Differential Geometry over General Base Fields and Rings. Part IV: Geometric multilinear algebra

Wolfgang Bertram
Institut Elie Cartan – U.M.R. 7502
Université Henri Poincaré (Nancy I)
B.P. 239
54506 Vandœuvre-lès-Nancy Cedex
France
e-mail: bertram@iecn.u-nancy.fr

Abstract. We introduce the purely algebraic concept of *multilinear spaces.* These spaces can serve as models for higher order tangent bundles in differential geometry. Particular attention is paid to groups of automorphisms, in particular, we introduce the *general multilinear group* of a multilinear space. This group is a *polynomial group*, i.e. a group modelled on a $\mathbb{K}$-module such that its group structure is polynomial. Transposing certain results from the theory of analytic Lie groups and of formal (power series) groups, we develop some basic Lie theory for polynomial groups.

Contents.

AMS subject classification (2000): 14L05, 15A69

Key words: formal group, general multilinear group, multilinear algebra

Introduction

The present Part IV of *Differential Geometry over General Base Fields and Rings* is completely independent of the other parts and contains all purely algebraic complements to the main text. We treat two topics which are closely related to each other:

(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 $T^k F$, 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. Following the general pattern of the theory of formal groups (or analytic Lie groups), we develop some basic "Lie theory" for such groups.

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* (Part II, 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.

1. *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* $E := \oplus_{\alpha \in I_k} V_\alpha$ on which the *general multilinear group* $\mathrm{Gm}(E)$ 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 is already present in [Be02] and [Be04]. 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 general 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 $\mathrm{Gm}(E)$ 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. The structure of the fibers $(T^k F)_x$ for a general vector bundle F over M is slightly more complicated: some of the V_α are canonically isomorphic to $T_x M$, others to the fiber F_x; then we say that the multilinear space is of *(general) bundle type*. In both cases, the symmetric group Σ_k acts on the total space E, where the action is induced by the canonical action of Σ_k on the set $\mathbb{N}_k$ (resp. on the first k elements of $\mathbb{N}_{k+1}$). 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.

Finally, let us say 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 $\mathrm{Gm}(E)$. 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 [Gl03]).

2. *Polynomial groups.* The general multilinear group $\mathrm{Gm}(E)$ 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 having a direct and not merely "formal" intepretation in terms of mappings. In particular, following the general pattern of Lie theory for formal or analytic groups, we associate a Lie algebra to a polynomial group and define (in characteristic zero) an exponential map and a logarithm (Theorem PG.6). Our proof of Theorem PG.6 is much more elementary and closer to "usual" analysis on Lie groups than the proofs given for general formal groups in [Bou72] or [Se65], and this is all the more interesting because the polynomial theory in a sense already contains the general theory – every general Lie group G can be analyzed via the sequence of its polynomial tangent groups $(T^k G)_e$, $k \in \mathbb{N}$ (see Part III of this work). The basic lines of the theory of polynomial groups are described in Chapter PG; although the contents appears to be rather classical in nature, we have not been able to find appropriate references in the literature and therefore decided to add this chapter here.

Notation. $\mathbb{K}$ denotes a commutative base ring with unit 1. Notation concerning set-theory and partitions is introduced in Sections MA.1, MA.3 and MA.4.

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:

$$\begin{aligned} -_b(u,v,w) &= (-u,-v,-w+2b(u,v)), \\ (u,v,w) -_b (u',v',w') &= (u-u', v-v', w-w'+ \\ &\qquad 2b(u',v') - b(u,v') - b(u',v)), \\ (1-r)_b(x,y,z) +_b r_b(x',y',z') &= ((1-r)x+rx', (1-r)y+ry', \\ &\qquad (1-r)z+rz'+r(r-1)b(x-x',y-y')). \end{aligned} \tag{BA.4}$$

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

$$\begin{aligned} &\tau_{b,x} : E \to E, \\ &(u,v,w) \mapsto (x_1,x_2,x_3) +_b (u,v,w) = (x_1+u, x_2+v, x_3+w+b(x_1,v)+b(u,x_2)), \end{aligned} \tag{BA.5}$$

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.* As explained above, the additive group $\mathrm{Bil}(V_1 \times V_2, V_3)$ acts (simply transitively) on the set $\mathrm{brs}(V_1, V_2; V_3)$ 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 (equivalently, that are invariant under all maps f_b, $b \in \mathrm{Bil}(V_1 \times V_2, V_3)$), and we will say that an object on E is *intrinsic* if its definition is independent of b. For instance, the origin $(0,0,0) \in E$ is invariant under all f_b, 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,

$$f_b(V_1 \times V_2 \times 0) = V_1 \oplus_b V_2 \oplus_b 0 = \{(u, v, b(u,v)) \,|\, u \in V_1, v \in V_2\} \tag{BA.7}$$

which clearly depends on b (it is the graph of b). Thus we have the following intrinsic diagram of inclusion of "axes" and "planes":

$$\begin{array}{ccccc} V_1 & \to & E^{13} := V_1 \times V_3 & & \\ & \nearrow & & \searrow & \\ V_3 & & & & E = V_1 \times V_2 \times V_3 \\ & \searrow & & \nearrow & \\ V_2 & \to & E^{23} := V_2 \times V_3 & & \end{array} \tag{BA.8}$$

Next, the notion of being "parallel to an axis" is not always intrinsic: it is so for the parallels to the third axis $0 \times 0 \times V_3$ through any point $z \in V_1 \times V_2 \times V_3$:

$$E_z^3 := z +_b V_3 = \{(z_1, z_2, z_3 + w) \,|\, w \in V_3\} \tag{BA.9}$$

does not depend on b; we call it the *vertical space (through z)*. But the parallel to, say, $0 \times V_2 \times 0$ through $(u, 0, 0)$ is the "horizontal space"

$$H_{u,b} := (u,0,0) +_b V_2 = \{(u,0,0) +_b (0,v,0) = (u, v, b(u,v)) \,|\, v \in V_2\} \tag{BA.10}$$

which clearly depends on b. We claim that not only the parallels to the third axis are intrinsic, but also the parallels to the "planes" $E^{13} = V_1 \times 0 \times V_3$ and $E^{12} = 0 \times V_2 \times V_3$:

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, and the vertical spaces are precisely intersections of fibers of this double fibration. 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 $V_{(u,v,w)}^{13}$ 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 $(u, v, w) \mapsto (\lambda u, \lambda v, \lambda^2 w)$.

Next we note 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 $\mathrm{pr}_j \circ f_b = f_b$, and hence they are intrinsically defined linear maps. (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:

$$\begin{array}{ccccccccc} 0 & \to & V_3 & \to & V_1\times V_2\times V_3 & \stackrel{\mathrm{pr}_{12}}{\longrightarrow} & V_1\times V_2 & \to 0 \\ 0 & \to & V_2\times V_3 & \to & V_1\times V_2\times V_3 & \stackrel{\mathrm{pr}_1}{\longrightarrow} & V_1 & \to 0 \\ 0 & \to & V_1\times V_3 & \to & V_1\times V_2\times V_3 & \stackrel{\mathrm{pr}_2}{\longrightarrow} & V_2 & \to 0 \end{array} \tag{BA.13}$$

The fibers of pr_j are nothing but the spaces E_z^{j3}, $j=1,2$.

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 the *shift*

$$S = \begin{pmatrix} 1 & & \\ & 0 & 0 \\ & 1 & 0 \end{pmatrix} : V_1 \times V_2 \times V_3 \to V_1 \times V_2 \times V_3, \quad (u,v,w) \mapsto (u,0,v) \tag{BA.14}$$

commutes with all f_b: $f_b \circ S = S \circ f_b$, and hence S is an "intrinsic linear map", i.e. it is linear with respect to all b. It stabilizes the fibers E_z^{23}, and the restriction to E_z^{23}, 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_z^{23} 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 τ-fixed space $\{(v_1,v_1,v_3) \mid v_1 \in V_1, v_3 \in V_3\}$ is a submodule of (E,L^b), isomorphic to $V_1 \times V_3$ with 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

$$\begin{aligned} (v_1,v_3) +_q (v_1',v_3') &= (v_1+v_1', v_2+v_2' + q(v_1+v_1') - q(v_1') - q(v_1)) \\ \lambda(v_1,v_3) &= (\lambda v_1, \lambda v_3 + \lambda(\lambda-1) q(v_1)) \end{aligned}$$

and 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*

$$V_1 \to V_1 \times V_1 \times V_3, \quad v \mapsto (v,0,0) +_b (0,v,0) = (v,v,b(v,v))$$

which clearly depends on (the quadratic form of) 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*

$$\omega_z^1 := f|_{E_z^{13}} : E_z^{13} \to W, \quad \omega_z^2 := f|_{E_z^{23}} : E_z^{23} \to W$$

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.* It can also be seen directly that (1) is in fact independent of the linear structure because $\omega(f_b(u,v,w)) = \nu(u,v) + \lambda(b(u,v)) + \lambda(w)$, showing that the space of intrinsically bilinear maps is stable under the action of the group $\mathrm{Bil}(V_1, V_2; V_3)$.

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,*

$$\forall w \in V_3 : \quad \omega(z_1, z_2, z_3 + w) = \omega(z_1, z_2, z_3).$$ ■

If ω satisfies (1) and (2) then we say that ω is *homogeneous bilinear*, and we let $M_h(E,W)$ the space of homogeneous bilinear maps $E \to W$. The preceding statements show that

$$\mathrm{Bil}(V_1, V_2; W) \to M_h(E,W), \quad \nu \mapsto \omega := \nu \circ \mathrm{pr}_{12}$$

is a bijection with inverse $\nu \mapsto \nu \circ \iota_{12}$. In particular, the inverse map does not depend on the linear structure.

BA.9. *The skew-symmetrization operator.* Assume we are in the special case $V_1 = V_2$. Then to every intrinsically bilinear map $f : E \to W$ we can associate a *homogeneous* intrinsically bilinear map

$$\mathrm{alt} f := f - f \circ \tau$$

In fact, f and $f\circ\tau$ are intrinsically bilinear, hence so is $f - f\circ\tau$, and $\mathrm{alt} f$ is homogeneous: writing f in the form $f(u,v,w) = \omega(u,v) + \lambda(w)$ as in the preceding paragraph, we get

$$\mathrm{alt}\, f(u,v,w) = f(u,v,w) - f(v,u,w) = \omega(u,v) - \omega(v,u).$$

which clearly is homogeneous bilinear. The bilinear map $V_1 \times V_1 \to V_3$ corresponding to $\mathrm{alt} f$ is

$$\omega_{\mathrm{alt}\, f}(u,v) = f(u,v,w) - f(v,u,w).$$

For instance, if $f = \mathrm{pr}_3^b$, then $\mathrm{alt} f$ is the torsion of L^b.

Relation with the tensor product

BA.10. *The tensor space.* As explained in the introduction to this part, 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

$$V_1 \times V_2 \times V_3 \to Z, \quad V_1 \times V_2 \times V_3 \to \widehat{Z}, \quad (v_1,v_2,v_3) \mapsto v_1 + v_2 + v_3 + v_1 \otimes v_2. \tag{BA.15}$$

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

$$\begin{array}{ccc} (V_1 \times V_2) \times V_3 & \to & Z \\ f_b \downarrow & & \downarrow F_b \\ (V_1 \times V_2) \times V_3 & \to & Z \end{array} \tag{BA.16}$$

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 automorphisms of E, as defined later in Section MA.5, induce different automorphisms of $\widehat{Z}$. For the space Z this is not the case.) Note that the splitting of Z into a direct sum is not intrinsic, but the subspace $0 \oplus V_3$ is intrinsic: the sequence

$$0 \to V_3 \to Z \to V_1 \otimes V_2 \to 0 \tag{BA.17}$$

is intrinsic. (If $V_1 = V_2$, then, with respect to symmetric b's, then we may define $Z' := S^2 V_1 \oplus V_3$ and we get an intrinsic sequence $V_3 \to Z' \to S^2 V_1$.) 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.16) 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: the cardinality of A corresponds to $|\alpha| := \sum_i \alpha_i$; the singletons $A = \{i\}$ correspond to the "canonical basis vectors" $e_1, \ldots, e_k$; 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 $\alpha \leq \beta$); 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 $\{0,1\}^k$.

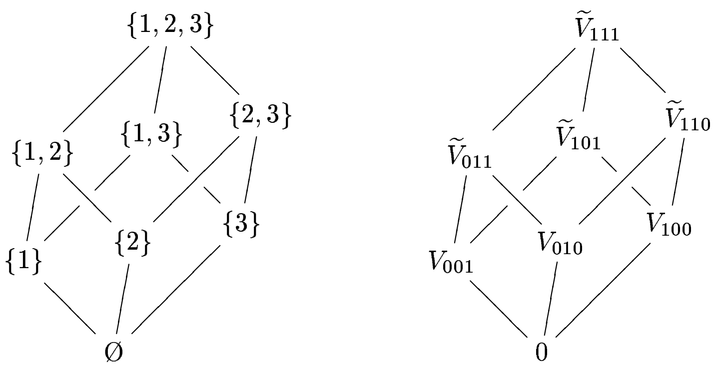
MA.2. *Cubes of $\mathbb{K}$-modules.* A *k-dimensional cube of $\mathbb{K}$-modules* is given by a family V_α, $\alpha \in I_k \setminus 0$, of $\mathbb{K}$-modules. We define the *total space*

$$E := \bigoplus_{\alpha \in I \setminus 0} V_\alpha.$$

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 $\Lambda = \{\Lambda^1, \ldots, \Lambda^l\}$ of the power set of A such that all Λ^i are non-empty and A is the disjoint union of the Λ^i: $A = \dot\cup_{i=1,\ldots,l}\Lambda^i$. We then say that $l := l(\Lambda)$ is the *length* of the partition. We denote by $\mathcal{P}(A)$ the set of all partitions of A and by $\mathcal{P}_l(A)$ (resp. by $\mathcal{P}_{\geq l}(\alpha)$) the set of all partitions of α of length l (resp. at least l):

$$\mathcal{P}(A) = \bigcup_{l=1}^{|A|} \mathcal{P}_l(A), \quad \mathcal{P}_l(A) = \{\Lambda = \{\Lambda_1, \ldots, \Lambda_l\} \in 2^{(2^A)} | \quad \dot\cup_{i=1,\ldots,l} \Lambda^i = A; \, \emptyset \notin \Lambda\}. \quad \text{(MA.7)}$$

Finally, for a finite set M, we denote by

$$\mathrm{Part}(M) := \bigcup_{\substack{A \in 2^M \\ A \neq \emptyset}} \mathcal{P}(A) \subset 2^{2^M} \quad \text{(MA.8)}$$

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

$$\underline{\Lambda} := A = \bigcup_{\nu \in \Lambda} \nu. \quad \text{(MA.9)}$$

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:

$$\Lambda = (\Lambda^1, \ldots, \Lambda^l), \quad \emptyset < \Lambda^1 < \ldots < \Lambda^l \leq A.$$

In terms of multi-indices, a *partition of a multi-index* $\alpha \in \{0,1\}^k$ is $\lambda = \{\lambda^1, \ldots, \lambda^l\} \subset I_k$ such that $\sum_{i=1}^l \lambda^i = \alpha$ and all $\lambda^i > 0$, or, using ordered sequences:

$$\lambda = (\lambda^1, \ldots, \lambda^l) \in (I_k)^l, \quad 0 < \lambda^1 < \ldots < \lambda^l, \quad \sum_{i=1}^l \lambda^i = \alpha. \quad \text{(MA.10)}$$

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 $l \times k$-matrix with rows $\lambda^1, \ldots, \lambda^l$. So, for instance,

$$\mathcal{P}(\{1,2,3\}) = \{\{1,2,3\}, \{\{1\},\{2,3\}\}, \{\{2\},\{1,3\}\}, \{\{1,2\},\{3\}\}, \{\{1\},\{2\},\{3\}\}$$

corresponds to

$$\mathcal{P}((111)) = \{(1 \quad 1 \quad 1), \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\}. \tag{MA.11}$$

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

$$l(\mu) = l(\lambda) =: l \quad \text{and} \quad \forall i = 1, \ldots, l : |\lambda^i| = |\mu^i|. \tag{MA.12}$$

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:

$$\underline{\Lambda} = A = \{a_1, \ldots, a_l\} \preceq \Lambda \preceq \overline{\Lambda} := \{\{a_1\}, \ldots, \{a_j\}\} = \{\{a\} | \, a \in A\}. \tag{MA.14}$$

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

$$b^\Lambda : V_{\Lambda^1} \times \ldots \times V_{\Lambda^{l(\Lambda)}} \to V'_\alpha \tag{MA.16}$$

be given. Then a *multinear map* or a *homomorphism of cubes of* $\mathbb{K}$*-modules* is a map of the form

$$f := f_b : E \to E', \quad v = \sum_{\alpha > 0} v_\alpha \mapsto \sum_{\alpha > 0} \sum_{\Lambda \in \mathcal{P}(\alpha)} b^\Lambda(v_{\Lambda^1}, \ldots, v_{\Lambda^{l(\Lambda)}}). \tag{MA.17}$$

Written out more explicitly,

$$f_b(v) = \sum_{\substack{\alpha \in I \\ \alpha > 0}} \sum_{l=1}^{|\alpha|} \sum_{\Lambda \in \mathcal{P}_l(\alpha)} b^\Lambda(v_{\Lambda^1}, \ldots, v_{\Lambda^l}).$$

To get a more concise notation, we let

$$V_\Lambda := V_{\Lambda^1} \times \ldots \times V_{\Lambda^l} \tag{MA.18}$$

and write $b^\Lambda : V_\Lambda \to V'_{\underline{\Lambda}}$. Note that $\Lambda = \{\Lambda^1, \ldots, \Lambda^l\}$ 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" $\Lambda^1 < \ldots < \Lambda^l$, 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 $l(\Lambda) = 1$ (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

$$\mathrm{Gm}^{0,k}(E) := \{f_b \in \mathrm{End}(E) \,|\, b \in \mathrm{Mult}(E),\ l(\Lambda) = 1 \Rightarrow b^\Lambda \in \mathrm{Gl}_{\mathbb{K}}(V_{\lambda^1})\}$$

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

$$\mathrm{Gm}^{1,k}(E) := \{f_b \in \mathrm{Gm}^{0,k}(E) \,|\, b \in \mathrm{Mult}(E),\ l(\Lambda) = 1 \Rightarrow b^\Lambda = \mathrm{id}_{V_{\lambda^1}}\}$$

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 $l(\Lambda) \leq j$, 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

$$\mathrm{Gm}^{j,k}(E) := \{f_b \in \mathrm{Gm}^{1,k}(E) \,|\, l(\lambda) = 2, \ldots, j \Rightarrow b^\lambda = 0\}$$

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 $w_\alpha = \sum_{\nu \in \mathcal{P}(\alpha)} f^\nu(v_{\nu_1}, \ldots, v_{\nu_l})$, then

$$\begin{aligned}
g(f(v)) &= g(\sum_\alpha w_\alpha) \\
&= \sum_\alpha \sum_{\Omega \in \mathcal{P}(\alpha)} g^\Omega(w_{\Omega^1}, \ldots, w_{\Omega^r}) \\
&= \sum_\alpha \sum_{\Omega \in \mathcal{P}(\alpha)} \sum_{\substack{\Lambda_1 \in \mathcal{P}(\Omega^1) \ldots \\ \Lambda_r \in \mathcal{P}(\Omega^r)}} g^\Omega(f^{\Lambda_1}(v_{\Omega^1_1}, \ldots, v_{\Omega^1_{l_1}}), \ldots, f^{\Lambda_r}(v_{\Omega^r_1}, \ldots, v_{\Omega^r_{l_r}})) \\
&= \sum_\alpha \sum_{\Lambda \in \mathcal{P}(\alpha)} \sum_{\Omega \preceq \Lambda} g^\Omega(f^{\Omega^1|\Lambda}(v_{\Omega^1_1}, \ldots, v_{\Omega^1_{l_1}}), \ldots, f^{\Omega^r|\Lambda}(v_{\Omega^r_1}, \ldots, v_{\Omega^r_{l_r}}))
\end{aligned}$$

where $r = l(\Omega)$ and $l_j = l(\Lambda_j)$ 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

$$(g \circ f)^\Lambda(v_\Lambda) = \sum_{\Omega \preceq \Lambda} g^\Omega(f^{\Omega^1|\Lambda}(v_{\Omega^1_1}, \ldots, v_{\Omega^1_{l_1}}), \ldots, f^{\Omega^r|\Lambda}(v_{\Omega^r_1}, \ldots, v_{\Omega^r_{l_r}})). \tag{MA.19}$$

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

$$T_0 f := \times_\alpha b^\alpha : E \to E', \quad \sum_\alpha v_\alpha \mapsto \sum_\alpha b^\alpha(v_\alpha), \tag{MA.20}$$

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 multilinear group.* The group $\mathrm{Gm}^{1,k}(E)$ 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 $\mathrm{Gm}^{0,k}(E)$, and we call $\mathrm{Gm}^{1,k}(E)$ the *general multilinear group of E*. The preceding proof shows that we have a (splitting) exact sequence of groups

$$1 \to \mathrm{Gm}^{1,k}(E) \subset \mathrm{Gm}^{0,k}(E) \overset{T_0}{\to} \times_{\alpha>0} \mathrm{Gl}(V_\alpha) \to 1 \tag{MA.21}$$

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.

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

$$f_b(v) = v + b^{001,010}(v_{001}, v_{010}) + b^{001,100}(v_{001}, v_{100}) + b^{010,100}(v_{010}, v_{100}) +$$
$$b^{001,110}(v_{001}, v_{110}) + b^{010,101}(v_{010}, v_{010}) + b^{011,100}(v_{011}, v_{100}) + b^{001,010,100}(v_{001}, v_{010}, v_{100}),$$

where the term in the second 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 $l(\lambda) = 2$, then we have simply $c^\lambda = b^\lambda + a^\lambda$. If $l(\lambda) = 3$, i.e., $\lambda = ((001),(010),(100))$, then there are three non-trivial partitions Ω that are strictly coarser than Λ, and we have:

$$c^\lambda(u,v,w) = a^\lambda(u,v,w) + b^\lambda(u,v,w) + b^{001,110}(u, a^{010,100}(v,w)) + \\ b^{010,101}(v, a^{001,100}(u,w)) + b^{011,100}(a^{001,010}(u,v), w).$$

Inversion of a unipotent $f = f_b$ is described as follows: if $l(\lambda) = 2$, then $(f^{-1})^\lambda = -f^\lambda$, and the coefficient belonging to $\lambda = ((100)(010)(001))$ is

$$(f^{-1})_\lambda(u,v,w) = -b^\lambda(u,v,w) + b^{001,110}(u, b^{010,100}(v,w)) + \\ b^{010,101}(v, b^{001,100}(u,w)) + b^{011,100}(b^{001,010}(u,v), w).$$

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-linearly 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-linearly 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 in case $k = 2$ (cf. Section BA.1) and will not be written out here. Let us just make the following remarks. It is easy to write the sum of elements taken from axes: since f_b acts trivially an 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 automorphism group $\mathrm{Gm}^{0,k}(E)$.

MA.10. *Intrinsic structures: subgeometries.* The group $\mathrm{Gm}^{1,k}(E)$ acts trivially on all axes; hence the axes are intrinsic subspaces, not only as sets, but also as sets with their induced $\mathbb{K}$-module structure. For any $\alpha \in I$, 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 $\Lambda^1, \ldots, \Lambda^l$.

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 $f_b \in \mathrm{Gm}^{1,k}(E)$:

$$\begin{array}{ccc} E & \overset{f_b}{\to} & E \\ p_\beta \downarrow & & p_\beta \downarrow \\ \widetilde{V}_\beta & \overset{f_b}{\to} & \widetilde{V}_\beta \end{array} \qquad \text{(MA.24)}$$

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 $j \in \{1,\dots,k\}$ and let $\beta := e_j^* := \sum_{i\neq j} e_i$ correspond to the complement $\mathbb{N}_k \setminus \{j\}$. The corresponding projection is

$$p_{e_j^*} : E \to \widetilde{V}_{(1\dots 101 \dots 1)} = \bigoplus_{\substack{\alpha\in I \\ \alpha_j = 0}} V_\alpha, \quad \sum_{\alpha > 0} v_\alpha \mapsto \sum_{\alpha_j = 0} v_\alpha. \qquad \text{(MA.25)}$$

The kernel is called the *j-th hyperplane*:

$$H_j := K_{e_j^*} = \ker(p_j) = \bigoplus_{\substack{\alpha\in I \\ \alpha_j = 1}} V_\alpha = \bigoplus_{\substack{\alpha\in I \\ e_j \subseteq \alpha}} V_\alpha = H_{e_j} \qquad \text{(MA.26)}$$

We claim that $\mathrm{Gm}^{1,k}(E)$ acts trivially on H_j. Indeed, assume $v = \sum_{\alpha_j = 1} v_\alpha \in H_{e_j}$. Then $f_b(v) = v + \sum_\lambda b^\lambda(v_{\lambda^1},\dots,v_{\lambda^l}) = v$ since $\lambda^i_j = 1$ for $i = 1,\dots,l$ 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

$$\mathrm{pr}_j := p_{e_j} : E \to V_j := V_{e_j}, \quad v = \sum v_\alpha \mapsto v_{e_j}. \qquad \text{(MA.27)}$$

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 $i \neq j$,

$$H_{ij} := H_{e_i^*} \cap H_{e_j^*} = \ker(p_{e_i^*} \times p_{e_j^*}) = \bigoplus_{\substack{\alpha\in I \\ \alpha_i = 1 = \alpha_j}} V_\alpha = H_{e_i + e_j} \qquad \text{(MA.28)}$$

is an intrinsic subspace, and so is the "vertical subspace" $\bigoplus_{\alpha:|\alpha|\geq 2} V_\alpha$, kernel of the projection

$$\mathrm{pr} := \mathrm{pr}_1 \times \ldots \times \mathrm{pr}_k : E \to V_1 \times \ldots \times V_k, \quad v \mapsto (\mathrm{pr}_1(v),\dots,\mathrm{pr}_k(v)). \qquad \text{(MA.29)}$$

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 connection in a multilinear space should be seen as an analog of choosing a base in a free $\mathbb{K}$-module V, and the analog of the map $\mathbb{K}^n \to V$ induced by a basis of n elements is the map

$$\Phi^b : \bigoplus_{\alpha>0} V_\alpha \to E, \quad v = (v_\alpha)_\alpha \mapsto f_b(v) \tag{MA.30}$$

(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

$$\begin{array}{ccc} \bigoplus_{\alpha>0} V_\alpha & \overset{\Phi^b}{\to} & E \\ D^b_c(f) \downarrow & & \downarrow f \\ \bigoplus_{\alpha>0} V'_\alpha & \overset{\Phi^c}{\to} & E' \end{array} \tag{MA.31}$$

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',

$$f^{\Omega|\Lambda} := \mathrm{pr}_\Omega \circ f \circ \iota_\Lambda : V_\Lambda \to V'_\Omega, \qquad \begin{array}{ccc} E & \overset{f}{\to} & E' \\ \iota_\Lambda \uparrow & & \downarrow \mathrm{pr}_\Omega \\ V_\Lambda & \overset{f^{\Omega|\Lambda}}{\to} & V'_\Omega \end{array} \tag{MA.32}$$

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 $l(\Omega) = r$, $r \leq l(\Lambda)$, $\Omega^i = \{\Omega^i_1, \ldots, \Omega^i_{l_i}\}$. Then the effective matrix coefficient is explicitly given by

$$\begin{aligned} f^{\Omega|\Lambda} &: V_\Lambda \to V_\Omega = V_{\Omega^1} \times \ldots \times V_{\Omega^r}, \\ (v_{\Lambda^1}, \ldots, v_{\Lambda^l}) &\mapsto (b^{\Omega^1|\Lambda}(v_{\Omega^1_1}, \ldots, v_{\Omega^1_{l_1}}), \ldots, b^{\Omega^r|\Lambda}(v_{\Omega^r_1}, \ldots, v_{\Omega^r_{l_r}})) \end{aligned} \tag{MA.33}$$

(2) We say that a matrix coefficient $f^{\Omega|\Lambda}$ is *elementary* if $l(\Omega) = 1$ 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 $f = f^b$:

$$b^\Lambda = f^{\underline{\Lambda}|\Lambda} =: f^\Lambda : V_\Lambda \to V'_{\underline{\Lambda}} \qquad \begin{array}{ccc} E & \overset{f_b}{\to} & E' \\ \iota_\Lambda \uparrow & & \downarrow \mathrm{pr}_\alpha \\ V_\lambda & \overset{b^\Lambda}{\to} & V'_\alpha \end{array} \tag{MA.34}$$

where $\alpha = \underline{\Lambda}$. In this situation, we may also use the notation $f^{\Omega|\Lambda} =: b^{\Omega|\Lambda}$.

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*

$$(g \circ f)^\Lambda = g^\Lambda + f^\Lambda + \sum_{r=2}^{l(\Lambda)-1} \sum_{\substack{\Omega \prec \Lambda \\ l(\Omega)=r}} g^\Omega \circ f^{\Omega|\Lambda}.$$

Proof. Using the terminology of matrix coefficients, Formula (MA.19) from the proof of Theorem MA.6 can be re-written

$$(g \circ f)^\Lambda = \sum_{\Omega \preceq \Lambda} g^\Omega \circ f^{\Omega|\Lambda} = \sum_{r=1}^{l(\Lambda)} \sum_{\substack{\Omega \preceq \Lambda \\ l(\Omega)=r}} g^\Omega \circ f^{\Omega|\Lambda}. \tag{MA.35}$$

The proposition follows from this formula. ■

Corollary MA.14. *The subgroup* $\mathrm{Gm}^{k-1,k}(E)$ *is central in* $\mathrm{Gm}^{1,k}(E)$. *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 $f \in \mathrm{Gm}^{k-1,k}(E)$ iff $f^\Lambda = 0$ whenever $l(\Lambda) \neq 1, k$. 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

$$l(\Lambda) < k \Rightarrow c^\Lambda = a^\Lambda, \qquad l(\Lambda) = k \Rightarrow c^\Lambda = a^\Lambda + h$$

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 $\mathbb{K}$). We let also

$$\iota^b : V_1 \times \ldots \times V_k \to E, \quad (v_{e_1}, \ldots, v_{e_k}) \mapsto (b \textstyle\sum) v_{e_i}$$

be the inclusion map for the longest partition (sum with respect to the linear structure L^b).

Proposition MA.16. *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 with inverse given by

$$\mathcal{M}_h(E, V') \to \mathrm{Hom}(V_1, \ldots, V_k; V'), \quad f \mapsto \underline{f} := f \circ \iota^b.$$

The map defined by the preceding formula does not depend on the linear structure L^b, *but its extension*

$$\mathcal{M}(E, V') \to \mathrm{Hom}(V_1, \ldots, V_k; V'), \quad f \mapsto \underline{f} := f \circ \iota^b.$$

does depend on L^b.

Proof. We prove first that $\overline{f}$ is intrinsically multilinear: the restrictions $\mathrm{pr}_i|_{H_j}$ are zero except for $i=j$ (for $|\alpha|=1$, only the e_j-component is non-zero on H_j), and it follows that the restriction of f to $y+H_j$ is affine. Moreover, $\overline{f}$ is homogeneous since the kernel of pr is E_2 and hence $\underline{f}$ will be constant on E_2 and on all other fibers of pr.

Next we prove that $\underline{f}$ is multilinear in the usual sense (where f may be homogeneous or not): in fact, $\underline{f}(v_1,\ldots,v_k)$ (with $v_j\in V_{e_j}$) is linear in v_j since f is linear on $(v_1,\ldots,0,\ldots,v_k)+H_j$. Thus $\underline{f}$ is multilinear in the usual sense. If f is homogeneous, then the definition of $\underline{f}$ does not depend on the linear structure on E because $f(\sum^{(b)} v_{e_i}) = f(\sum v_{e_i} + \sum_\lambda b^\lambda(v_{\lambda_1},\ldots,v_{\lambda^l})) = f(\sum v_{e_i})$ by homogenity, using that b^λ takes values in an V_α with $|\alpha|\geq 2$, and such a V_α lies in some H_{ij}. If f is non-homogeneous, then $\underline{f}$ may very well depend on L^b.

Finally, a direct calculation shows that $f\mapsto \underline{f}$ and $f\mapsto \overline{f}$ are inverse to each other. ∎

Summing up, the following diagram encodes two different ways of seeing the same object f:

$$\begin{array}{ccc} E & & \\ \mathrm{pr}\downarrow\uparrow \iota^b & \overset{\overline{f}}{\searrow} & \\ V_1\times\ldots\times V_k & \overset{\underline{f}}{\to} & V' \end{array} \tag{MA.36}$$

Tensor space and tensor functor

MA.17. *The tensor space.* For a partition Λ we let

$$\widehat{V}_\Lambda := V_{\Lambda^1}\otimes\ldots\otimes V_{\Lambda^l}$$

and we define the *(big) tensor space of* E

$$\widehat{E} := \bigoplus_{\Lambda\in\mathrm{Part}(I)} \widehat{V}_\Lambda = \bigoplus_{\alpha>0}\bigoplus_{l=1}^{|\alpha|}\bigoplus_{\Lambda\in\mathcal{P}_l(\alpha)} \widehat{V}_\Lambda .$$

There is a canonical map

$$E\to\widehat{E},\quad (v_\alpha)_\alpha\mapsto \sum_\Lambda v_{\Lambda^1}\otimes\ldots\otimes v_{\Lambda^l}.$$

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

$$f^{\Omega|\Lambda} = b^{\Omega^1|\Lambda}\times\ldots\times b^{\Omega^r|\Lambda} : \times V_{\Omega^i|\Lambda}\to \times \underline{V_{\Omega^i|\Lambda}}$$

and define

$$\widehat{f}^{\Omega|\Lambda} := \widehat{b}^{\Omega^1|\Lambda}\otimes\ldots\otimes\widehat{b}^{\Omega^r|\Lambda} : \otimes V_{\Omega^i|\Lambda}\to\otimes\underline{V_{\Omega^i|\Lambda}}.$$

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:

$$(\widehat{g \circ f})^{\Omega|\Lambda} = \sum_\nu \widehat{g}^{\Omega|\nu} \circ \widehat{f}^{\nu|\Lambda},$$

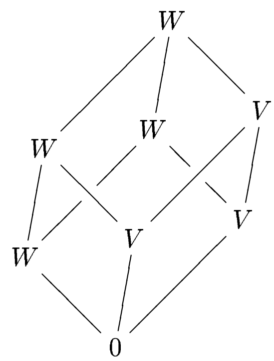
where the sum is over all partitions ν; but, by the preceding definitions, only terms with $\Omega \preceq \nu \preceq \Lambda$ really contribute to the sum, and hence we get the component $(\widehat{g \circ f})^{\Omega|\Lambda}$. ■

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 $\widehat{f}$ (it is here that in BA.10 we needed $\widehat{E}$ and not Z). 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 symmetric powers and identifying some of the $\widehat{V}_\Lambda$, 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].

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 polyomials 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 $\alpha = (\alpha_2, \alpha_1, \alpha_0)$.) 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:

$$\sigma : E \to E, \quad \sigma(\sum_\alpha v_\alpha) = \sum_\alpha \mathrm{id}_{\alpha,\sigma(\alpha)}\, v_\alpha \tag{SA.2}$$

(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 $E^\sigma = \{v \in E \mid \sigma v = v\}$ and denote by

$$E^{\Sigma_k} = \{v \in E \mid \forall \sigma \in \Sigma_k : \sigma v = v\} = \bigcap_{\sigma\in\Sigma_k} E^\sigma$$

the fixed point space of the action of Σ_k, and we define subgroups of $\mathrm{Gm}(E)^{0,k}$ by

$$\begin{aligned} \mathrm{Vsm}^{j,k}(E) &:= \{f^b \in \mathrm{Gm}^{j,k}(E) \mid \forall \sigma \in \Sigma_k : f_b \circ \sigma = \sigma \circ f_b\}, \\ \mathrm{Sm}^{j,k}(E) &:= \{f^b \in \mathrm{Gm}^{j,k}(E) \mid f^b(E^{\Sigma_k}) = E^{\Sigma_k}\}, \end{aligned} \tag{SA.3}$$

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 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

$$\begin{aligned} \sigma f_b \sigma^{-1}(\sum_\alpha v_\alpha) &= \sigma f_b(\sum_\alpha \mathrm{id}_{\alpha,\sigma^{-1}(\alpha)}(v_\alpha)) \\ &= \sigma(\sum_\alpha \sum_{\Lambda \in \mathcal{P}(\alpha)} b^\Lambda(\mathrm{id}_{\sigma.\Lambda,\Lambda}(v_{\sigma(\Lambda)}))) \\ &= \sum_\alpha \sum_{\Lambda \in \mathcal{P}(\alpha)} \mathrm{id}_{\alpha,\sigma.\alpha}\, b^\Lambda(\mathrm{id}_{\sigma.\Lambda,\Lambda}(v_{\sigma(\Lambda)})) \\ &= \sum_\beta \sum_{\Omega \in \mathcal{P}(\beta)} \mathrm{id}_{\sigma^{-1}.\beta,\beta}\, b^{\sigma^{-1}\Omega}(\mathrm{id}_{\Omega,\sigma^{-1}\Omega}(v_\Omega)) \\ &= f_{\sigma.b}(\sum_\beta v_\beta). \end{aligned}$$

Finally, since $\sigma \cdot L^0 = L^0$, it follows that

$$\sigma \cdot L^b = \sigma \cdot f^b \cdot L^0 = (\sigma \circ f^b \circ \sigma^{-1}) \cdot L^0 = f^{\sigma \cdot b} L^0 = L^{\sigma \cdot b} \qquad \blacksquare$$

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* $\Lambda \in \mathrm{Part}(\mathbb{N}_k)$ *and all* $\sigma \in \Sigma_k$, *we have* $(\sigma \cdot b)^\Lambda = b^\Lambda$, *i.e.,* b^Λ *and* $b^{\sigma \cdot \Lambda}$ *are conjugate under* σ:

$$b^{\sigma\cdot\Lambda} = \sigma \circ b^\Lambda \circ (\sigma^{-1} \times \ldots \times \sigma^{-1}) : \qquad \begin{array}{ccc} V_{\Lambda^1} \times \ldots \times V_{\Lambda^l} & \overset{b^\Lambda}{\to} & V_\alpha \\ \sigma \times \ldots \sigma \downarrow & & \downarrow \sigma \\ V_{\sigma(\Lambda^1)} \times \ldots \times V_{\sigma(\Lambda^l)} & \overset{b^{\sigma\cdot\Lambda}}{\to} & V_{\sigma\cdot\alpha} \end{array} \qquad \text{(SA.4)}$$

$\blacksquare$

Let us 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 $k = 3$, 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 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 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 $\cap_{\sigma\in\Sigma_k} E^\sigma = \cap_{\tau\in T} E^\sigma$ 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 $f_b \in \mathrm{Gm}^{1,k}(E)$ and $\tau\in\Sigma_k$. Then* (2) *implies* (1)*:*

(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.*

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.

$$\begin{aligned}\mathrm{id}_{\alpha,\tau\cdot\alpha}(v_\alpha) + \sum_{\Lambda\in\mathcal{P}(\alpha)} b^\Lambda(v_{\Lambda^1},\ldots,v_{\Lambda^l})) &= v_{\tau\cdot\alpha} + \sum_{\Omega\in\mathcal{P}(\tau\cdot\alpha)} b^\Omega(v_{\Omega^1},\ldots,v_{\Omega^l}) \\ &= v_{\tau\cdot\alpha} + \sum_{\Lambda'\in\mathcal{P}(\alpha)} b^{\tau\cdot\Lambda'}(v_{(\tau\cdot\Lambda')^1},\ldots,v_{(\tau\cdot\Lambda')^l})\end{aligned}$$

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 $\alpha\in I_k$,

$$\sum_{\Lambda\in\mathcal{P}(\alpha)} \mathrm{id}_{\alpha,\tau\cdot\alpha}(b^\Lambda(v_{\Lambda^1},\ldots,v_{\Lambda^l})) = \sum_{\Lambda'\in\mathcal{P}(\alpha)} b^{\tau\cdot\Lambda'}(v_{(\tau\cdot\Lambda')^1},\ldots,v_{(\tau\cdot\Lambda')^l}) \tag{SA.5}$$

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 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

$$R^{\sigma,b}\cdot L = \sigma\cdot L. \tag{SA.6}$$

This element is called the *curvature operator* and is given by

$$R^{\sigma,L} = \sigma\circ f_b\circ\sigma^{-1}\circ(f_b)^{-1} = f_{\sigma\cdot b}\circ(f_b)^{-1}. \tag{SA.7}$$

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. $\sigma \cdot L^0 = L^0$. If $f_b \in \mathrm{Gm}^{j,k}(E)$, then also $\Omega^{\sigma,b} \in \mathrm{Gm}^{j,k}(E)$. Directly from the definition of the curvature operators, we get the *cocycle relations*

$$R^{\sigma\circ\tau} = \sigma \circ R^\tau \circ \sigma^{-1} \circ R^\sigma, \qquad R^{\mathrm{id}} = \mathrm{id}. \tag{SA.8}$$

The elementary matrix coefficients of R with respect to L,

$$R^\Lambda := (R^{\sigma,L})^\Lambda : V_\Lambda \to V_{\underline{\Lambda}} \tag{SA.9}$$

are called *curvature forms (of σ, L)*.

SA.7. *Symmetrizability.* We say that a 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 $l(\Lambda) < k$, or equivalently,

$$\forall \Lambda : l(\Lambda) < k \Rightarrow (\sigma \cdot b)^\Lambda = b^\Lambda.$$

Lemma SA.8. *Assume $L = L^b$ is symmetrizable. Then the orbit $\mathcal{O}_L := \mathrm{Gm}^{k-1,k}(E).L$ is an affine space over $\mathbb{K}$. 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 both $f_{\sigma\cdot c}$ and $R^{\sigma,L}$ belong to $\mathrm{Gm}^{k-1,k}(E)$. 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". We say that the shift is *positive* if $i > j$ ("base index > direction index"; this is the natural way shifts arise in differential geometry). More generally, one could define shifts into a general direction $B \in I_k$.

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*

$$S := S_{ij} : E \to E, \quad S(v_\alpha) = \mathrm{id}_{\alpha, s(\alpha)}(v_\alpha) = \begin{cases} v_\alpha & \text{if } \alpha_i = 0, \\ \mathrm{id}_{\alpha,\alpha+e_j}(v_\alpha) & \text{if } \alpha_i = 1, \alpha_j = 0, \\ 0 & \text{if } \alpha_i = 1, \alpha_j = 1. \end{cases} \tag{SA.16}$$

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 $S \circ S = p_{1-e_i}$ is the projection given by (MA.18) corresponding to $\beta = 1 - e_i$, and that S_{ij} preserves the factors of the decomposition $E = \widetilde{V}_{1-e_i} \oplus H_i$ (cf. Eqn. (MA.20)); 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:*

$$\begin{array}{ccc} V_{\Lambda^1} \times \ldots \times V_{\Lambda^l} & \stackrel{b^\Lambda}{\to} & V_\alpha \\ S \times \ldots \times S \downarrow & & \downarrow S \\ V_{\Lambda^1} \times \ldots \times V_{\Lambda^m \cup \{i\}} \times \ldots \times V_{\Lambda^l} & \stackrel{b^{s \cdot \Lambda}}{\to} & V_{s\cdot\alpha} \end{array} \tag{SA.18}$$

where m is the (unique) index such that $i \in \Lambda^m$. In this case, $s\Lambda = \{s\Lambda^1, \ldots, s\Lambda^l\} = \{\Lambda^1, \ldots, \Lambda^m \cup \{i\}, \ldots, \Lambda^l\}$ 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,

$$\begin{aligned} f_b(S_{ij}(v)) &= \sum_\alpha \sum_{\Lambda \in \mathcal{P}(\alpha)} b^\Lambda((S_{ij}v)_{\Lambda^1}, \ldots, (S_{ij}v)_{\Lambda^l}) \\ &= \sum_{\substack{\alpha \\ \alpha_i = 0}} b^\Lambda(v_{\Lambda^1}, \ldots, v_{\Lambda^l}) + \sum_{\substack{\alpha \\ \alpha_i = 1 = \alpha_j}} b^\Lambda((S_{ij}v)_{\Lambda^1}, \ldots, (S_{ij}v)_{\Lambda^l}) =: A + B, \end{aligned}$$

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 $r = r(\Lambda)$ 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 $i, j \in \Lambda^m$; then $b^\Lambda((S_{ij}v)_{\Lambda^1}, \ldots, (S_{ij}v)_{\Lambda^l}) = b^\Lambda(v_{\Lambda^1}, \ldots, \mathrm{id}_{\Lambda^m - e_j, \Lambda^m}(v_{\Lambda^m - e_j}), \ldots, v_{\Lambda^l})$, or

(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 $\gamma := \alpha - e_j$, $\Lambda' := \{\Lambda^1, \ldots, \Lambda^m - e_j, \ldots, \Lambda^l\}$; 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 $\mathrm{Gm}^{1,k}(E)$:

$$\begin{aligned} \mathrm{Shi}^{l,k}(E) &:= \{f_b \in \mathrm{Gm}^{l,k}(E) \,|\, \forall i \neq j : S_{ij} \circ f_b = f_b \circ S_{ij}\}, \\ \mathrm{Shi}^{l,k}_+(E) &:= \{f_b \in \mathrm{Gm}^{l,k}(E) \,|\, \forall i > j : S_{ij} \circ f_b = f_b \circ S_{ij}\}. \end{aligned} \tag{SA.19}$$

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 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 $k = 3$:

$$
\begin{array}{ccccccccccc}
 & & (\{2\},\{3,1\}) & & & & (\{2,3\},\{1\}) & & & & (\{3\},\{1,2\}) \\
\overset{s_{31}}{\nearrow} & & & & \overset{s_{13}}{\nwarrow} \quad \overset{s_{23}}{\nearrow} & & & & \overset{s_{32}}{\nwarrow} \quad \overset{s_{12}}{\nearrow} & & \overset{s_{21}}{\nwarrow} \\
(\{2\},\{3\}) & & \overset{(13)}{\leftrightarrow} & & (\{2\},\{1\}) & \overset{(23)}{\leftrightarrow} & & & (\{3\},\{1\}) & \overset{(12)}{\leftrightarrow} & (\{3\},\{2\})
\end{array}
$$

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 $l(\Lambda) = 2$, all bilinear maps $b^{\sigma.\Lambda}$, $\sigma \in \Sigma_3$, are conjugate to b^Λ under σ, and according to (SA.10), this means that f_b is symmetrizable. This argument can be generalized to arbitrary $k > 2$ and Λ with $l(\Lambda) < k$ because for all such Λ there exists Ω such that one of the relations $\Lambda \overset{s_{ij}}{\to} \Omega$ or $\Omega \overset{s_{ij}}{\to} \Lambda$ 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 $b^{\tau.\Lambda}$ 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 $l(\Lambda) = l$ 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" $\{\{1\}, \ldots, \{l\}\}$ 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

$$f_b(v) = \sum_\alpha \varepsilon^\alpha \Big(\sum_{j=1}^{|\alpha|} \sum_{\Lambda \in \mathcal{P}_j(\alpha)} d^j(v_\Lambda) \Big).$$

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* $d^l : V^l \to V$ *such that, if* $l(\Lambda) = l$,

$$b^\Lambda(\varepsilon^{\Lambda_1} v_{\Lambda_1}, \ldots, \varepsilon^{\Lambda_l} v_{\Lambda_l}) = \varepsilon^{\underline{\Lambda}} d^l(v_{\Lambda_1}, \ldots, v_{\Lambda_l}).$$

(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* Formula (7.19)*:*

$$f_b(v) = \sum_\alpha \sum_{\Lambda \in \mathcal{P}(\alpha)} (d^l f)(0)(v_\Lambda)$$

where $(d^l f)(0)$ *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 by permutations and shifts every partition Λ can be transformed into the the standard partition $\widehat{\Lambda} := \{\{1\},\ldots,\{l\}\}$ of length $l = l(\Lambda)$. Then we let $d^l := b^{\widehat{\Lambda}}$. 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 (MG.15 and MG.8) 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

$$F(x) := \sum_{l=1}^{k} \frac{1}{l!} d^l(x,\ldots,x).$$

Then we have $(d^l F)(0) = d^l$, and hence F satisfies the requirement from (2). Conversely, given F, we let $d^l := (d^l F)(0)$ and then define f_b as in Part (1). ∎

Summing up, in characteristic zero, the group $\mathrm{Vsm}^{1,k}(E) \cap \mathrm{Shi}^{1,k}(E)$ can be identified, via the k-th order tangent functor, with the group $S^\times$ of invertible elements of the semigroup

$$S = \mathrm{Pol}(V,V)_0 \mod I_k$$

of polynomial self-maps of V fixing the origin quotiented by polynomials vanishing at 0 of order at least k. This is essentially the *jet group* G^k from [KMS93, Ch. 13].

PG. Polynomial groups

In the following, we will use the notion of *polynomial mappings between modules over a ring* $\mathbb{K}$. If the integers are invertible in $\mathbb{K}$, then this concept is the standard one (polynomial maps come from sums of multilinear maps which moreover may be assumed to be symmetric); otherwise there are several concepts which may slightly differ from each other (see Appendix A of [BGN04]). In order to keep the text in reasonable bounds, we rely on the reader to adapt the following to the concept he prefers ; at any rate, the main results concerning the exponential map are only valid if the integers are invertible, and in this case there is no ambiguity on the concept of polynomial mappings.

PG.1. *Definition of polynomial groups.* A *polynomial group (over a ring* $\mathbb{K}$*) of degree at most* k (where $k \in \mathbb{N}$) is a $\mathbb{K}$-module M together with a group structure (M, m, i, e) such that $e = 0$, the product map $m : M \times M \to M$ is polynomial, and all iterated product maps

$$m^{(j)} : M^j \to M, \quad (x_1, \ldots, x_j) \mapsto x_1 \cdots x_j \tag{PG.1}$$

(which are polynomial maps since so is $m = m^{(2)}$) are polynomial maps of degree bounded by k. (We will see below that this implies that the inversion map also is polynomial.)

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 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 $m^{(j)}$.

(2) For any Lie group G, the group $(T^k G)_e$, 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

$$m(x, y) = \sum_{i=0}^{k} m_i(x, y) = \sum_{p+q \le k} m_{p,q}(x, y) \tag{PG.3}$$

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

$$\begin{aligned} x^{-1} &= i(x) = -x + m_{1,1}(x,x) \operatorname{mod} \deg 3 \\ xyx^{-1} &= y + m_{1,1}(x,y) - m_{1,1}(y,x) \operatorname{mod} \deg 3 \\ xyx^{-1}y^{-1} &= m_{1,1}(x,y) - m_{1,1}(y,x) \operatorname{mod} \deg 3, \end{aligned} \tag{PG.4}$$

(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. Let us describe this Lie bracket more explicitly in case of the examples 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:

$$\deg(\Omega|\Lambda) := \operatorname{card}\{\omega \in \Omega |\, \omega \notin \Lambda\},$$

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,

$$m^{(j)} = \sum_{\substack{p_1,\ldots,p_j \geq 0 \\ p_1+\cdots+p_j \leq k}} m^{(j)}_{p_1,\ldots,p_j},$$

(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

$$\psi_j : M \to M, \quad x \mapsto \sum_{p_1,\ldots,p_j>0} m^{(j)}_{p_1,\ldots,p_j}(x,\ldots,x). \tag{PG.6}$$

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

$$\begin{aligned}
x^2 &= m(x,x) = x + x + \sum_{i>1} m_i(x,x) = 2\psi_1(x) + \psi_2(x),\\
x^3 &= m^{(3)}(x,x,x) = \sum_{p_1,p_2,p_3\geq 0} m^{(3)}_{p_1,p_2,p_3}(x,x,x)\\
&= \Big(\sum_{p_1,p_2,p_3\neq 0} + \sum_{p_1=0,p_2\neq 0,p_3\neq 0} + \sum_{p_2=0,p_1\neq 0,p_3\neq 0} + \sum_{p_3=0,p_2\neq 0,p_1\neq 0} +\\
&\quad \sum_{p_1=0,p_2=0,p_3\neq 0} + \sum_{p_1=0,p_3=0,p_2\neq 0} + \sum_{p_2=0,p_3=0,p_1\neq 0} + \sum_{p_1=p_2=p_3=0} \Big)(m^{(3)}_{p_1,p_2,p_3}(x,x,x))\\
&= \psi_3(x) + 3\psi_2(x) + 3\psi_1(x) + \psi_0(x)\\
x^n &= \sum_{j=0}^{k} \binom{n}{j} \psi_j(x).
\end{aligned}$$

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 the integers are invertible in $\mathbb{K}$. 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 $t,s \in \mathbb{N}$. But both sides of this equality depend polynomially on $(t,s) \in \mathbb{K}^2$, and hence they agree for all $t,s \in \mathbb{K}$ (here we use that, by our assumption on $\mathbb{K}$, the cardinality of $\mathbb{K}^\times$ is infinite). 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 $f'(0)$ is not equal to x but is given by the linear term in

$$\begin{aligned}
f_x(t) &= tx + \frac{t(t-1)}{2}\psi_2(x) + \frac{t(t-1)(t-2)}{3!}\psi_3(x) + \ldots\\
&= t(x - \frac{1}{2}\psi_2(x) + \frac{1}{3}\psi_3(x) \pm \ldots) \bmod t^2
\end{aligned}.$$

Thus, in order to define "the one-parameter subgroup in direction x", all we need is to invert the polynomial $\sum_{p=1}^{k} \frac{(-1)^{p-1}}{p} \psi_p(x)$. This is essentially the content of the following theorem:

Theorem PG.6. *Assume M is a polynomial group over a base ring $\mathbb{K}$ such that the integers are invertible in $\mathbb{K}$. Then there exists a unique polynomial $\exp : M \to M$ such that*

(1) *for all $X \in M$ and $n \in \mathbb{Z}$, $\exp(nX) = (\exp X)^n$,*

(2) *the linear term of $\exp$ is the identity map of M.*

The polynomial $\exp$ is bijective, and its inverse map $\log : M \to M$ 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 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 $\log \circ \exp = \mathrm{id}_M$. 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

$$\varphi_{1,m}(x) = \sum_{p \le m} \frac{(-1)^{p-1}}{p} \psi_{p,m}(x), \qquad \varphi_{m,m}(x) = \frac{1}{m!} \psi_{m,m}(x).$$

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. 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. On the other hand, we have the impression that, with the necessary modifications, our arguments should carry over to the case of power series instead of polynomials. This would have the advantage that the existence of an exponential is proved by "elementary" methods which are close to the usual notions of analysis on Lie groups.

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 $l(\Lambda) > 1$,

$$\exp(X^b)^\Lambda = b^\Lambda + \sum_{j=2}^{k} \frac{1}{j!} \sum_{\substack{\Omega^{(1)} \prec \ldots \prec \Omega^{(j)} = \Lambda \\ \deg(\Omega^{(i)} | \Omega^{(i+1)}) = 1}} b^{\Omega^{(1)}} \circ b^{\Omega^{(1)}|\Omega^{(2)}} \circ \ldots \circ b^{\Omega^{(j-1)}|\Lambda}. \qquad \text{(PG.10)}$$

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 "jet group" $\mathrm{Vsm}^{1,k}(E) \cap \mathrm{Gm}^{1,k}(E)$ (cf. [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 grou $\mathrm{Gm}^{1,k}(E)$ 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 $[\mathrm{Gm}^{i,k}(E), \mathrm{Gm}^{j,k}(E)] \subset \mathrm{Gm}^{i+j,k}(E)$. Finally, we note that in all three examples the group M is nilpotent. One should think that this is a general property of polynomial groups, but this is not obvious.

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. ■

References

[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.

[BGN04] Bertram, W., H. Glöckner and K.-H. Neeb, *Differential Calculus, manifolds and Lie groups over arbitrary infinite fields*, Expo. Math. **22** (2004), 213–282. arXiv: math.GM/0303300.

[Bou72] Bourbaki, N. "Groupes et algèbres de Lie. Chapitres 1 – 3.", Hermann, Paris 1972.

[D73] Dieudonné, J., "Introduction to the Theory of Formal Groups", Marcel Dekker, New York 1973.

[Gl03] Glöckner, H., *Tensor products in the category of topological vector spaces are not associative*, to appear: Comment. Math. Univ. Carol.; arXiv:math.FA/0307041.

[Haz78] Hazewinkel, M., "Formal Groups and Applications", Academic Press, New York 1978.

[KMS93] Kolář, I., Michor P. W. and J. Slovák, "Natural Operations in Differential Geometry", Springer-Verlag, Berlin 1993.

[P62] Pohl, W.F., *Differential Geometry of Higher Order*, Topology **1** (1962), 169–211.

[Se65] Serre, J.-P., "Lie Algebras and Lie Groups", Benjamin, New York 1965.