tensor products, we associate a linear map $\widehat{b}^\Lambda : \widehat{V}_\Lambda \to V_\alpha$. More generally, whenever $\Omega \preceq \Lambda$, to the matrix coefficient $f^{\Omega|\Lambda} : V_\Lambda \to V_\Omega$ we associate a linear map $\widehat{f}^{\Omega|\Lambda} : \widehat{V}_\Lambda \to \widehat{V}_\Omega$: we can write

$$f^{\Omega|\Lambda} = b^{\Omega^1|\Lambda} \times \ldots \times b^{\Omega^r|\Lambda} : \times V_{\Omega^i|\Lambda} \to \times V_{\underline{\Omega^i|\Lambda}}$$

and define

$$\widehat{f}^{\Omega|\Lambda} := \widehat{b}^{\Omega^1|\Lambda} \otimes \ldots \otimes \widehat{b}^{\Omega^r|\Lambda} : \otimes V_{\Omega^i|\Lambda} \to \otimes V_{\underline{\Omega^i|\Lambda}}.$$

We define all other components of $\widehat{f} : \widehat{E} \to \widehat{E}$ to be zero. In other words, we simply forget all non-effective matrix coefficients; since f is determined by the elementary coefficients, this is no loss of information.

Proposition MA.18. *The preceding construction is functorial:*

$$\widehat{\mathrm{id}}_E = \mathrm{id}_{\widehat{E}}, \qquad \widehat{g \circ f} = \widehat{g} \circ \widehat{f}.$$

Proof. The first equality is trivial. The second one follows from the matrix multiplication rule MA.13: the component $(\widehat{g} \circ \widehat{f})^{\Omega|\Lambda}$ of $\widehat{g} \circ \widehat{f}$ from $\widehat{V}_\Lambda$ to $\widehat{V}_\Omega$ is calculated by usual matrix multiplication:

$$(\widehat{g} \circ \widehat{f})^{\Omega|\Lambda} = \sum_\nu \widehat{g}^{\Omega|\nu} \circ \widehat{f}^{\nu|\Lambda},$$

where the sum is over all partitions ν; but, by the preceding definitions, only terms with $\Omega \preceq \nu \preceq \Lambda$ really contribute to the sum, and hence, by MA.13, we get the component $(\widehat{g \circ f})^{\Omega|\Lambda}$. ■

We call the functor $E \to \widehat{E}$ the *tensor functor*. This functor is an equivalence of categories: f can be recovered from $\widehat{f}$ since all $\widehat{b}^\Lambda$ are components of $\widehat{f}$. Finally, let us add the remark that in the special case of E of tangent or bundle type (see next chapter), we may define a smaller tensor space, essentially by taking fixed subspaces of the permutation group, which is functorial with respect to totally symmetric homomorphisms; it is this smaller tensor space that plays the role of the "higher osculating bundle of a vector bundle" in [Po62].

The affine-multilinear group

MA.19. Just as the usual general (or special) *linear* group can be extended to the usual general (or special) *affine* group, the groups $\mathrm{Gm}^{j,k}(E)$ may be extended to the corresponding *affine multilinear groups* $\mathrm{Am}^{j,k}(E)$. We give the basic definitions and leave details to the reader: assume that for every $\alpha \in I_k$ and $\beta \subseteq \alpha$ and $\Lambda \in \mathcal{P}(\beta)$, a multilinear map

$$C^{\Lambda;\alpha} : V_{\Lambda^1} \times \ldots \times V_{\Lambda^\ell} \to V_\alpha$$

is given. (Note that β is recovered from Λ as $\beta = \underline{\Lambda}$ and hence can be suppressed in the notation.) We say that $f : E \to E$ is *affine-multilinear* if it is of the form

$$f(\sum_\alpha v_\alpha) = \sum_\alpha \sum_{\beta \subset \alpha} \sum_{\Lambda \in \mathcal{P}(\beta)} C^{\Lambda;\alpha}(v_{\Lambda^1}, \ldots, v_{\Lambda^\ell})$$

for some family $(C^{\Lambda;\alpha})$. Note that for $\beta = \emptyset$, $C^{\Lambda;\alpha}$ is a constant belonging to V_α; in particular, for $k = 1$, we get the usual affine group of $E = V_1$. For $\beta = \alpha$, $C^{\Lambda;\alpha} =: b^\Lambda$ contributes to the "multilinear part" of f which is defined by

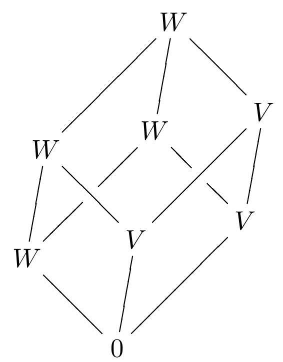
$$F : E \to E, \quad \sum_\alpha v_\alpha \mapsto \sum_\alpha \sum_{\Lambda \in \mathcal{P}(\alpha)} b^\Lambda(v_{\Lambda^1}, \ldots, v_{\Lambda^\ell}).$$

Clearly, F is a multilinear map in the sense of Section MA.5. We say that f is *regular* (resp. *unipotent*) if so is F. Then the following analog of Theorem MA.6 (2) holds: *An affine-multilinear map f is invertible iff it is regular, and then the inverse of f is again affine-multilinear.* (The proof is left to the reader.) Thus we have defined groups $\mathrm{Am}^{0,k}(E)$, resp. $\mathrm{Am}^{1,k}(E)$, of regular (resp. unipotent) affine-multilinear maps. The structure of these groups is rather interesting and will be investigated elsewhere.

SA. Symmetric and shift invariant multilinear algebra

Fibers of higher order tangent bundles are cubes of $\mathbb{K}$-modules having the additional property that all or at least some of the V_α are canonically isomorphic as $\mathbb{K}$-modules; for this reason we call such tubes *of tangent type*. This situation gives rise to interesting new symmetries: the main actors are the *permutation operators* which act horizontally, on the rows of the cube, like the castle in a game of chess, and the *shift operators* which act diagonally, like the bishops. The commutants in the general multilinear group of these actions are the *very symmetric*, resp. the *shift-invariant* subgroups. The intersection of these two groups is (in characteristic zero) the unit group of a semigroup of truncated polynomials on V.

SA.1. *Cubes of tangent type and of bundle type.* We say that a cube of $\mathbb{K}$-modules *is of tangent type* if all V_α are "canonically" isomorphic to each other and hence to a given model space denoted by V. Formally, this is expressed by assuming that a family $\mathrm{id}_{\alpha,\beta} : V_\alpha \to V_\beta$ of $\mathbb{K}$-linear "identification isomorphisms" be given, satisfying the compatibility conditions $\mathrm{id}_{\alpha,\beta} \circ \mathrm{id}_{\beta,\gamma} = \mathrm{id}_{\alpha,\gamma}$ and $\mathrm{id}_{\alpha,\alpha} = \mathrm{id}_{V_\alpha}$. We say that the cube *is of (general) bundle type* if there are two $\mathbb{K}$-modules V and W such that each V_α is canonically isomorphic to V or to W; more precisely, we require that all V_α with $\alpha_1 = 1$ are canonically isomorphic to W and all other V_α are canonically isomorphic to V. (In a differential geometric context, we then often use also the the notation $\alpha = (\alpha_k, \ldots, \alpha_1, \alpha_0)$.) Thus for a cube of bundle type the axes of the cube (MA.6) are represented by



(SA.1)

For simplicity, in the following text we treat only the case of a cube of tangent type. In the differential geometric part, we apply these results as well to the general bundle type, leaving the slight modifications in the statements to the reader.

SA.2. *Action of the permutation group Σ_k.* If E is of tangent type, fixing for the moment the linear structure L^0, there is a linear action of the permutation group Σ_k on E such that, for all $\sigma \in \Sigma_k$, $\sigma : V_\alpha \to V_{\sigma\cdot\alpha}$ agrees with the canonical identification isomorphism:

$$\sigma : E \to E, \quad \sigma\Big(\sum_\alpha v_\alpha\Big) = \sum_\alpha \mathrm{id}_{\alpha,\sigma(\alpha)}\, v_\alpha \tag{SA.2}$$

(Strictly speaking, in order to define such an action of Σ_k we do not need the assumption that E is of tangent type, but only the weaker assumption that $V_\alpha \cong V_\beta$ whenever $|\alpha| = |\beta|$.) By our definition, the Σ_k-action is linear with respect to the linear structure L^0 on E. It will not be linear with respect to all L^b, or, put differently, it will not commute with the action of $\mathrm{Gm}^{1,k}(E)$. We let $E^\sigma = \{v \in E | \sigma v = v\}$ and denote by $E^{\Sigma_k} = \{v \in E | \forall \sigma \in \Sigma_k : \sigma v = v\}$ the fixed point space of the action of Σ_k. Then we define subgroups of $\mathrm{Gm}(E)^{0,k}$ by

$$\begin{aligned} \mathrm{Vsm}^{j,k}(E) &:= \{f^b \in \mathrm{Gm}^{j,k}(E) | \forall \sigma \in \Sigma_k : f_b \circ \sigma = \sigma \circ f_b\}, \\ \mathrm{Sm}^{j,k}(E) &:= \{f^b \in \mathrm{Gm}^{j,k}(E) | \ f^b(E^{\Sigma_k}) = E^{\Sigma_k}\}, \end{aligned} \tag{SA.3}$$

called the *group of very symmetric elements*, resp. the *group of symmetric elements of* $\mathrm{Gm}^{j,k}(E)$. It is clear that $\mathrm{Vsm}^{j,k}(E) \subset \mathrm{Sm}^{j,k}(E)$; we will see that in general this inclusion is strict. We say that a linear structure L^b is *very* or *totally symmetric* if the action of Σ_k is linear with respect to L^b, and that it is *(weakly) symmetric* if all E^σ, $\sigma \in \Sigma_k$, are linear subspaces of E with respect to L^b. Clearly, L^b is (very) symmetric if so is f^b.

Proposition SA.3. *The group* $\mathrm{Gm}^{j,k}(E)$ *is normalized by the action of* Σ_k. *More precisely, for every multilinear family* $b = (b^\Lambda)_\Lambda$ *and every* $\sigma \in \Sigma_k$,

$$\sigma \circ f_b \circ \sigma^{-1} = f_{\sigma \cdot b},$$

where $\sigma \cdot b$ *is the multilinear family*

$$(\sigma \cdot b)^\Lambda := \sigma \circ b^{\sigma^{-1} \cdot \Lambda} \circ (\sigma^{-1} \times \ldots \times \sigma^{-1}).$$

It follows that $\sigma \cdot L^b$ *is again a multilinear connection on* E*; more precisely,* $\sigma \cdot L^b = L^{\sigma \cdot b}$.

Proof. By a direct calculation, making the change of variables $\beta = \sigma.\alpha$ and $\Omega := \sigma.\Lambda$, we obtain

$$\begin{aligned} \sigma f_b \sigma^{-1}(\sum_\alpha v_\alpha) &= \sigma f_b(\sum_\alpha \mathrm{id}_{\alpha,\sigma^{-1}(\alpha)}(v_\alpha)) \\ &= \sigma\Big(\sum_\alpha \sum_{\Lambda \in \mathcal{P}(\alpha)} b^\Lambda\big(\mathrm{id}_{\sigma.\Lambda,\Lambda}(v_{\sigma(\Lambda)})\big)\Big) \\ &= \sum_\alpha \sum_{\Lambda \in \mathcal{P}(\alpha)} \mathrm{id}_{\alpha,\sigma.\alpha}\, b^\Lambda(\mathrm{id}_{\sigma.\Lambda,\Lambda}(v_{\sigma(\Lambda)})) \\ &= \sum_\beta \sum_{\Omega \in \mathcal{P}(\beta)} \mathrm{id}_{\sigma^{-1}.\beta,\beta}\, b^{\sigma^{-1}\Omega}(\mathrm{id}_{\Omega,\sigma^{-1}\Omega}(v_\Omega)) \\ &= f_{\sigma.b}(\sum_\beta v_\beta). \end{aligned}$$

Finally, since $\sigma \cdot L^0 = L^0$, it follows that

$$\sigma \cdot L^b = \sigma \cdot f^b \cdot L^0 = (\sigma \circ f^b \circ \sigma^{-1}) \cdot L^0 = f^{\sigma \cdot b} L^0 = L^{\sigma \cdot b}$$ ■

Corollary SA.4. *The following are equivalent:*

(1) $f_b \in \mathrm{Vsm}^{1,k}(E)$

(2) $\forall \sigma \in \Sigma_k\,,\ L^b = \sigma \cdot L^b$

(3) *the multilinear family b is Σ_k-equivariant in the sense that, for all $\alpha \in I_k$, all $\Lambda \in \mathcal{P}(\alpha)$ and all $\sigma \in \Sigma_k$, we have $(\sigma \cdot b)^\Lambda = b^\Lambda$, i.e., $b^{\sigma\cdot\Lambda} = \sigma \circ b^\Lambda \circ (\sigma^{-1} \times \ldots \times \sigma^{-1})$:*

$$\begin{array}{ccc} V_{\Lambda^1} \times \ldots \times V_{\Lambda^\ell} & \stackrel{b^\Lambda}{\to} & V_\alpha \\ \sigma \times \ldots \sigma \downarrow & & \downarrow \sigma \\ V_{\sigma(\Lambda^1)} \times \ldots \times V_{\sigma(\Lambda^\ell)} & \stackrel{b^{\sigma\cdot\Lambda}}{\to} & V_{\sigma\cdot\alpha} \end{array} \tag{SA.4}$$

■

The condition from the corollary implies that the b^Λ are *symmetric multilinear maps*. Let us explain this. We assume that f_b is very symmetric, i.e. (SA.4) commutes for all Λ and σ, and let us fix $\Lambda \in \mathcal{P}(\alpha)$ such that, for some $i \neq j$, we have $|\Lambda^i| = |\Lambda^j|$. Then there exists σ such that $\sigma(\Lambda^i) = \Lambda^j$ and $\sigma(\Lambda^m) = \Lambda^m$ for all other m, whence $\sigma(\Lambda) = \Lambda$ and hence $\sigma(\alpha) = \alpha$. The action of σ on V_Λ is determined by exchanging V_{Λ^i} and V_{Λ^j} and fixing all other V_{Λ^m}. Therefore (SA.6) can be interpreted in this situation by saying that b^Λ is symmetric with respect to exchange of arguments from V_{Λ^i} and V_{Λ^j}. In other words, the multilinear map b^Λ has all possible permutation symmetries in the sense that it is invariant under the subgroup $\{\sigma \in \Sigma_k |\, \sigma \cdot \Lambda = \Lambda\}$ of Σ_k. For instance, if $k = 2$, f_b is very symmetric iff $b^{01,10}$ “is a symmetric bilinear map $V \times V \to V$”, and if $k = 3$ (cf. Section MA.8), then f_b is very symmetric iff

(a) the three bilinear maps $b^{001,110}, b^{101,010}, b^{011,100}$ are conjugate to each other: $(23) \cdot b^{101,010} = b^{011,100} = (13) \cdot b^{110,001}$,

(b) the three bilinear maps $b^{001,010}, b^{001,100}, b^{010,100}$ are conjugate to each other: $(23) \cdot b^{001,010} = b^{001,100} = (12) \cdot b^{010,100}$ *and* are symmetric as bilinear maps,

(c) the trilinear map $b^{001,010,100}$ is a symmetric trilinear map.

Next we will give a sufficient condition for f_b to be weakly symmetric. For $k = 2$, all $f_b \in \mathrm{Gm}^{1,2}(E)$ are weakly symmetric (in the notation of BA.1, $f_b(v, v, w) = (v, v, w + b(v, v))$), and for $k = 3$, the reader may check by hand that f_b is weakly symmetric if the preceding condition (a) holds in combination with

(b') the three bilinear maps $b^{001,010}, b^{001,100}, b^{010,100}$ are conjugate to each other: $(23) \cdot b^{001,010} = b^{001,100} = (12) \cdot b^{010,100}$.

Our sufficient criterion will be based on the simple observation that $E^{\Sigma_k} = \cap_{\tau\in T} E^\tau$ where T is any set of generators of Σ_k; for instance, we may choose T to be the set of transpositions.

Proposition SA.5. *Let $f_b \in \mathrm{Gm}^{1,k}(E)$, $\tau \in \Sigma_k$ and consider the following conditions* (1) *and* (2)*.*

(1) $f_b(E^\tau) = E^\tau$

(2) *For all $\alpha \in I_k$ such that $\tau \cdot \alpha \neq \alpha$ and for all partitions $\Lambda \in \mathcal{P}(\alpha)$, the condition $(\tau \cdot b)^\Lambda = b^\Lambda$ holds.*

Then (2) *implies* (1)*. It follows that, if f_b satisfies* Condition (2) *for all elements τ of some set T of generators of Σ_k, then $f_b \in \mathrm{Sm}^{1,k}(E)$.*

Proof. By definition of E^τ,

$$v = \sum_\alpha v_\alpha \in E^\tau \Leftrightarrow \forall \alpha \in I_k : \mathrm{id}_{\alpha,\tau\cdot\alpha}(v_\alpha) = v_{\tau\cdot\alpha}.$$

It follows that $f^b(v)$ belongs to E^τ if and only if $\tau((f^b v)_\alpha) = \mathrm{id}_{\alpha,\tau\cdot\alpha}(f^b v)_{\tau\cdot\alpha}$ for all α, i.e.

$$\begin{aligned}\mathrm{id}_{\alpha,\tau\cdot\alpha}(v_\alpha) + \sum_{\Lambda\in\mathcal{P}(\alpha)} b^\Lambda(v_{\Lambda^1},\ldots,v_{\Lambda^\ell})) &= v_{\tau\cdot\alpha} + \sum_{\Omega\in\mathcal{P}(\tau\cdot\alpha)} b^\Omega(v_{\Omega^1},\ldots,v_{\Omega^\ell}) \\ &= v_{\tau\cdot\alpha} + \sum_{\Lambda'\in\mathcal{P}(\alpha)} b^{\tau\cdot\Lambda'}(v_{(\tau\cdot\Lambda')^1},\ldots,v_{(\tau\cdot\Lambda')^\ell})\end{aligned}$$

where we made the change of variables $\Lambda' := \tau^{-1}\cdot\Omega$. If we assume that $v \in E^\tau$, then this condition is equivalent to: for all $\alpha \in I_k$,

$$\sum_{\Lambda\in\mathcal{P}(\alpha)} \mathrm{id}_{\alpha,\tau\cdot\alpha}(b^\Lambda(v_{\Lambda^1},\ldots,v_{\Lambda^\ell})) = \sum_{\Lambda'\in\mathcal{P}(\alpha)} b^{\tau\cdot\Lambda'}(v_{(\tau\cdot\Lambda')^1},\ldots,v_{(\tau\cdot\Lambda')^\ell}) \qquad \text{(SA.5)}$$

If $\tau\cdot\alpha = \alpha$, then, since $\mathrm{id}_{\alpha,\alpha} = \mathrm{id}_{V_\alpha}$, this condition is automatically fulfilled, and if $\tau\cdot\alpha \neq \alpha$, then Condition (2) guarantees that it holds. Thus (SA.5) holds for all α, and (1) follows. The final conclusion is now immediate. ∎

One can show that, if in the preceding claim that τ is a transposition, then (1) and (2) are equivalent.

SA.6. *Curvature forms.* Let $L = L^b$ be a multilinear connection on E. We have seen that L and $\sigma\cdot L$ are multilinearly related (Prop. SA.3), and hence there exists a unique element $R := R^{\sigma,L} \in \mathrm{Gm}^{1,k}(E)$ such that

$$R^{\sigma,b}\cdot L = \sigma\cdot L. \qquad \text{(SA.6)}$$

This element is called the *curvature operator* and is given by

$$R^{\sigma,L} = \sigma\circ f_b\circ\sigma^{-1}\circ(f_b)^{-1} = f_{\sigma\cdot b}\circ(f_b)^{-1}. \qquad \text{(SA.7)}$$

In fact, $\sigma\circ f_b\circ\sigma^{-1}\circ(f_b)^{-1}\cdot L^b = \sigma\cdot f_b\cdot\sigma^{-1}L^0 = \sigma\cdot L^b$ since L^0 is very symmetric, i.e. $\sigma\cdot L^0 = L^0$. Directly from the definition of the curvature operators, we get the *cocycle relations*

$$R^{\sigma\circ\tau} = \sigma\circ R^\tau\circ\sigma^{-1}\circ R^\sigma, \qquad R^{\mathrm{id}} = \mathrm{id}\,. \qquad \text{(SA.8)}$$

The elementary matrix coefficients of R with respect to L,

$$R^\Lambda := (R^{\sigma,L})^\Lambda : V_\Lambda \to V_{\underline{\Lambda}} \qquad \text{(SA.9)}$$

are called *curvature forms (of σ, L).*

SA.7. *Symmetrizability.* We say that a multilinear connection $L = L^b$ (resp. an endomorphism $f_b \in \mathrm{Gm}^{1,k}(E)$) is *symmetrizable* if all its curvature operators $R^{\sigma,L}$,

$\sigma \in \Sigma_k$, belong to the central vector group $\mathrm{Gm}^{k-1,k}(E)$, i.e. $(\sigma \circ f_b \circ \sigma^{-1} \circ (f_b)^{-1})^\Lambda = 0$ whenver $\ell(\Lambda) < k$, or, equivalently,

$$\forall \Lambda : \ell(\Lambda) < k \Rightarrow (\sigma \cdot b)^\Lambda = b^\Lambda .$$

Lemma SA.8. *Assume L is symmetrizable. Then the orbit $\mathcal{O}_L := \mathrm{Gm}^{k-1,k}(E).L$ of L under the action of the vector group $\mathrm{Gm}^{k-1,k}(E)$ is an affine space over $\mathbb{K}$, isomorphic to $\mathrm{Gm}^{k-1,k}(E)$. This affine space is stable under the action of the group Σ_k which acts by affine maps on $\mathcal{O}_L$.*

Proof. All orbits of the action of the vector group $\mathrm{Gm}^{k-1,k}(E)$ on the space of connections have trivial stabilizer and hence are affine spaces. Let us prove that $\mathcal{O}_L$ is stable under the action of $\sigma \in \Sigma_k$: for $f_c \in \mathrm{Gm}^{k-1,k}(E)$,

$$\sigma . f_c . L = (\sigma \circ f_c \circ \sigma^{-1}) . \sigma . L = f_{\sigma \cdot c} . R^{\sigma, L} . L$$

belongs to the orbit $\mathrm{Gm}^{k-1,k}(E).L = \mathcal{O}_L$ (since $R^{\sigma,L} \in \mathrm{Gm}^{k-1,k}(E)$ by assumption of symmetrizability, and $f_{\sigma \cdot c} \in \mathrm{Gm}^{k-1,k}(E)$ by Prop. SA.3). Next, we show that the last expression depends affinely on c: indeed, conjugation by σ is a *linear* automorphism of the vector group $\mathrm{Gm}^{k-1,k}(E)$ (the action $c^\Lambda \mapsto (\sigma.c)^\Lambda$, for the only non-trivial component $\Lambda = \{\{1\}, \ldots, \{k\}\}$, is the usual action of the permutation group on k-multilinear maps $V^k \to V$, which is linear), and hence $f_c.L \mapsto f_{\sigma.c}.L$ is a linear automorphism of the linear space $(\mathcal{O}_L, L)$. Composing with the translation by $R^{\sigma,L}$, we see that $f_c.L \mapsto R^{\sigma,L}.f_{\sigma.c}.L = \sigma.(f_c.L)$ is affine.■

Corollary SA.9. *Assume that $L = L^b$ is symmetrizable and that the integers $2, \ldots, k$ are invertible in $\mathbb{K}$. Then the barycenter of the Σ_k-orbit of L,*

$$L_{\mathrm{Sym}} := \frac{1}{k!} \sum_{\nu \in \Sigma_k} \nu . L \in \mathcal{O}_L \tag{SA.10}$$

is well-defined and is a very symmetric linear structure on E. Assume that, in addition, L is invariant under Σ_{k-1}. Then

$$L_{\mathrm{Sym}} = \frac{1}{k} \sum_{j=1}^{k} (12 \cdots k)^j . L. \tag{SA.11}$$

Proof. For any finite group Γ, acting affinely on an affine space over $\mathbb{K}$, under the suitable assumption on $\mathbb{K}$, the barycenter of Γ-orbits is well-defined and is fixed under Γ since affine maps are compatible with barycenters. This can be applied to the situation of the preceding lemma in order to prove the first claim. Moreover, if $\Sigma_{k-1}.L = L$, then $\Sigma_k.L = \mathbb{Z}/(k).L$, where $\mathbb{Z}/(k)$ is the cyclic group generated by the permutation $(12 \cdots k)$, and we get (SA.11). ■

Summing up, $P := \frac{1}{k!} \sum_{\nu \in \Sigma_k} \nu$ may be seen as a well-defined projection operator on the space of all symmetrizable linear structures, acting affinely on orbits of the type of $\mathcal{O}_L$ and having as image the space of all very symmetric linear structures. The fibers of this operator are linear spaces, isomorphic to the kernel of the symmetrization operator on multilinear maps $V^k \to V$.

SA.10. *Bianchi's identities.* Assume L satisfies the assumptions leading to Equation (SA.11). Then $R^{L_{\mathrm{Sym}}} = 0$ since L_{Sym} is very symmetric; on the other hand,

$$0 = R^{L_{\mathrm{Sym}}} = \frac{1}{k}\sum_{j=1}^{k} R^{(12\cdots k)^j . L} = \frac{1}{k}\sum_{j=1}^{k}(12\cdots k)^j . R^L. \tag{SA.12}$$

Thus the sum of the cyclically permuted k-multilinear maps corresponding to R^L vanishes; for $k=3$ this is Bianchi's identity. Also, if σ is a transposition, then R^σ has the familier skew-symmetry of a curvature tensor: this follows from the cocylce relations

$$\mathrm{id} = R^{\mathrm{id}} = R^{\sigma^2} = \sigma \circ R^\sigma \circ \sigma \circ R^\sigma;$$

but under the assumption of SA.9, the composition of the curvature operators is just vector addition in the vector group $\mathrm{Gm}^{k-1,k}(E)$, whence

$$0 = \sigma \circ R^\sigma \circ \sigma + R^\sigma, \tag{SA.13}$$

expressing that the multilinear map R^σ is skew-symmetric in the i-th and j-th component if $\sigma = (ij)$.

Shift operators and shift invariance

SA.11. *Shift of the index set.* For $i \neq j$, the *(elementary) shift of the index set* $I_k \cong 2^{\mathbb{N}_k}$ *(in direction j with basis i)* is defined by

$$s := s_{ij} : 2^{\mathbb{N}_k} \to 2^{\mathbb{N}_k}, \quad s(A) = \begin{cases} A & \text{if } i \notin A, \\ A \cup \{j\} & \text{if } i \in A, j \notin A, \\ \emptyset & \text{if } i, j \in A. \end{cases} \tag{SA.14}$$

In terms of multi-indices,

$$s := s_{ij} : I_k \to I_k, \quad s(\alpha) = \begin{cases} \alpha & \text{if } \alpha_i = 0, \\ \alpha + e_j & \text{if } \alpha_i = 1, \alpha_j = 0, \\ 0 & \text{if } \alpha_i = 1, \alpha_j = 1. \end{cases} \tag{SA.15}$$

Then $s \circ s(A) = A$ if $i \in A$ and $s \circ s(A) = \emptyset$ else; hence $s \circ s$ is idempotent and may be considered as a "projection with base i".

SA.12. *Shift operators.* We continue to assume that E is of tangent type. Then, with respect to the linear structure L^0, there is a unique linear map $S := S_{ij} : E \to E$ corresponding to the elementay shift s of the index set, called an *(elementary) shift operator*

$$S_{ij} : E \to E, \quad S_{ij}(v_\alpha) = \mathrm{id}_{\alpha, s(\alpha)}(v_\alpha) = \begin{cases} v_\alpha & \text{if } \alpha_i = 0, \\ \mathrm{id}_{\alpha,\alpha+e_j}(v_\alpha) & \text{if } \alpha_i = 1, \alpha_j = 0, \\ 0 & \text{if } \alpha_i = 1, \alpha_j = 1. \end{cases} \tag{SA.16}$$

The components of $S_{ij}v$ are then given by

$$(S_{ij}(v))_\beta = \mathrm{pr}_\beta(S_{ij}v) = \mathrm{pr}_\beta(\sum_\alpha S_{ij}v_\alpha) = \begin{cases} v_\beta & \text{if } \beta_i = 0, \\ \mathrm{id}_{\beta-e_j,\beta}(v_{\beta-e_j}) & \text{if } \beta_i = 1 = \beta_j, \\ 0 & \text{if } \beta_i = 1, \beta_j = 0. \end{cases} \tag{SA.17}$$

We say that S_{ij} is a *positive shift* if $i > j$. Note that $S \circ S = p_{1-e_i}p_{e_i^*}$ is the projection onto a maximal subgeometry considered in Section MA.11 (Equation (MA.25)), and that S_{ij} preserves the factors of the decomposition $E = \widetilde{V}_{1-e_i} \oplus H_i$ (Section MA.11); it acts trivially on the first factor and by a two-step nilpotent map ε_j on the second factor, as may be summarized by the following matrix notation:

$$S_{ij} = \begin{pmatrix} \mathbf{1} & \\ & \begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix} \end{pmatrix} = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & \varepsilon_j \end{pmatrix}.$$

The shift operators are not intrinsic with respect to the whole general linear group $\mathrm{Gm}^{1,k}(E)$ but only with respect to a suitable subgroup.

Proposition SA.13. *The map f^b commutes with the shift $S = S_{ij}$ if and only if the following holds: for all $\alpha \in I_k$ such that $\alpha_i = 1$ and $\alpha_j = 0$ and for all $\Lambda \in \mathcal{P}(\alpha)$ the following diagram commutes:*

$$\begin{array}{ccc} V_{\Lambda^1} \times \ldots \times V_{\Lambda^\ell} & \xrightarrow{b^\Lambda} & V_\alpha \\ S \times \ldots \times S \downarrow & & \downarrow S \\ V_{\Lambda^1} \times \ldots \times V_{\Lambda^m \cup \{i\}} \times \ldots \times V_{\Lambda^\ell} & \xrightarrow{b^{s\cdot\Lambda}} & V_{s\cdot\alpha} \end{array} \tag{SA.18}$$

where m is the (unique) index such that $i \in \Lambda^m$. In this case,

$$s\Lambda = \{s\Lambda^1, \ldots, s\Lambda^\ell\} = \{\Lambda^1, \ldots, \Lambda^m \cup \{i\}, \ldots, \Lambda^\ell\}$$

is a partition of $s\alpha = \alpha \cup \{i\}$, so that $b^{s\Lambda}$ is a matrix coefficient of f^b.

Proof. On the one hand,

$$\begin{aligned} f_b(S_{ij}(v)) &= \sum_\alpha \sum_{\Lambda \in \mathcal{P}(\alpha)} b^\Lambda((S_{ij}v)_{\Lambda^1}, \ldots, (S_{ij}v)_{\Lambda^\ell}) \\ &= \sum_{\substack{\alpha \\ \alpha_i = 0}} b^\Lambda(v_{\Lambda^1}, \ldots, v_{\Lambda^\ell}) + \sum_{\substack{\alpha \\ \alpha_i = 1 = \alpha_j}} b^\Lambda((S_{ij}v)_{\Lambda^1}, \ldots, (S_{ij}v)_{\Lambda^\ell}) =: A + B, \end{aligned}$$

where A corresponds to sum over all α with $\alpha_i = 0$ (in this case we have also $(\Lambda^r)_i = 0$ for all i, and hence the shift s_{ij} fixes these indices), B corresponds to the sum over all α with $\alpha_i = 1 = \alpha_j$, and the term corresponding to the sum over all α with $\alpha_i = 1$, $\alpha_j = 0$ is zero since then there exists r with $(\Lambda^r)_i = 1$ and $(\Lambda^r)_j = 0$, and hence $(S_{ij}v)_{\Lambda^r} = 0$. On the other hand, by definition of S_{ij},

$$\begin{aligned} S_{ij}(f_b(v)) &= S_{ij}(\sum_\alpha \sum_{\Lambda \in \mathcal{P}(\alpha)} b^\Lambda(v_\Lambda)) = \sum_\alpha \sum_{\Lambda \in \mathcal{P}(\alpha)} S_{ij}(b^\Lambda(v_\Lambda)) \\ &= \sum_{\substack{\alpha \\ \alpha_i = 0}} \sum_{\Lambda \in \mathcal{P}(\alpha)} b^\Lambda(v_\Lambda) + \sum_{\substack{\alpha \\ \alpha_i = 1, \alpha_j = 0}} \sum_{\Lambda \in \mathcal{P}(\alpha)} \mathrm{id}_{\alpha,\alpha+e_j}(b^\Lambda(v_\Lambda)) =: A + B', \end{aligned}$$

where the first term A coincides with the first term of the preceding calculation. Comparing, we see that $f_b \circ S_{ij} = S_{ij} \circ f_b$ iff $B = B'$. Let us have a look at B: if $\alpha_i = 1 = \alpha_j$, then for a given $\Lambda \in \mathcal{P}(\alpha)$ two cases may arise:

(a) either there exists $m = m(\Lambda)$ such that $i, j \in \Lambda^m$; in this case

$$b^\Lambda((S_{ij}v)_{\Lambda^1}, \ldots, (S_{ij}v)_{\Lambda^\ell}) = b^\Lambda(v_{\Lambda^1}, \ldots, \mathrm{id}_{\Lambda^m - e_j, \Lambda^m}(v_{\Lambda^m - e_j}), \ldots, v_{\Lambda^\ell}),$$

(b) there exists $m \neq n$ such that $i \in \Lambda^m$, $j \in \Lambda^n$. In this case $(S_{ij}v)_{\Lambda^m} = 0$, and the corresponding term is zero.

Thus we are left with Case (a). In this case we let $\gamma := \alpha - e_j$ and define a partition $\Lambda' := \{\Lambda^1, \ldots, \Lambda^m - e_j, \ldots, \Lambda^\ell\}$; then $\Lambda' \in \mathcal{P}(\gamma)$ and $\Lambda = s(\Lambda')$, $\alpha = s(\gamma)$, and the condition $B = B'$ is seen to be equivalent to the condition from the claim. ■

Let us define the following subgroups of $\mathrm{Gm}^{1,k}(E)$:

$$\begin{aligned} \mathrm{Shi}^{\ell,k}(E) &:= \{f_b \in \mathrm{Gm}^{\ell,k}(E) \,|\, \forall i \neq j : \, S_{ij} \circ f_b = f_b \circ S_{ij}\}, \\ \mathrm{Shi}^{\ell,k}_+(E) &:= \{f_b \in \mathrm{Gm}^{\ell,k}(E) \,|\, \forall i > j : \, S_{ij} \circ f_b = f_b \circ S_{ij}\}. \end{aligned} \tag{SA.19}$$

For $k = 2$ the only target space for non-trivial multilinear maps is V_{11} and the only non-trivial partition is $\Lambda = ((01),(10))$, which corresponds to Case (b) in the preceding proof, whence $\mathrm{Gm}^{1,2}(E) = \mathrm{Shi}^{1,2}(E)$ (cf. also Section BA.4). We will see that the condition from Prop. SA.13 really has non-trivial consequences if $k \geq 3$.

Theorem SA.14. *Assume that $k > 2$ and that the multilinear connection L is invariant under all shifts. Then L is symmetrizable, and (if the integers are invertible in $\mathbb{K}$) L_{Sym} is again invariant under all shifts. Equivalently, if $f_b \in \mathrm{Shi}^{1,k}(E)$, then f^b is symmetrizable, and $f^{\mathrm{Sym}(b)}$ belongs to* $\mathrm{Shi}^{1,k}(E) \cap \mathrm{Vsm}^{1,k}(E)$.

Proof. We prepare the proof by some remarks on shifts. The shift maps $s_{ij} : I_k \to I_k$ are "essentially injective" in the sense that, if $\beta \neq 0$, then there is at most one α such that $s_{ij}(\alpha) = \beta$, and similarly for partitions. Then we write $\alpha \overset{s_{ij}}{\to} \beta$, having in mind that α is uniquely determined by β. Now let us look at the following pattern of correspondences of (ordered) partitions of length 2 for $k = 3$:

$$\begin{array}{ccccccc} & \{2\},\{3,1\} & & \{2,3\},\{1\} & & \{3\},\{1,2\} & \\ \overset{s_{31}}{\nearrow} & & \overset{s_{13}}{\nwarrow} \ \overset{s_{23}}{\nearrow} & & \overset{s_{32}}{\nwarrow} \ \overset{s_{12}}{\nearrow} & & \overset{s_{21}}{\nwarrow} \\ \{2\},\{3\} & \overset{(13)}{\leftrightarrow} & \{2\},\{1\} & \overset{(23)}{\leftrightarrow} & \{3\},\{1\} & \overset{(12)}{\leftrightarrow} & \{3\},\{2\} \end{array}$$

If f_b is invariant under all shifts, then the maps b^Λ corresponding to the partitions of the preceding pattern are all conjugate to each other under the respective shifts. But this implies that that they are also conjugate to each other under the transpositions as indicated in the bottom line, and by composing the three transpositions, we see that also the effect of the transposition (32) on the partition $(\{2\},\{3\})$ is induced by the shift-invariance. (At this point, the reader may recognize the famous "braid lemma" in disguise.) Also, the partitions from the top line are conjugate to each other under permutations (here the order is not important since their two elements are already distinguished by their cardinality). Summing up, for $\ell(\Lambda) = 2$, all bilinear maps $b^{\sigma.\Lambda}$, $\sigma \in \Sigma_3$, are conjugate to b^Λ under σ, and according to SA.7, this means that f_b is symmetrizable. This argument can be generalized to arbitrary $k > 2$ and Λ with $\ell(\Lambda) < k$ because for all such Λ there exists Ω such that one of the relations $\Lambda \stackrel{s_{ij}}{\to} \Omega$ or $\Omega \stackrel{s_{ij}}{\to} \Lambda$ holds (the only partition for which no such relation holds is the longest one, $\Lambda = \{\{1\}, \ldots, \{k\}\}$). Then, using the above arguments, one sees that b^Λ and $b^{\tau \cdot \Lambda}$ for $\tau = (i,j)$ are conjugate, and so on, leading to invariance under a set of generating transpositions of Σ_k.

Finally, note that, if $L = L^b$, then $L_{\mathrm{Sym}} = L^{\mathrm{Sym}(b)}$, where $\mathrm{Sym}(b)$ is obtained by symmetrising the highest coefficient of b. By assumption, f_b commutes with all shifts; then also all elements of the $\mathrm{Gm}^{k-1,k}(E)$-coset commute with all shifts since elements of $\mathrm{Gm}^{k-1,k}(E)$ commute with all shifts, by Prop. SA.13. Since $f_{\mathrm{Sym}(b)}$ belongs to this coset, it commutes with all shifts. Summing up, $f_{\mathrm{Sym}(b)} \in \mathrm{Shi}^{1,k}(E) \cap \mathrm{Vsm}^{1,k}(E)$. ■

SA.15. *The intersection* $\mathrm{Vsm}^{1,k}(E) \cap \mathrm{Shi}^{1,k}(E)$. Assume $f_b \in \mathrm{Vsm}^{1,k}(E) \cap \mathrm{Shi}^{1,k}(E)$. Then all components b^Λ of f_b for with fixed length $\ell(\Lambda) = \ell$ are conjugate to each other under a combination of shifts and permutations: by these operations, every partition Λ of length ℓ can be transformed into the the "standard partition" $\{\{1\}, \ldots, \{\ell\}\}$ of length ℓ. Thus f_b is determined by a collection of k symmetric multilinear maps $d^j : V^j \to V$ where V is the model space for all V_α. Let us denote by $\varepsilon^\alpha : V \to V_\alpha$ the isomorphism with the model space (so that $\mathrm{id}_{\alpha,\beta} = \varepsilon^\alpha \circ (\varepsilon^\beta)^{-1}$). Then f_b is given by

$$f_b(v) = \sum_\alpha \varepsilon^\alpha (\sum_{j=1}^{|\alpha|} \sum_{\Lambda \in \mathcal{P}_j(\alpha)} d^j(v_\Lambda)).$$

Theorem SA.16.

(1) *A multilinear map* $f_b \in \mathrm{Gm}(E)$ *belongs to* $\mathrm{Shi}^k(E) \cap \mathrm{Vsm}^k(E)$ *if and only if there are symmetric multilinear maps* $d^\ell : V^\ell \to V$ *such that, if* $\ell(\Lambda) = \ell$,

$$b^\Lambda(\varepsilon^{\Lambda_1} v_{\Lambda_1}, \ldots, \varepsilon^{\Lambda_\ell} v_{\Lambda_\ell}) = \varepsilon^{\underline{\Lambda}} d^l(v_{\Lambda_1}, \ldots, v_{\Lambda_\ell}).$$

(2) *If* $2, \ldots, k$ *are invertible in* $\mathbb{K}$, *then a multilinear map* $f_b \in \mathrm{Gm}(E)$ *belongs to* $\mathrm{Shi}^k(E) \cap \mathrm{Vsm}^k(E)$ *if and only if there is a polynomial map* $F : V \to V$ *of degree at most* k *such that* f_b *is given by the following purely algebraic analog of* Theorem 7.5*:*

$$f_b(v) = \sum_\alpha \sum_{\Lambda \in \mathcal{P}(\alpha)} (d^{\ell(\Lambda)} F)(0)(v_\Lambda),$$

where $(d^\ell F)(0)$ *is the* ℓ*-th differential of* F *at the origin. We write this also in the form* $f_b = (T^k F)_0$.

Proof. The condition from (1) is necessary since, as remarked above, by permutations and shifts every partition Λ can be transformed into the the standard partition $\widehat{\Lambda} := \{\{1\}, \ldots, \{\ell\}\}$ of length $\ell = \ell(\Lambda)$. Then we let $d^\ell := b^{\widehat{\Lambda}}$. The condition is also sufficient: with arbitrary choices of symmetric multilinear maps d^ℓ, the formula defines a multilinear map f_b; this map satisfies the shift-invariance and the permutation-invariance conditions (SA.13 and SA.4) and hence belongs to the intersection of both groups.

(2) Given $f_b \in \mathrm{Shi}^k(E) \cap \mathrm{Vsm}^k(E)$, we define d^ℓ as in (1) and define a polynomial map $F : V \to V$ by

$$F(x) := \sum_{\ell=1}^{k} \frac{1}{\ell!} d^\ell(x, \ldots, x).$$

Then we have $(d^\ell F)(0) = d^\ell$, and hence F satisfies the requirement from (2). Conversely, given F, we let $d^\ell := (d^\ell F)(0)$ and then define f_b as in Part (1). ■

The preceding theorem shows that the group $\mathrm{Vsm}^{1,k}(E) \cap \mathrm{Shi}^{1,k}(E)$ is a purely algebraic version of the "truncated polynomial group", denoted in [KMS93, Ch. 13.1] by B_1^k. Its Lie algebra (see next chapter, or [KMS93]) is the truncated (modulo degre $> k$) Lie algebra of polynomial vector fields on V.

PG. Polynomial groups

In the following, we will use the notion of *polynomial mappings between modules over a ring or field* $\mathbb{K}$, which we define as finite sums $p = \sum_{k=0}^{m} p_k$ of homogeneous polynomial mappings $p_k : V \to W$, where $p_k(x) = b_k(x, \ldots, x)$ is given by evaluating a (not necessary symmetric) $\mathbb{K}$-multilinear map $b_k : V^k \to W$ on the diagonal. However, in general it is not possible to recover the homogeneous parts p_k from the mapping $p : V \to W$. Therefore one either has to work with a more formal concept of polynomials (as developed by N. Roby in [Rob63]), or one has to make assumptions on the base ring. In order to simplify our presentation, we will choose the second option and make rather strong assumptions on $\mathbb{K}$: in Sections PG.1 – PG.4 we will assume that $\mathbb{K}$ is an infinite field, and from Section PG.5 on we will assume that $\mathbb{K}$ is a field of characteristic zero. Under these assumptions, the homogeneous parts can be recovered from p by algebraic differential calculus for polynomials – this is classical, and it can also be recovered in the general framework of differential calculus developed in [BGN04] since polynomial maps over infinite fields form a "C^0-concept" in the sense of [BGN04] (cf. also Appendix G.1). Another choice of assumptions could be to assume that $\mathbb{K}$ is a topological ring with dense unit group, and to admit only continuous polynomials.

PG.1. *Definition of polynomial groups.* A *polynomial group (over an infinite field* $\mathbb{K}$*) of degree at most* k (where $k \in \mathbb{N}$) is a $\mathbb{K}$-vector space M together with a group structure (M, m, i, e) such that $e = 0$, the product map $m : M \times M \to M$ and the inversion $j : M \to M$ are polynomial, and all iterated product maps, for $i \in \mathbb{N}$,

$$m^{(i)} : M^i \to M, \quad (x_1, \ldots, x_i) \mapsto x_1 \cdots x_i \tag{PG.1}$$

(which are polynomial maps since so is $m = m^{(2)}$) are polynomial maps of degree bounded by k.

PG.2. Our three main examples of polynomial groups are:

(1) The general multilinear group $M = \mathrm{Gm}^{1,k}(E)$, with its natural chart, is a polynomial group of degree at most k: with respect to a fixed multilinear connection $L = L^0$ on E, every $f \in \mathrm{Gm}^{1,k}$ is of the form $f = f^b = \mathrm{id} + X^b$ with a multilinear family b, and in this way $\mathrm{Gm}^{1,k}(E)$ is identified with the $\mathbb{K}$-module $M = \mathrm{Mult}_{>1}$ of multilinear families that are singular of order 1. The multiplication map $m : M \times M \to M$ is described by Prop. MA.13: it is clearly polynomial in (g, f), with degree bounded by k. Applying twice the matrix multiplication rule MA.13, we get the iterated product map

$$(m^{(3)}(h, g, f))^{\Lambda} = \sum_{\Xi \preceq \Omega \preceq \Lambda} h^{\Xi} \circ g^{\Xi|\Omega} \circ f^{\Omega|\Lambda} \tag{PG.2}$$

which is again polynomial of degree bounded by k: applied to an element $v \in E$, (PG.2) is a sum of multilinear terms; every occurence of f, g or h corresponds to a non-trivial contraction; but in a multilinear space of degree k, one cannot contract more than k times, so the degree is bounded by k.

Similar arguments apply to $m^{(j)}$. Finally, the inversion procedure used in the proof of Theorem MA.6 is also polynomial (details are left to the reader).

(2) For any Lie group G, the group $(T^kG)_e$, equipped with the $\mathbb{K}$-module structure coming from left or right trivialization, is a polynomial group of degree at most k – this is clear from Theorem 24.7 since the Lie algebra $(T^k\mathfrak{g})_e$ is nilpotent of order k. (Examples (1) and (2) make sense in arbitrary characteristic.)

(3) If $\mathbb{K}$ is of characteristic zero, then any nilpotent Lie algebra over $\mathbb{K}$, equipped with the Campbell-Hausdorff multiplication, is a polynomial group of degree bounded by the index of nilpotence of the Lie algebra.

PG.3. *The Lie bracket.* We write

$$m(x,y) = \sum_{i=0}^{k} m_i(x,y) = \sum_{p+q\leq k} m_{p,q}(x,y) \tag{PG.3}$$

for the decomposition of $m : M \times M \to M$ into homogeneous polynomial maps $m_i : M \times M \to M$, resp. into parts $m_{p,q}(x,y)$, homogeneous in x of degree p and in y of degree q. The constant term is $m_0(x,y) = m_0(0,0) = 0$, and the linear term is $m_1(x,y) = m_1((x,0)+(0,y)) = m_1(x,0)+m_1(0,y) = x+y$, whence $m_{1,0}(x,y) = x$, $m_{0,1}(x,y) = y$. For $j > 1$, we have $m_{j,0}(x,y) = 0 = m_{0,j}(x,y)$ since $m_{j,0}(x,y) = m_{j,0}(x,0)$ is the homogeneous component of degree j of $x \mapsto m(x,0) = x$. Thus

$$m(x,y) = x + y + m_{1,1}(x,y) + \sum_{\substack{p+q=3,\ldots,k \\ p,q\geq 1}} m_{p,q}(x,y)$$

with a bilinear map $m_{1,1} : M \times M \to M$. Similar arguments (cf. [Bou72, Ch.III, Par. 5, Prop. 1]) show that

$$\begin{aligned} x^{-1} = i(x) &= -x + m_{1,1}(x,x) \operatorname{mod} \deg 3 \\ xyx^{-1} &= y + m_{1,1}(x,y) - m_{1,1}(y,x) \operatorname{mod} \deg 3 \\ c_1(x,y) := xyx^{-1}y^{-1} &= m_{1,1}(x,y) - m_{1,1}(y,x) \operatorname{mod} \deg 3, \end{aligned} \tag{PG.4}$$

(where $\operatorname{mod} \deg 3$ has the same meaning as in [Bou72, loc.cit.]), and one proves that

$$[x,y] := m_{1,1}(x,y) - m_{1,1}(y,x) \tag{PG.5}$$

is a Lie bracket on M. If one defines the iterated group commutator by induction,

$$c_{j+1}(x_1,\ldots,x_{j+2}) := c_1(x_1, c_j(x_2,\ldots,x_{j+2})),$$

it is easily seen that

$$c_j(x_1,\ldots,x_{j+1}) = [x_1,[x_2,\ldots,[x_j,x_{j+1}]\ldots]] \operatorname{mod} \deg(j+2).$$

The maps c_j are polynomials, and their degree is uniformely bounded by some integer $N \in \mathbb{N}$ (write $c_j = m_r \circ \iota_j$, where $\iota_j : M^{j+1} \to M^r$ is a polynomial map whose degree is equal to the degree of inversion $j : M \to M$ and which is a certain composition of diagonal imbeddings and copies of j). Since the iterated Lie bracket is of degree $j+1$, it follows that it vanishes as soon as $j+1 > N$, and hence the Lie bracket is nilpotent. Let us now describe this nilpotent Lie algebra for the examples mentioned above (PG.2, (1) – (3)):

(1) Recall Formula (MA.19) (or Prop. MA.13) for the components of the product $m(g,f) = g \circ f$ in the group $\mathrm{Gm}^{1,k}(E)$. It is linear in g, but non-linear in f. The part which is linear in f is obtained by summing over all partitions Ω such that Λ is a refinement of Ω *of degree one*, i.e. there is exactly one element of Ω which is a union of at least two elements of Λ, whereas all other elements of Ω are also elements of Λ. Thus, if we define the *degree* of the refinement (Λ, Ω) to be the number of non-trivial contractions of Λ induced by Ω:
$$\deg(\Omega|\Lambda) := \mathrm{card}\{\omega \in \Omega | \, \omega \notin \Lambda\},$$
then
$$m_{1,1}(g,f)^\Lambda = \sum_{\substack{\Omega \prec \Lambda \\ \deg(\Omega|\Lambda)=1}} g^\Omega \circ f^{\Omega|\Lambda}.$$
Viewing f and g as polynomial mappings $E \to E$ and with the usual definition of differentials of polynomial maps, this can be written $m_{1,1}(g,f)(x) = dg(x) \cdot f(x)$, and thus
$$[g,f](x) = dg(x) \cdot f(x) - df(x) \cdot g(x)$$
is the usual Lie bracket of the algebra of polynomial vector fields on E. We denote the Lie algebra thus obtained by $\mathfrak{gm}^{1,k}(E)$. It is a *graded Lie algebra*; the grading depends on the linear structure L^0, whereas the *associated filtration* is an intrinsic feature, given by
$$\mathfrak{gm}^{1,k}(E) \supset \mathfrak{gm}^{2,k}(E) \supset \ldots \supset \mathfrak{gm}^{k-1,k}(E),$$
where $\mathfrak{gm}^{j,k}(E) = \mathrm{Mult}_{>j}(E)$ is the $\mathbb{K}$-module of multilinear families that are singular of order j.

(2) For $M = (T^k G)_e$, we recover the Lie algebra $(T^k \mathfrak{g})_0$ of $(T^k G)_e$, as follows easily from the fundamental commutation rule (24.1).

(3) Since the first terms of the Campbell-Hausdorff formula are $X + Y + \frac{1}{2}[X,Y] + \ldots$, we get the Lie bracket we started with: $\frac{1}{2}[X,Y] - \frac{1}{2}[Y,X] = [X,Y]$.

PG.4. *The power maps.* We decompose the iterated product maps into multihomogeneous parts,
$$m^{(j)} = \sum_{\substack{p_1,\ldots,p_j \geq 0 \\ p_1+\ldots+p_j \leq k}} m^{(j)}_{p_1,\ldots,p_j},$$
(where the $m^{(j)}_{p_1,\ldots,p_j}$ can be calculated in terms of $m_{p,q}$, but we will not need the explicit formula) and (following the notation from [Bou72] and from [Se65, LG 4.19]) we let
$$\psi_j : M \to M, \quad x \mapsto \sum_{p_1,\ldots,p_j > 0} m^{(j)}_{p_1,\ldots,p_j}(x,\ldots,x). \qquad \text{(PG.6)}$$

Since the degree of $m^{(j)}$ is bounded by k and $\sum_i p_i$ is the degree of a homogeneous component of $m^{(j)}$, it follows that $\psi_j = 0$ for $j > k$. By convention, $\psi_0(x) = 0$; then $\psi_1(x) = x$ and

$$\begin{aligned} x^2 &= m(x,x) = x + x + \sum_{i>1} m_i(x,x) = 2\psi_1(x) + \psi_2(x), \\ x^3 &= m^{(3)}(x,x,x) = \sum_{p_1,p_2,p_3 \geq 0} m^{(3)}_{p_1,p_2,p_3}(x,x,x) \\ &= (\sum_{p_1,p_2,p_3 \neq 0} + \sum_{p_1=0,p_2\neq 0,p_3\neq 0} + \sum_{p_2=0,p_1\neq 0,p_3\neq 0} + \sum_{p_3=0,p_2\neq 0,p_1\neq 0} + \\ &\quad \sum_{p_1=0,p_2=0,p_3\neq 0} + \sum_{p_1=0,p_3=0,p_2\neq 0} + \sum_{p_2=0,p_3=0,p_1\neq 0} + \sum_{p_1=p_2=p_3=0})(m^{(3)}_{p_1,p_2,p_3}(x,x,x)) \\ &= \psi_3(x) + 3\psi_2(x) + 3\psi_1(x) + \psi_0(x) \\ x^n &= \sum_{j=0}^{k} \binom{n}{j} \psi_j(x). \end{aligned}$$

See [Bou72, III, Par. 5, Prop. 2] for the details of the proof (mind that $\psi_j = 0$ for $j > k$). The last formula holds for all $n \in \mathbb{N}$ and in arbitrary characteristic.

PG.5. *One-parameter subgroups.* From now on we assume that $\mathbb{K}$ is a field of characteristic zero. Then, for fixed $x \in M$, we let

$$f := f_x : \mathbb{K} \to M, \quad t \mapsto x^t := \sum_{j=0}^{k} \binom{t}{j} \psi_j(x). \tag{PG.7}$$

Clearly, f is a polynomial map (of degree bounded by k), and we have just seen that $f(t+s) = f(t)f(s)$ for all $t,s \in \mathbb{N}$. For fixed $s \in \mathbb{N}$, both sides of this equality depend polynomially on $t \in \mathbb{K}$; they agree for $t \in \mathbb{N}$ and hence for all $t \in \mathbb{K}$ (since $\mathbb{K}$ is a field of characteristic zero). Since $f(n) = x^n$ for $n \in \mathbb{N}$, we may say that f_x is a *one-parameter subgroup through* x. In particular, it follows that

$$x^{-1} = \sum_{j=0}^{k} \binom{-1}{j} \psi_j(x) = \sum_{j=0}^{k} (-1)^j \psi_j(x), \tag{PG.8}$$

and hence the inversion map is a polynomial map of degree bounded by k. Note that f_x is not a one-parameter subgroup "in direction of x" since the derivative $f_x'(0)$ is not equal to x but is given by the linear term in

$$\begin{aligned} f_x(t) &= tx + \frac{t(t-1)}{2}\psi_2(x) + \frac{t(t-1)(t-2)}{3!}\psi_3(x) + \dots \\ &= t(x - \frac{1}{2}\psi_2(x) + \frac{1}{3}\psi_3(x) \pm \dots) \bmod (t^2), \end{aligned}$$

and thus f_x is a one-parameter subgroup in direction $f_x'(0) = \sum_{p=1}^{k} \frac{(-1)^{p-1}}{p}\psi_p(x)$. Thus, in order to define "the one-parameter subgroup in direction y", all we need is to invert the polynomial $\sum_{p=1}^{k} \frac{(-1)^{p-1}}{p}\psi_p(x)$, if possible. This is essentially the content of the following theorem:

Theorem PG.6. *Assume M is a polynomial group over a field $\mathbb{K}$ of characteristic zero. We denote by $\mathfrak{m}$ a formal copy of the $\mathbb{K}$-vector space M (where we may forget about the group structure). Then there exists a unique polynomial* $\exp : \mathfrak{m} \to M$ *such that*

(1) *for all $X \in \mathfrak{m}$ and $n \in \mathbb{Z}$,* $\exp(nX) = (\exp X)^n$,

(2) *the linear term of* $\exp$ *is the identity map* $\mathfrak{m} = M$.

The polynomial $\exp$ *is bijective, and its inverse map* $\log : M \to \mathfrak{m}$ *is again polynomial. We have the explicit formulae*

$$\exp(x) = \sum_{p=1}^{k} \frac{1}{p!} \psi_{p,p}(x)$$

$$\log(x) = \sum_{p=1}^{k} \frac{(-1)^{p-1}}{p} \psi_p(x)$$

where $\psi_{p,m}$ is the homogeneous component of degree m of ψ_p.

Proof. We proceed in several steps.

Step 1. Uniqueness of the exponential map. We claim that two polynomial homomorphisms $f, g : \mathbb{K} \to M$ with $f'(0) = g'(0)$ are equal (here we use the standard definition of the derivative of a polynomial). Indeed, this follows from general arguments given in [Se65, p. LG 5.33]: f, g being homomorphisms, they satisfy the same formal differential equation with same initial condition and hence must agree. It follows that, for every $v \in M$, there exists at most one polynomial homomorphism $\gamma_v : \mathbb{K} \to M$ such that $\gamma_v'(0) = v$. On the other hand, Property (1) of the exponential map implies (by the same "integer density" argument as above) that $\exp(tX) = (\exp X)^t$ for all $t \in \mathbb{K}$ and hence $\exp((t+s)X) = \exp(tX)\cdot\exp(sX)$ for all $t, s \in \mathbb{K}$. Invoking Property (2) of the exponential map, it follows that the exponential map, if it exists, has to be given by $\exp(v) = \gamma_v(1)$.

Step 2. We define the logarithm by the polynomial expression given in the claim and establish the functional equation of the logarithm. As remarked before stating the theorem, $\log(x) = f_x'(0)$. This implies the functional equation

$$\forall n \in \mathbb{Z} : \quad \log(x^n) = n \log(x).$$

Indeed, from $f_x(n) = x^n = f_{x^n}(1)$ we get $f_x(nt) = f_{x^n}(t)$ for all $t \in \mathbb{K}$ and hence

$$\log(x^n) = \frac{d}{dt}|_{t=0} f_{x^n}(t) = \frac{d}{dt}|_{t=0} f_x(nt) = n \log(x).$$

This implies $\log(x^t) = t \log(x)$ for all $t \in \mathbb{K}$.

Step 3. We define exp by the polynomial expression given in the claim and show that $\log \circ \exp = \mathrm{id}_{\mathfrak{m}}$. In fact, here the proof from [Bou72, III. Par. 5 no. 4] carries over word by word. We briefly repeat the main arguments : for $t \in \mathbb{K}^\times$,

$$\log(tx) = t \log((tx)^{t^{-1}}) = t \log(\sum_{r \le m} t^{m-r} \varphi_{r,m}(x)), \tag{PG.9}$$

where the polynomials $\varphi_{r,m}$ are defined by the expansion

$$x^t = \sum_{r,m} t^r \varphi_{r,m}(x)$$

into homogeneous terms of degree r in t and degree m in x. From the definition of x^t and of the components $\psi_{p,m}$ one gets in particular

$$\varphi_{1,m}(x) = \sum_{p \leq m} \frac{(-1)^{p-1}}{p} \psi_{p,m}(x), \qquad \varphi_{m,m}(x) = \frac{1}{m!} \psi_{m,m}(x).$$

One notes also that $\psi_{p,m} = 0$ if $m < p$. Taking $\frac{d}{dt}|_{t=0}$ in (PG.9), we get

$$x = \log(\sum_m \varphi_{m,m}(x)) = \log(\sum_m \frac{1}{m!} \psi_{m,m}(x)) = \log(\exp(x)).$$

Step 4. We show that $\exp \circ \log = \mathrm{id}_M$. We let

$$\gamma_v(t) := f_{\exp(v)}(t).$$

Then $\gamma_v'(0) = f'_{\exp(v)}(0) = \log(\exp(v)) = v$ by Step 3, and hence the preceding definition is consistent with the definition of γ_v in Step 1. We have

$$\gamma_v(1) = f_{\exp(v)}(1) = \exp(v).$$

The uniqueness statement on one-parameter subgroups from Step 1 shows that $f_x = \gamma_{\log(x)}$ since both have derivative $\log(x)$ at $t = 0$. Thus $x = f_x(1) = \gamma_{\log(x)}(1) = \exp(\log(x))$.

Step 5. We prove that exp satisfies (1) and (2). In fact, Property (2) follows from the fact that $\Psi_1(x) = \Psi_{1,1}(x) = x$. Since $\exp = (\log)^{-1}$, the functional equation (1) follows from the functional equation of the logarithm established in Step 2. (One may also adapt the usual "analytic" argument: from the uniqueness statement on one-parameter subgroups from Step 1 we get $\gamma_{tv}(1) = \gamma_v(t)$, which yields $\exp(tv) = (\exp v)^t$ and hence the functional equation (1).) ■

Let us add some comments on the theorem and on its proof. We have chosen notation $\mathfrak{m}$ and M such that on the side of $\mathfrak{m}$ only the $\mathbb{K}$-vector space structure is involved (see Theorem PG.8 below for relation with the Lie algebra structure) and on M only the group structure and its "regularity". In the theory of general formal groups, the existence and uniqueness of formal power series exp and log with the desired properties is proved by quite different arguments: in [Bou72, Ch.III, Par. 4, Thm. 4 and Def. 1] the existence of an exponential map is established by using the theory of universal envelopping algebras of Lie algebras, and the argument from [Se65, LG 5.35. – Cor. 2] relies heaviliy on the existence of a Campbell-Hausdorff multiplication.

PG.7. *Examples.* Let us give a more explicit description of the exponential maps for our examples PG.2 (1) – (3).

(1) Let $X = X^b \in M = \mathfrak{gm}^{1,k}(E)$, where b is a multilinear family b, singular of order 1. Then, for $\ell(\Lambda) > 1$,

$$\exp(X^b)^\Lambda = b^\Lambda + \sum_{j=2}^{k} \frac{1}{j!} \sum_{\substack{\Omega^{(1)} \prec \ldots \prec \Omega^{(j)} = \Lambda \\ \deg(\Omega^{(i)} | \Omega^{(i+1)}) = 1}} b^{\Omega^{(1)}} \circ b^{\Omega^{(1)} | \Omega^{(2)}} \circ \ldots \circ b^{\Omega^{(j-1)} | \Lambda}. \quad \text{(PG.10)}$$

Moreover, if E is of tangent or bundle type (Chapter SA), then the very symmetric and the (weakly) symmetric subgroups of $\mathrm{Gm}^{1,k}(E)$ are again polynomial groups, and the polynomial exp commutes with the action of the symmetric group (this follows from functoriality of exp). Thus the exponential map of $\mathrm{Sm}^{1,k}(E)$ and of $\mathrm{Vsm}^{1,k}(E)$ is given by restriction of the one of $\mathrm{Gm}^{1,k}(E)$. In particular, we get a combinatorial formula for the exponential map of the group $\mathrm{Vsm}^{1,k}(E) \cap \mathrm{Gm}^{1,k}(E)$ which, as seen at the end of Chapter SA, is a version of "jet group" from [KMS93, Section 13.1].

(2) See Theorem 25.2.

(3) If M is a Lie algebra with Campbell-Hausdorff multiplication, then exp is the identity map (here $\psi_j = 0$ for $j \geq 2$, and this property characterizes the "canonical chart", cf. [Bou72, III. Par. 5, Prop. 4]).

Concerning Example (1), we add the remark that, even if the integers are not invertible in $\mathbb{K}$, it is possible to express the group law of the general multilinear group $\mathrm{Gm}^{1,k}(E)$ in terms of the Lie bracket, by a formula similar to the one given for higher order tangent groups in Theorem 24.7. In any case, it can be shown that the filtration of the Lie algebra $\mathfrak{gm}^{1,k}(E)$ has an analog on the group level: the chain of subgroups from part (3) of Theorem MA.6 is a *filtration* of $\mathrm{Gm}^{1,k}(E)$ by normal subgroups in the sense that the group commutator satisfies $[\mathrm{Gm}^{i,k}(E), \mathrm{Gm}^{j,k}(E)] \subset \mathrm{Gm}^{i+j,k}(E)$.

Theorem PG.8. *Under the assumptions of the preceding theorem, the multiplication on a polynomial group is the Campbell-Hausdorff multiplication with respect to the polynomial* exp *from the preceding theorem. In other words,* $x * y := \log(\exp(x)\exp(y))$ *is given by the Campbell-Hausdorff formula, with the Lie bracket defined by* (PG.4)*. The Campbell-Hausdorff multiplication is again polynomial.*

Proof. The first claim follows from a general result on formal groups characterizing the exponential map as the unique formal homomorphism from the Campbell-Hausdorff group chunk to the group having the identity as linear term (cf. [Bou72, Ch.III, Par. 4, Th. 4 (v)], see also [Se65, LG 5.35]). In our case, the Campbell-Hausdorff multiplication is polynomial since since so are exp, log and the "original" multiplication. ∎

PG.9 *The projective limit and formal groups.* We could define now a *formal (power series) group* as the projective limit M_∞ of a projective system $\pi_{\ell,k} : M_k \to M_\ell$, $k \geq \ell$, of polynomial groups M_k, $k \in \mathbb{N}$, where M_k is of degree (at most) k. A formal power series group in the usual sense (see, e.g., [Haz78], [Se65]) gives rise to a sequence of truncated polynomial groups of which it is the projective limit. The exponential maps and logarithms $\exp_k$, resp. $\log_k$, then define, via their projektive limits, formal power series $\exp_\infty$ and $\log_\infty$ for M_∞. This way of proving the existence of the exponential and logarithm series has the advantage of using only "elementary" methods which are close to the usual notions of analysis on Lie groups.

References

[ALV91] Alekseevski, D.V., Lychagin, V.V and A.M. Vinagradov, "Geometry I: Basic Ideas and Concepts of Differential Geometry", vol. **28** of the Encyclopedia of Mathematical Sciences (Ed. R.V. Gamkrelidze), Springer-Verlag, Berlin 1991.

[B57] Berger, M., *Les espaces symétriques non-compacts*, Ann. Sci. Ec. Norm. Sup. (3) **74** (1957), 85 – 177.

[Bes78] Besse, A.L. "Manifolds all of whose Geodesics are Closed", Springer EMG **93**, New York 1978.

[BBG98] Bertelson, M., Bielavsky, P. et S. Gutt, *Parametrizing equivalence classes of invariant star products*, Lett. Math. Phys. **46** (4) (1998), 339 – 345.

[Be96] Bertram, W., *Un théorème de Liouville pour les algèbres de Jordan*, Bull. Soc. Math. Française **124** (1996), 299 – 327.

[Be00] Bertram, W., "The Geometry of Jordan and Lie Structures", Springer LNM **1754**, Berlin 2000.

[Be02] Bertram, W. *Generalized projective geometries: general theory and equivalence with Jordan structures*, Advances in Geometry **2** (2002), 329–369.

[Be04] Bertram, W. *From linear algebra via affine algebra to projective algebra*, Linear Algebra and its Applications **378** (2004), 109 – 134.

[Be07] Bertram, W., "Calcul différentiel topologique élémentaire" (book in preparation), Calvage et Mounet, Paris (to appear).

[Be08] Bertram, W., *Difference problems and differential problems*, preprint 2008, available on the arXiv.

[BD07] Bertram, W. and M. Didry, *Symmetric bundles and representations of Lie triple systems*, preprint, arXiv: math.DG/0710.1543.

[BGN04] Bertram, W., H. Glöckner and K.-H. Neeb, *Differential Calculus over General Base Fields and Rings*, Expo. Math. **22** (2004), 213–282. arXiv: math.GM/0303300.

[BeNe04] Bertram, W. and K.-H. Neeb, *Projective completions of Jordan pairs. Part I: The generalized projective geometry of a Lie algebra*. J. of Algebra **227** (2004), 474 – 519. arXiv: math.RA/0306272.

[BeNe05] Bertram, W. and K.-H. Neeb, *Projective completions of Jordan pairs. Part II: Manifold structures and symmetric spaces*. Geom. Ded. **112** (2) (2005), 73 – 113. arXiv: math.GR/0401236.

[BGV92] Berline, N., Getzler E. and M. Vergne, "Heat Kernels and Dirac Operators", Springer-Verlag, New York 1992.

[Bor02] Bordemann, M., *Sur l'existence d'une prescription d'ordre naturelle projectivement invariante*, preprint 2002, arXiv: math.DG/0208171.

[Bou67] Bourbaki, N. "Variétés différentielles et analytiques – Fascicule de résultats", Hermann, Paris 1967 – 1971.

[Bou72] Bourbaki, N. "Groupes et algèbres de Lie. Chapitres 1 – 3.", Hermann, Paris 1972.

[Bue95] Buekenhout, F. (editor), " Handbook of Incidence Geometry", Elsevier, Amsterdam 1995.

[Ca71] Cartan, H., "Calcul Différentiel", Hermann, Paris 1971.

[CFT01] Cattaneo, A., Felder, G. and L. Tomassini, *From local to global deformation quantization of Poisson manifolds*, preprint, arXiv math. QA/ 0012228v3.

[Chap03] Chaperon, M., "Calcul différentiel et calcul intégral", Dunod, Paris 2003.

[Ch01] Chapoton, F., *Algèbres de pré-Lie et algèbres de Hopf liées à la renormalisation*, C.R. Acad. Sci. Paris, t. **332** (2001), Série I, p. 1 – 4.

[Did07] Didry, M., *Construction of groups associated to Lie- and Leibniz-algebras*, Journal of Lie Theory **17** (2007), 399–426.

[D73] Dieudonné, J., "Introduction to the Theory of Formal Groups", Marcel Dekker, New York 1973.

[GD71] Grothendieck, A. and J. Dieudonné, "Eléments de Géométrie Algébrique. I.", Springer, Berlin 1971.

[Gi03] Giordano, P., *Differential Geometry in spaces of mappings using nilpotent infinitesimals*, preprint, arXiv: math.DG/0308119v1.

[Gl03a] Glöckner, H., *Implicit Functions from Topological Vector Spaces to Banach Spaces*, Israel J. Math. **155** (2006), 205 – 252; arXiv: math. GM/0303320.

[Gl03b] Glöckner, H., *Bundles of locally convex spaces, group actions, and hypocontinuous bilinear mappings*, submitted.

[Gl03c] Glöckner, H., *Every smooth p-adic Lie group admits a compatible analytic structure*, to appear in Forum Math. **18** (2006), 45 – 84. (also TU Darmstadt Preprint 2307, December 2003; arXiv:math. GR/0312113).

[Gl04a] Glöckner, H., *Examples of differentiable mappings into non-locally convex spaces*, TU Darmstadt Preprint 2272, March 2003, 6 pp, also arXiv: math. FA/0307040; Topol. Proceedings. **28** (2004), No. 2, 47 – 486.

[Gl04b] Glöckner, H., *Tensor products in the category of topological vector spaces are not associative*, Comment. Math. Univ. Carol. **45** (2004), no. 4, 607 – 614.

[Gl04c] Glöckner, H., *Fundamentals of direct limit Lie theory*, Compositio Math. **141** (2005), 1551-1577. (cf. TU Darmstadt Preprint 2324, March 2004; arXiv: math. GR/0403093).

[Gl04d] Glöckner, H., *Smooth Lie groups over local fields of positive characteristic need not be analytic*, J. Algebra **285** (2005), 356 – 371; arXiv: math. GR/0408239.

[Gl04e] Glöckner, H., *Lie groups over non-discrete topological fields*, TU Darmstadt Preprint 2356, July 2004; arXiv: math. GR/0408008.

[Haz78] Hazewinkel, M., "Formal Groups and Applications", Academic Press, New York 1978.

[Hel78] Helgason, S., "Differential Geometry and Symmetric Spaces", Academic Press 1978.

[Hel84] Helgason, S., "Groups and Geometric Analysis", Academic Press 1984.

[Ho01] Holdgrün, H.S., "Analysis. Band 2: Differential- und Integralrechnung mehrerer Variablen", Leins-Verlag, Göttingen 2001.

[KM87] Kainz, G. and P. Michor, *Natural transformations in differential geometry*, Czech. Math. Journal **37** (112) (1987), 584 – 607.

[Ko81] Kock, A., "Synthetic Differential Geometry", London Math. Soc. Lect. Note Series **51**, Cambridge Univ. Press, 1981.

[Koe69] Koecher, M., *Gruppen und Lie-Algebren von rationalen Funktionen*, Math. Z. **109**, 349–392 (1969).

[KMS93] Kolář, I., Michor P. W. and J. Slovák, "Natural Operations in Differential Geometry", Springer-Verlag, Berlin 1993.

[KoNo63] Kobayashi, S. and S. Nomizu, "Foundations of Differential Geometry, vol. I", Wiley, New York 1963.

[La99] Lang, S., "Fundamentals of Differential Geometry", Graduate Texts in Math. **191**, Springer-Verlag, Berlin 1999.

[Lau77] Laugwitz, D., "Differentialgeometrie", Teubner, Stuttgart 1977.

[Lau86] Laugwitz, D., "Zahlen und Kontinuum", BI Wissenschaftsverlag, Mannheim 1986.

[Lav87] Lavendhomme, R., "Leçons de Géométrie Différentielle Synthétique Naïve", Monographies de mathématique **3**, Louvain-la-Neuve 1987.

[Lo67] Loos, O., *Spiegelungsräume und homogene symmetrische Räume*, Math. Z. **99** (1967), 141 – 170.

[Lo69] Loos, O., "Symmetric Spaces I", Benjamin, New York 1969.

[Mac79] Macdonald, I. G., "Symmetric Functions and Hall Polynomials", Clarendon Press, Oxford 1979.

[MR91] Moerdijk, I. and G.E. Reyes, "Models for Smooth Infinitesimal Analysis", Springer-Verlag, New York 1991.

[Ne02a] Neeb, K.-H., *A Cartan-Hadamard Theorem for Banach-Finsler Manifolds*, Geom. Dedicata **95** (2002), 115 – 156.

[Ne02b] Neeb, K.-H., "Nancy lectures on infinite dimensional Lie groups", Lecture Notes, Nancy 2002.

[Ne06] Neeb, K.-H., *Towards a Lie theory of locally convex groups*, Jap. J. of Math. **1** (2006), 291–468.

[N88] Nel, L.D., *Categorial Differential Calculus for infinite dimensional Spaces*, Cahiers de Top. et Géom. Diff. Catégoriques, **29**, t. 4 (1988), 257–286.

[Nes03] Nestruev, J., "Smooth Manifolds and Observables", Springer-Verlag, New York 2003.

[P62] Pohl, W.F., *Differential Geometry of Higher Order*, Topology **1** (1962), 169–211.

[Pos01] Postnikov, M.M., “Riemannian Geometry”, vol. **91** of the Encyclopedia of Mathematical Sciences, Springer-Verlag, New York 2001.

[Ro66] Robinson, A, “Non-Standard Analysis”, North Holland, Amsterdam 1966.

[Rob63] Roby, N., *Lois polynômes et lois formelles en théorie de modules*, Ann. Sci. Ec. Norm. Sup. 3e série t. **80** (1963), 213 – 348.

[Sch84] Schikhof, W.H., “Ultrametric Calculus – An Introduction to p-adic Analysis”, Cambridge University Press, Cambridge 1984.

[Se65] Serre, J.-P., “Lie Algebras and Lie Groups”, Benjamin, New York 1965.

[Sh97] Sharpe, R.W., “Differential Geometry Cartan’s Generalization of Klein’s Erlangen Program”, Springer, New York 1997.

[Sp79] Spivak, M., “A comprehensive Introduction to Differential Geometry”, vol. 5 (Scd. Ed.), Publish and Perish, Houston 1979.

[W53] Weil, A., *Théorie des points proches sur les variétés différentiables*, Colloque de Géom. Diff., Strasbourg 1953, pp. 111–117 (in: A. Weil, Œuvres scientifiques, Vol. 2, Springer, New York 1980, pp. 103–109).

[T04] Tu, L.W., *Une courte démonstration de la formule de Campbell-Hausdorff*, J. of Lie Theory **14**, no. 2 (2004), 501–508.

[Tr67] Trèves, F., “Topological Vector Spaces, Distributions and Kernels”, Academic Press, New York 1967.