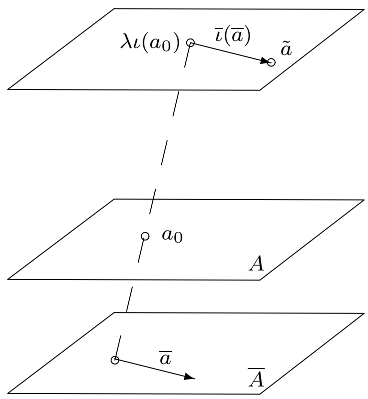
# Berwald-type connections in time-dependent mechanics and dynamics on affine Lie algebroids

Tom Mestdag




Proefschrift ingediend aan de Faculteit Wetenschappen
tot het behalen van de graad van
Doctor in de Wetenschappen: Wiskunde

Promotor: Prof. Dr. W. Sarlet

Universiteit Gent
Faculteit Wetenschappen
Vakgroep Wiskundige Natuurkunde en Sterrenkunde
Academiejaar 2002-2003

# *Woord vooraf*

Het voorwoord is vaak het eerste, en ook het enige deel van een verhandeling dat door iedereen gelezen wordt. Het is dan ook de ideale gelegenheid om even de aandacht van de lezers te vestigen op enkele personen die ik veel verschuldigd ben.

In de eerste plaats wil ik mijn promotor Willy Sarlet vermelden. Ik ben Willy zeer erkentelijk omdat hij de voorbije jaren altijd klaar stond om met veel geduld de zwakke plekken in mijn redeneringen aan te wijzen. Zijn oplossend vermogen tijdens onze samenwerking bepaalde voor een belangrijk deel de uiteindelijke inhoud van dit proefschrift.

Verder wil ik niet nalaten om ook Frans Cantrijn, József Szilasi en Eduardo Martínez te bedanken voor onze vele discussies en voor hun oprechte interesse in mijn werk; alsook alle collega's van de vakgroep, met in het bijzonder mijn dichtste buur Bavo, voor de gezellige werksfeer.

Ik wil zeker mijn familie en vrienden in dit dankwoord niet vergeten, bovenal mijn ouders voor hun onvoorwaardelijke steun in alles wat ik onderneem. Tot slot wens ik Viki te bedanken omdat ze steeds met veel brio de noodzakelijke functie van klankbord op zich nam.

Tom Mestdag, 3 mei 2003

# *Contents*

# *Introduction*

Berwald's covariant derivative.[1] Before entering into the contents of this dissertation, it is useful to explain the original context in which 'the Berwald-type connection' was defined. The Berwald connection is an important concept in Finsler geometry. In many ways, Finsler geometry can be regarded as a generalisation of Riemann geometry. Let $M$ be a manifold and $c : t \in [a,b] \mapsto x^i(t)$ a curve in it. In his famous inaugural lecture of 1854 at the Göttingen University [73] (see also [82], Ch. 4A), Riemann himself proposed (translated into the terminology of today) the following idea for defining a concept of length of such a curve (between the endpoints $a$ and $b$):

$$s = \int_a^b F\big(x(t), \frac{dx}{dt}(t)\big) dt. \tag{1}$$

We will not enter the discussion on the precise differentiability assumptions which should be imposed on the function $F$. Here, it is only important to know that, to keep the definition of length independent of the choice of the parameter $t$, the function $F$ is supposed to be homogeneous of degree 1 in its velocity components

$$F(x, py) = p\ F(x,y), \qquad \text{for any } p > 0.$$

In addition, we will assume that $(g_{ij}) = \frac{1}{2}\left(\frac{\partial^2 F^2}{\partial y^i \partial y^j}\right)$ defines, at any point, a non-singular matrix.

Apart from the 'Riemann' case $F(x,y) = \sqrt{a_{ij}(x) y^i y^j}$, Riemann suggested in his lecture the study of the fourth root of a function of degree four. He immediately added, however, that he thought that the investigation of such a more general class would merely be time-consuming and would not bring any new insights into the geometry of metric spaces. So, shortly after his talk, Riemann and his contemporaries forgot about Riemann's proposal. More

[1] My account on the history of Finsler geometry is, to a large extent, based on [58] and [82]. Details on Berwald's life and work can be found in [71] and [31].

than sixty years later, in 1918, Finsler studied some subclasses of metrics which were more general than Riemann metrics in his thesis [30]. However, he does not mention the most general form (1) and Matsumoto [58] thinks that Finsler was at that time not aware of the suggestion made by Riemann. Later on, Finsler felt quite uncomfortable with the name 'Finsler geometry' which was adopted for the general case (1) (for the first time by Taylor). In the following decades, it was rather Cartan and Berwald who were the leading figures in Finsler geometry.

Ludwig Berwald was born in 1883 in Prague. He contributed to many domains in differential geometry. His most famous contribution to Finsler geometry is the paper [8]; we will come back to this work later on. Next to his connection, Berwald's name also survived in the form of e.g. 'Berwald space', 'Berwald curvature' and 'Berwald inequalities'.

Let us come back to the length function (1). Having defined distance, we check (for fixed endpoints) which curves on the manifold have the shortest length. These curves are the geodesics of the Finsler manifold. They are solutions of the variational problem $\delta \int_a^b F(x, \frac{dx}{dt})dt = 0$, or equivalently, solutions of the differential equation (parameterised by arclength)

$$\frac{d^2x^i}{ds^2} + \gamma^i_{jk}\big(x, \frac{dx}{ds}\big)\frac{dx^j}{ds}\frac{dx^k}{ds} = 0,$$

where $\gamma^i_{jk} = \frac{1}{2}g^{ir}\big(\frac{\partial g_{jr}}{\partial x^k} + \frac{\partial g_{kr}}{\partial x^j} - \frac{\partial g_{jk}}{\partial x^r}\big)$. Throughout this section, we refer to e.g. [67] for proofs and calculations. Important to notice here is that the functions $\gamma^i_{jk}$ formally have the same expression as the Christoffel symbols, although $g_{ij}(x,y)$ is not a Riemannian metric.

Let us take a step backwards and suppose that $a_{ij}(x)$ is a Riemann tensor and $\Gamma^i_{jk}(x)$ its Christoffel functions. It is well known that the functions $\Gamma^i_{jk}(x)$ are the connection coefficients of a linear connection on $M$, the Levi–Civita connection. After a coordinate change $\hat{x}^i = \hat{x}^i(x)$, connection coefficients transform according to the law

$$\hat{\Gamma}^a_{bc} = \Gamma^i_{jk}\frac{\partial \hat{x}^a}{\partial x^i}\frac{\partial x^j}{\partial \hat{x}^b}\frac{\partial x^k}{\partial \hat{x}^c} + \frac{\partial \hat{x}^a}{\partial x^i}\frac{\partial^2 x^i}{\partial \hat{x}^c \partial \hat{x}^b}. \tag{2}$$

Linear connections are used to define covariant differentiation. A set of functions $X^i(x)$ which transform according to the rule

$$\hat{X}^a = X^i \frac{\partial \hat{x}^a}{\partial x^i} \tag{3}$$

constitute a vector field on $M$. If $X^i(x)$ is a vector field on $M$, then

$$X^i{}_{,j} = \frac{\partial X^i}{\partial x^j} + X^r \Gamma^i_{rj} \tag{4}$$

is a type (1,1)-tensor field on $M$.

Berwald is the first to introduce the notion of connection in Finsler geometry [8]. The way he proceeds goes as follows. Suppose that we have a set of functions $\Gamma^k_{ij}(x,y)$, satisfying formally the transformation law (2). It is then easy to find a rule of covariant differentiation for (vertical) vector fields $X^i(x,y)$ (which transform also according to (3)): one can simply adjust the rule (4) to

$$X^i{}_{,j} = \frac{\partial X^i}{\partial x^j} - \frac{\partial X^i}{\partial y^r}\Gamma^r_{0j} + X^r \Gamma^i_{rj}, \tag{5}$$

where $\Gamma^r_{0j} = y^i \Gamma^r_{ij}$.

But how does one find functions which satisfy the transformation rule (2)? Obviously there is a link between covariant derivation and the geodesic equation. The functions $\gamma^k_{ij}$, however, do *not* transform according to the rule (2). Berwald's idea to overcome this problem was the following. Put $G^i(x,y) = \frac{1}{2}\gamma^i_{jk}(x,y)y^j y^k$, and differentiate twice: $G^i_j = \frac{\partial G^i}{\partial y^j}$, $G^i_{jk} = \frac{\partial G^i_j}{\partial y^k}$. It can be shown that, due to the homogeneity of $F$, the geodesic equation can equivalently be rewritten in the form

$$\frac{d^2x^i}{ds^2} + G^i_{jk}\big(x, \frac{dx}{ds}\big)\frac{dx^j}{ds}\frac{dx^k}{ds} = 0.$$

Although nothing seems to have changed, we have entered a different world: the functions $G^i_{jk}$ do satisfy (2) now.

Having defined covariant differentiation, Berwald also introduced a corresponding rule of parallel transport (in such a way that autoparallel curves coincide with the geodesics of the Finsler space). Almost at the same time, a different way to introduce parallel displacement in Finsler geometry was found, independently from each other, by Synge [83] and Taylor [87]. Wondering about the geometric meaning of the functions $\gamma^i_{jk}$ in the geodesic equation, they introduced a new concept, that of a 'non-linear connection'. A non-linear connection (or 'horizontal structure') is a set of quantities $\Gamma^i_j(x,y)$ transforming according to the rule

$$\hat{\Gamma}^a_b(\hat{x},\hat{y}) = \Gamma^i_j(x,y)\frac{\partial \hat{x}^a}{\partial x^i}\frac{\partial x^j}{\partial \hat{x}^b} + \frac{\partial \hat{x}^a}{\partial x^i}\frac{\partial^2 x^i}{\partial \hat{x}^c \partial \hat{x}^b}\hat{y}^c. \tag{6}$$

In the approach of Synge and Taylor, a vector field $X^i$ along a curve $c$ is said to be parallel if

$$\frac{dX^i}{ds} + \Gamma^i_j\big(x, \frac{dx}{ds}\big) X^j = 0.$$

The terminology 'non-linear' connection was used to point to the fact that functions can be covariantly differentiated ($f_{;i} = \frac{\partial f}{\partial x^i} - \frac{\partial f}{\partial y^r}\Gamma^r_i$), but there exists no covariant derivation rule for vector fields.

The point that we would like to emphasise is the following: the non-linear connection of Synge and Taylor is not completely independent from the covariant derivative of Berwald: it turns out that $\Gamma^i_j = G^i_j$. Therefore we can conclude that *the linear connection of Berwald is a kind of 'linearisation' of the non-linear connection of Synge and Taylor!*

It is not clear whether Synge, Taylor and Berwald were aware of the above relation between their two theories. The papers of Synge [83] and Taylor [87] appeared in the same journal in 1925. After completion of his own article [8] in the same year, Berwald added a small note at the end of his introduction: *" Die vorliegende Abhandlung war schon vollendet, als eine Arbeit des Herrn J.L. Synge erschien ... Herr Synge betrachtet auch eine Parallelübertragung eines willkürlichen Vektors längs einer gegebenen Kurve; diese ist jedoch verschieden von der Parallelübertragung, die im Innsbrucker Vortrage und hier verwendet wird ... Das über die Arbeit von Synge Gesagte gilt auch von einer Abhandlung von J.H. Taylor, die während des Druckes der vorliegende Untersuchung publiziert wurde ...".* Berwald does not mention how the non-linear connection fits in his approach to parallel transport. In fact, according to Matsumoto [58], the three authors paid no longer attention to the concept of non-linear connection after the appearance of their papers. In present-day approaches to Finsler geometry (based on the geometry of the tangent bundle $\tau_M : TM \to M$), however, the above non-linear connection on $\tau_M$ plays an essential role. The previous observation can even be cast in a far more general statement: *any non-linear connection on $\tau_M$ generates a covariant derivative, said to be of Berwald type.* This fact can easily be deduced as follows: if we have functions $\Gamma^i_j$ satisfying (6) at our disposal, then their fibre derivatives $\frac{\partial \Gamma^i_j}{\partial y^k}$ satisfy (2). Matsumoto [58] states that Kawaguchi [40] is the first to emphasise the importance of non-linear connections in Finsler geometry. Finally, for completeness, we should mention that many other Finsler geometers found different rules of covariant differentiation. The most important one was found by Cartan [15] in 1934.

Brief outline of the dissertation. In the first chapter we briefly mention a modern framework in which all of the covariant derivatives in Finsler geometry can be studied within the context of linear connections D on the pullback bundle $\tau_M^*\tau_M$. Following Crampin [21], we indicate how an arbitrary non-linear connection on $\tau_M$ leads to a Berwald-type connection on $\tau_M^*\tau_M$. It is fairly easy to describe how the Berwald connection from Finsler geometry fits in this approach. Given a non-linear connection on $\tau_M$, a linear connection on $\tau_M^*\tau_M$ can be regarded as being composed of two covariant derivatives: a 'horizontal' one $\mathrm{D}^H$ and a 'vertical' one $\mathrm{D}^V$. In the case of the Berwald connection, $\mathrm{D}^H$ and $\mathrm{D}^V$ can be determined as follows. First, we let the horizontal covariant derivative $\mathrm{D}^H X$ of vector fields $X$ along $\tau_M$ coincide with Berwald's original rule, that is (5) with $\Gamma^i_{jk} = G^i_{jk}$. We add to this that $\mathrm{D}^V X$ is the tensor with components

$$X^i{}_{;j} = \frac{\partial X^i}{\partial y^j}.$$

A linear connection on $\tau_M^*\tau_M$ with this vertical covariant derivative is said 'to have complete parallelism in the fibres'.

Within the class of Berwald-type connections, a particular case of interest is when the non-linear connection is derived from a system of *autonomous* second-order differential equations (autonomous Sodes, for short). Such a Sode is governed by a vector field on $TM$. In the geometric study of certain qualitative aspects of a Sode, Berwald-type connections have proved to be very useful. Of further special intrest for mechanics is the subclass of Lagrangian-type Sodes. Schematically, the Berwald-type connection associated to a regular autonomous Lagrangian, can be defined by means of the following consecutive steps:

$$L \in C^\infty(TM) \xrightarrow{(1)} \text{Sode } \Gamma \in \mathcal{X}(TM) \xrightarrow{(2)} \text{n-l. con. on } \tau_M \xrightarrow{(3)} \text{D of Berwald-type}$$

An important part of this dissertation deals with a generalisation of the main ideas that lie at the basis of this sequence. In fact, we will explore generalisations in two different directions.

1. First, we shall look at an extension of the above sequence to *time-dependent* Lagrangians. The framework of time-dependent mechanics is different from the above autonomous situation. In order to include

time-dependent coordinate transformations, the manifold $M$ is supposed to be fibred over $I\!R$ and the carrier space is not a tangent bundle, but the first jet extension $\pi_M : J^1M \to M$ of $M \to I\!R$. A time-dependent Lagrangian is therefore a function on $J^1M$. Observe that the structure of $\tau_M : TM \to M$ and $\pi_M : J^1M \to M$ is different: the first is a vector bundle, the second an affine bundle.

2. Let $\pi : E \to M$ be an arbitrary affine bundle. We will introduce the concept of a 'Lie algebroid structure on $\pi$'. For an affine bundle with such a Lie algebroid structure, the classical notion of Lagrangian systems can be extended to Lagrangians $L$ on $E$.

*1. Berwald-type connections in time-dependent mechanics.* We shall study geometric structures that live on the pullback bundle $\pi_M^*\tau_M : \pi_M^*TM \to J^1M$. The set $\mathcal{X}(\pi_M)$ of sections of $\pi_M^*\tau_M$ contains a canonical element $\mathbf{T}$ which is such that any other section $X$, in the following called a vector field along $\pi_M$, can be decomposed in a $\mathbf{T}$-component and a component $\overline{X}$ that is annihilated by $dt$.

$$X = X^0\mathbf{T} + \overline{X}.$$

We will refer to $X^0\mathbf{T}$ as the time-component. Typical for the time-dependent extension of geometrical structures is the following methodology: the autonomous structure is translated to the component $\overline{X}$ and one further tries to find a natural way to fill in the remaining freedom for $X^0\mathbf{T}$.

In the SECOND CHAPTER, we recall how any regular Lagrangian $L \in C^\infty(J^1M)$ leads to a time-dependent SODE (which is in this case a vector field on $J^1M$). Further, we mention how a time-dependent SODE gives rise to a non-linear connection on $\pi_M$. The first two steps in the sequence are well-known; in the third step some extra complications occur. Independently from each other, three different constructions have appeared in the literature which associate a linear connection to a non-linear connection on $\pi_M$ (see [9], [57] and [23]). All of these can be regarded as time-dependent extensions of Berwald-type connections in some sense. However, they differ from each other in two ways. First of all, the freedom caused by the time component is being fixed in a different way. Secondly, the carrier space is not the same for the three connections: the linear connections of [9] and [57] live on $\tau_{J^1M} : TJ^1M \to J^1M$, while the one in [23] is a linear connection on $\pi_M^*\tau_M$. Therefore, it is of interest to explore a scheme in which the three constructions can be compared and it is useful to enter the discussion whether one of them should be preferred over the others. While doing so, we will find natural restrictions which one can impose on the class of acceptable Berwald-type connections. This will lead us to the introduction of a new Berwald-type connection on $\pi_M^*\tau_M$, different from the one in [23], which somehow is the most obvious candidate for satisfying these restrictions, and therefore is called the 'optimal' Berwald-type connection. In CHAPTER 3 we sketch some areas of application.

*2. Berwald-type connections in the dynamics on affine Lie algebroids.* Bits of information concerning the second generalisation of the sequence are spread throughout the literature. Usually, however, these results refer to a vector bundle, rather then to an affine bundle. For example, in the last years, a number of papers have appeared which investigate dynamical systems on a Lie algebroid (see e.g. [13, 14, 16, 45, 48, 49, 72, 91]). A Lie algebroid is a vector bundle $\tau : \mathsf{V} \to M$ with the property that its sections constitute a real Lie algebra (with structure functions $\mathsf{C}_{ab}^{\ c}$ say). Each section is 'anchored' on a vector field, by means of a linear bundle map $\varrho : \mathsf{V} \to TM$, which is further supposed to induce a Lie algebra homomorphism. We are in particular interested in the notion of Lagrangian systems on such a Lie algebroid, which is due to Weinstein [91]. If $L \in C^\infty(\mathsf{V})$, then Weinstein's extension of the Lagrangian formalism leads to dynamical systems of the form

$$\begin{aligned} \dot{x}^I &= \varrho_a^I(x)\mathsf{v}^a, \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \mathsf{v}^a}\right) &= \varrho_a^I \frac{\partial L}{\partial x^I} - \mathsf{C}_{ab}^{\ c}\mathsf{v}^b \frac{\partial L}{\partial \mathsf{v}^c}. \end{aligned} \tag{7}$$

Detailed information on Weinstein's construction can be found in the FOURTH CHAPTER.

So, the kind of generalisation of the sequence we have in mind is the following one. Let $\pi : E \to M$ now be an affine bundle with an affine anchor map $\rho : E \to TM$. Based on a 'rudimentary calculus of variations' we have clear indications that for a function $L \in C^\infty(E)$, an affine version of Weinstein's equations should look like

$$\begin{aligned} \dot{x}^I &= \rho_0^I(x) + \rho_\alpha^I(x)y^\alpha, \\ \frac{d}{dt}\left(\frac{\partial L}{\partial y^\alpha}\right) &= \rho_\alpha^I \frac{\partial L}{\partial x^I} - (C_{\alpha\beta}^\gamma y^\beta - C_{0\alpha}^\gamma)\frac{\partial L}{\partial y^\gamma}. \end{aligned} \tag{8}$$

The geometric description of these Lagrangian equations (in CHAPTER 7) will involve the introduction of the concept of a Lie algebroid structure on an affine bundle $\pi$ (in CHAPTER 6).

In [48] Martínez showed how one can associate to the system (7) a 'Lagrangian' section of a kind of 'prolonged' bundle. The same ideas can be applied to the system (8) and are explained in CHAPTER 5. The prolonged bundle in question, $\pi^1 : T^{\tilde{\rho}}E \to E$ is constructed as follows. First, we consider the bi-dual $\tilde{\pi} : \tilde{E} \to M$ of $\pi$. It is a vector bundle which is dual to $\pi^\dagger : E^\dagger \to M$, the fibres of which are the set of affine functions in fibres of $\pi : E \to M$. The bundle $\tilde{\pi}$ contains a copy of $\pi$ and of its underlying vector bundle. The affine map $\rho : E \to TM$ has a natural extension $\tilde{\rho} : \tilde{E} \to TM$. The total space $T^{\tilde{\rho}}E$ of the prolonged bundle now is the total space of the pullback bundle $\tilde{\rho}^*TE$. Its projection $\pi^1$ onto $E$ is the composition of the projection $\tilde{\rho}^1 : \tilde{\rho}^*TE \to TE$ with the tangent bundle projection $\tau_E$.

Let us now come back to the dynamical systems (8). For regular Lagrangians, the equations (8) can be written in the form

$$\begin{aligned} \dot{x}^I &= \rho^I_\alpha(x)y^\alpha + \rho^I_0(x), \\ \dot{y}^\alpha &= f^\alpha(x,y). \end{aligned} \tag{9}$$

These are equations which we call of pseudo-SODE type. If there is a Lie algebroid structure on the affine bundle $\pi$, then for any regular Lagrangian $L \in C^\infty(E)$, we can construct a section $\Gamma$ of $\pi^1$ which is such that the integral curves of the vector field $\tilde{\rho}^1 \circ \Gamma \in \mathcal{X}(E)$ are solutions of the above Lagrangian pseudo-SODE.

So far, we have explained the first step in the schematic sequence mentioned above. We can now bring the connections back into the picture. Related to the second step, the first thing which might come to one's mind is to look at non-linear connections on the affine bundle $\pi$. Again, in the literature mainly results concerning vector bundles are known. The tangent bundle of a vector bundle $\tau : \mathsf{V} \to M$ contains a canonical 'vertical' subbundle $V\mathsf{V}$. A direct complement of $V\mathsf{V}$ in $T\mathsf{V}$ is the horizontal subbundle of a non-linear connection on $\tau$. In [89] Vilms poses the question how one can intrinsically define curvature for such non-linear connections. His solution is related to the third step of the sequence: he introduces a linear connection on $V\mathsf{V}$ which 'linearises' in some sense the original non-linear connection. Using the associated covariant exterior derivative, curvature can then be defined. Similar observations were made by the school of Miron (see e.g. Ch. II and III in [67], although the formalism is different from the one of Vilms). Vilms' construction of Berwald-type connections can easily be extended to affine bundles. In fact, in a small note [90], Vilms later mentioned this possibility.

The construction of Vilms is not appropriate as third step in our programme, however, because it is not clear how the pseudo-SODEs (9) can be related to a non-linear connection on $\pi$. This is the reason why we shall consider the dynamical sections (9) as sections of the prolonged bundle $T^{\tilde{\rho}}E \to E$, rather than as sections of $TE \to E$. We will concentrate on 'horizontality' on $\pi^1$, rather than on $\tau_E$, which is possible because $\pi^1$ contains a well-defined 'vertical' subbundle (consisting of those elements that project by means of $\pi^2$ on the zero vector on $\tilde{E}$). This brings us to the notion of generalised connections on $\pi$. In CHAPTER 8, we show how a pseudo-SODE on an affine Lie algebroid generates a generalised non-linear connection on $\pi$. Generalised connections appear in quite more general situations and have proved to be useful in many domains (see e.g. [28, 42, 43, 44, 88]).

In view of what precedes, we pay particular attention to the subclass of 'affine' generalised connections on an affine bundle $\pi$ and show their relation to 'linear' generalised connections on the bi-dual and the underlying vector bundle. In CHAPTER 9, we define for every generalised connection on $\pi$ two affine generalised connections on $\pi^*\pi$. For this to work, it suffices that $\pi$ is an anchored bundle, not necessarily equipped with the additional structure of a Lie algebroid. However, the case of an affine Lie algebroid makes extra tools available, which for example lead to explicit defining relations for the two associated linear generalised connections on $\pi^*\tilde{\pi}$. We arrive finally at an interesting link with the first few chapters of this work. Indeed, by the fact that we explore in great detail the origin of the differences between these two connections, we reach a better understanding of the nature of the two Berwald-type connections for time-dependent mechanics, discussed in the beginning, because these are indeed the connections we recover when we specialise to the case $\pi = \pi_M$ and $\tilde{\rho}$ is the identity.

REFERENCES. A part of the work presented here has already been published. The SECOND CHAPTER on Berwald-type connections in time-dependent mechanics is based on [62, 61, 76]. The results in CHAPTERS 4, 5 AND 6 on affine Lie algebroids and in CHAPTER 7 on dynamics on affine Lie algebroids have appeared in [56, 77, 78]. The theory on pseudo-SODE connections and affine generalised connections in CHAPTER 8 can be found in [64, 60]. The LAST CHAPTER on generalised Berwald-type connections is based on [63]. The reader who is interested in Finsler geometry can take a look at [65, 66].

# Chapter 1

# *Berwald-type connections – Autonomous case*

## 1.1 Finsler geometry

As we stated in the introduction, we will spend a few words on a modern description of Finsler geometry. Let $\tau_M : TM \to M$ be the tangent bundle of a manifold $M$. If $0 : M \to TM$ denotes the zero section, then $\overset{\circ}{T}M = \bigcup_{m \in M} T_m M \setminus \{0_m\}$ is the (open) submanifold of $TM$ containing all nonzero tangent vectors. We will denote its natural projection on $M$ by $\overset{\circ}{\tau}_M$. In these introductory sections we will suppose that the reader is familiar with most of the basic notions of tangent bundle geometry. For example, two canonical objects play an important role: the Liouville dilation vector field $\Delta$ and the vertical endomorphism $S$. The vertical lift and the complete lift of a vector field $X$ on $M$ will be denoted, respectively, by $X^V$ and $X^C$. In a lot of situations, as additional input, the presence of a non-linear connection on $\tau_M$ will be required. Given a non-linear connection (or horizontal distribution on $TM$), with corresponding projection operators $P_H$ and $P_V$, every vector field $\xi$ on $TM$ uniquely decomposes into a horizontal and vertical lift of vector fields along $\tau_M$, which, as in [23], will be called $\xi_H$ and $\xi_V$, respectively:

$$\xi = {\xi_H}^H + {\xi_V}^V . \tag{1.1}$$

For more detailed information on the geometry of $\tau_M$, see e.g. [19, 35, 36, 41].

**Definition 1.1.** *Let $E : TM \to I\!R$ be a function on $TM$. The couple $(M, E)$ is said to be a Finsler manifold with energy $E$ if*

1. *$E$ is of class $C^1$ on $TM$ and smooth over $\overset{\circ}{T}M$.*
2. *$E$ is positively-homogeneous of degree 2, i.e. $\Delta E = 2E$.*

3. *The two-form* $\omega_E := dd_S E$ *on* $\overset{\circ}{T}M$ *is non-degenerate, i.e.* $E$ *is a regular Lagrangian on* $\overset{\circ}{T}M$.

4. $\forall\, v \in \overset{\circ}{T}M$: $E(v) \geq 0$ *and* $E(0) = 0$.

Then, the function $F$ we met in the introduction is given by $F = \sqrt{2E}$. The tensor $g_E$ along $\overset{\circ}{\tau}_M$, defined on basic vector fields $X, Y \in \mathcal{X}(M)$ by means of $g_E(X,Y) = \omega_E(X^V, Y^C)$, is symmetric and non-degenerate. In the following, $g_E$ is referred to as the *Finsler metric*. In any fibre $T_m M$ the coordinate components of the Finsler metric $g_E$ are given by the Hessian of the energy $E$.

Remark that in the above definition, the smoothness of $E$ is only guaranteed over $\overset{\circ}{T}M$. This restriction is necessary in this context, because condition 2 would otherwise imply that Finsler metrics are 'lifts' of Riemann metrics. Indeed, due to the Euler theorem, it follows that if $E$ is positive homogeneous of degree 2 on $\overset{\circ}{T}M$ and (at least) of class $C^2$ on $TM$, then $E$ is a polynomial of degree 2 on the fibres (see e.g. [85]). As a consequence, the Hessian of $E$ will be constant on the fibres and thus related to a Riemann metric.

Standard textbooks on Finsler geometry are [59, 74], while [1, 6, 85] are recent surveys.

## 1.2 Linear connections in Finsler geometry

In view of the close relation between Finsler geometry and Riemann geometry, it is obvious that an important issue in Finsler geometry is to find a suitable analogue for the (torsion-free and metrical) Levi-Cevita connection of Riemann geometry. The Berwald connection is one of many related linear connections which have been studied in this context, other often discussed connections being for example those attributed to Chern–Rund, Cartan and Hashiguchi (see e.g. [3, 5, 7, 36, 59, 68]). Starting from Definition 1.1, we will show first in a few consecutive steps how these connections are defined.

A vector field $\Gamma \in \mathcal{X}(TM)$ is said to be a *semispray* if it is of class $C^1$ on $TM$, smooth on $\overset{\circ}{T}M$ and if it satisfies $S(\Gamma) = \Delta$. If, in addition, $[\Delta, \Gamma] = \Gamma$, then $\Gamma$ is a *spray*. The smoothness restrictions are only important when $M$

is a Finsler manifold with energy $E$. On any Finsler manifold there exists a *canonical spray* $\Gamma$ which is uniquely determined over $\overset{\circ}{T}M$ by the formula

$$i_\Gamma \omega_E = -dE.$$

Any semispray gives rise to a horizontal distribution by means of the following formula

$$P_H = \tfrac{1}{2}(I - \mathcal{L}_\Gamma S) \tag{1.2}$$

for the corresponding horizontal projection operator (or 'horizontal endomorphism'). If, in particular, $\Gamma$ is the canonical spray of a Finsler manifold, the above constructed operator $P_H$ is called the *Barthel endomorphism* of the Finsler manifold. In the Finsler case $P_H$ will only be smooth over $\overset{\circ}{T}M$. In this dissertation, the differentiability restrictions of the involved geometric objects will not play an essential role. Therefore, we will usually assume them to be smooth over the whole (tangent) manifold. In particular, smooth vector fields $\Gamma$ on $TM$ satisfying the condition $S(\Gamma) = \Delta$ play the role of the semisprays in Finsler geometry. Since their integral curves are solutions of autonomous second-order ordinary differential equations, we refer to such vector fields for short as (autonomous) SODEs. Then, the above construction (1.2) simply shows how one associates to any SODE a non-linear connection on $\tau_M$.

Let us first recall the definition of a *linear connection on a vector bundle* $\mu : P \to M$. It is an $\mathbb{R}$-bilinear map $\mathrm{D} : \mathcal{X}(M) \times \mathrm{Sec}(\mu) \to \mathrm{Sec}(\mu)$, which satisfies the following requirements with respect to multiplication by functions $f$ on $M$:

$$\mathrm{D}_X(f\sigma) = f\mathrm{D}_X\sigma + X(f)\sigma \quad \text{and} \quad \mathrm{D}_{fX}\sigma = f\mathrm{D}_X\sigma$$

($\mathrm{Sec}(\mu)$ denotes the set of sections of $\mu$).

Coming back to the context of Finsler metrics, an illuminating discussion of the relationship between different linear connections used in Finsler geometry has been given by Szilasi [84]. All such connections in Szilasi's account live on the tangent bundle $\tau_{TM} : T(TM) \to TM$ of a tangent bundle $\tau_M : TM \to M$. Let us now take a step backwards and suppose first that an arbitrary horizontal distribution on $TM$ is given. If $J$ is the almost complex structure provided by this horizontal distribution, then a *Finsler connection* on $T(TM) \to TM$ is characterised by $\nabla P_H = \nabla J = 0$. It then follows that also $\nabla S = 0$ and that a Finsler connection is completely determined if one knows the covariant derivative of vertical vectors. Finsler connections correspond to 'normal-d connections' in the terminology of [67].

The standard torsion tensor of a Finsler connection, $T(\xi,\eta) = \nabla_\xi \eta - \nabla_\eta \xi - [\xi,\eta]$, when decomposed into its horizontal and vertical part for various combinations of the arguments, gives rise to six tensor fields of type (1,2) in principle, but one of them is identically zero:

$$\begin{aligned}
\mathcal{A}(\xi,\eta) &= P_H\big(T(P_H\xi, P_H\eta)\big), &\qquad \mathcal{B}(\xi,\eta) &= P_H\big(T(P_H\xi, P_V\eta)\big),\\
\mathcal{R}(\xi,\eta) &= P_V\big(T(P_H\xi, P_H\eta)\big), &\qquad \mathcal{P}(\xi,\eta) &= P_V\big(T(P_H\xi, P_V\eta)\big),\\
\mathcal{S}(\xi,\eta) &= P_V\big(T(P_V\xi, P_V\eta)\big). & & 
\end{aligned} \tag{1.3}$$

Szilasi's account of Berwald connections is partially based on earlier work by Okada [70].

**Proposition 1.2.** *[84] Let $(M,E)$ be a Finsler manifold and suppose that an arbitrary non-linear connection $P_H$ is given.*

*(i) There exists a unique Finsler connection $(\nabla, P_H)$ such that*

*(B1) The $\mathcal{P}$-torsion of $(\nabla, P_H)$ vanishes.*

*(B2) The $\mathcal{B}$-torsion of $(\nabla, P_H)$ vanishes.*

*(ii) If $(\nabla, P_H)$ further satisfies the conditions*

*(B3) $d_{P_H} E = 0$,*

*(B4) $\nabla_{X^H}\Delta = 0$ for all $X \in \mathcal{X}(\tau_M)$,*

*(B5) the $\mathcal{A}$-torsion of $(\nabla, P_H)$ vanishes,*

*then $P_H$ is the Barthel endomorphism of the Finsler manifold. The linear connection $\nabla$ is then the so-called Berwald connection of the Finsler manifold.*

The advantage of Szilasi's approach is that he singles out the Berwald connection of $(M,E)$ in two steps. In the first step, we get the minimal axioms which are needed to construct uniquely a linear connection on $\tau_{TM}$ from a given non-linear connection. Only in the second step the relation between the linear connection, the horizontal projector $P_H$ and the Finsler energy $E$ is specialised. In the following we will be mainly interested in the linear connection satisfying the conditions $(B1-B2)$ w.r.t. the horizontal distribution. Following Szilasi, such a Finsler connection is said to be *of Berwald* *type*. For Berwald-type connections the vertical lifts of basic vector fields (i.e. vector fields on $M$) are parallel with respect to vertical vector fields and $\nabla_{X^H} Y^V = P_V([X^H, Y^V])$. In Szilasi's overview [84], similar axiomatic definitions for the Cartan, Hashiguchi and Chern–Rund connections can be found.

Crampin [21] has pushed our understanding of this matter further ahead by explaining the more concise picture where all connections are constructed on the pullback bundle $\tau_M^*\tau_M : \tau_M^* TM \to TM$ and by putting thereby the Berwald-type connection in the spotlight as the one to which all others can be related. Similar observations were also made by Anastasiei [2]. The key point in Crampin's analysis is the following. The covariant derivative operator $\mathrm{D} : \mathcal{X}(TM) \times \mathcal{X}(\tau_M) \to \mathcal{X}(\tau_M)$, defined by

$$\mathrm{D}_\xi X = [P_H(\xi), X^V]_V + [P_V(\xi), X^H]_H, \tag{1.4}$$

determines the unique linear connection on $\tau_M^*\tau_M$ which has the properties:

(i) the restriction to fibres $T_m M$ is the canonical complete parallelism;

(ii) parallel translation along a horizontal curve is given by a rule of Lie transport.

It is easy to move back to the bigger space $T(TM)$, i.e. to consider another linear connection $\nabla$, this time defined on $T(TM) \to TM$, which is obtained from the one on $\tau_M^*\tau_M$ by "doubling the formulas", as follows:

$$\nabla_\xi X^H = (\mathrm{D}_\xi X)^H, \qquad \nabla_\xi X^V = (\mathrm{D}_\xi X)^V. \tag{1.5}$$

It turns out that the Berwald-type connection corresponds exactly to the lift $\nabla$ of the connection D in (1.4). Therefore, we will also refer to the linear connection (1.4) as the *Berwald-type connection on* $\tau_M^*\tau_M$ determined by the given horizontal distribution. If in particular the non-linear connection is the one canonically associated to a given Sode on $TM$ (by means of (1.2)), then the associated Berwald-type connection has all of its torsion tensor fields equal to zero, except for the one whose vanishing would require that the non-linear connection is flat.

The other type of linear connections referred to at the beginning, although they were originally introduced merely in the framework of Finsler manifolds, can also be given a quite more general meaning (cf. [2, 21, 65]). All they require is one extra tool, namely a metric tensor field $g$ along $\tau_M$. For example, in the case that the horizontal distribution comes from a spray $\Gamma$, Crampin defines *vertical and horizontal Cartan tensors* $\mathcal{C}_V$ and $\mathcal{C}_H$ by the following relations: $\forall X, Y, Z \in \mathcal{X}(\tau_M)$,

$$g(\mathcal{C}_V(X,Y),Z) = \mathrm{D}_{X^V} g\,(Y,Z), \tag{1.6}$$
$$g(\mathcal{C}_H(X,Y),Z) = \mathrm{D}_{X^H} g\,(Y,Z). \tag{1.7}$$

For completeness, we should remark here that it is the type (1,2) tensor field $\mathcal{C}_V$ which is related to the Cartan tensor known in Finsler geometry.

The main idea behind the connections associated to Chern–Rund, Hashiguchi and Cartan stems from various degrees of trying to obtain a metrical connection. Since the difference between two linear connections is tensorial and the decomposition (1.1) of $\xi$ will, in the present context, split this tensor field also in a vertical and horizontal component, denoted by $\delta^V$ and $\delta^H$ in [21], new connections can be derived from the Berwald connection by making assignments for $\delta^V$ and $\delta^H$. In [21], for the further special case that $\Gamma$ is the canonical spray of a Finsler manifold $(M, E)$ and $P_H$ its associated Barthel endomorphism, the other connections of interest are characterised as follows:

$$\begin{array}{lll} \delta^V = 0, & \delta^H = \frac{1}{2}\mathcal{C}_H, & \text{(Chern–Rund)} \\ \delta^V = \frac{1}{2}\mathcal{C}_V, & \delta^H = 0, & \text{(Hashiguchi)} \\ \delta^V = \frac{1}{2}\mathcal{C}_V, & \delta^H = \frac{1}{2}\mathcal{C}_H. & \text{(Cartan)} \end{array} \tag{1.8}$$

The price to pay with these modifications of the Berwald connection is that every step towards the ideal of a metrical connection (the Cartan connection is fully metrical) introduces more torsion, i.e. a larger deviation from the ideal of a maximally torsion-free connection.

# Chapter 2

# *Berwald-type connections – Time-dependent case*

## 2.1 Time-dependent second-order differential equations

In the last decade some applications have been developed in the study of second-order ordinary differential equations (SODE's), see e.g. [55, 23, 75, 57, 39], which make use of covariant derivative operators; these may be seen to come essentially from the Berwald-type connection associated to the non-linear connection of the given SODE (see the next chapter for a couple of examples). This by itself may be a sufficient reason for having a closer look at the relationship between various versions of such a connection which have been discovered independently in the literature.

Our interest in this subject comes in the first place from the study of *time-dependent* second-order differential equations. The 'time-dependency' of our approach will be a consequence of the simple requirement that the base manifold $M$ is fibred over the real numbers, $\pi_{\scriptscriptstyle R} : M \to I\!R$. We will in particular be interested in time-dependent Lagrangian systems: the coordinate on $I\!R$ then represents time, and the manifold $M$ space-time. The space-time-velocity space is then the total manifold $J^1M$ of the first jet bundle of $\pi_{\scriptscriptstyle R}$, $\pi_{\scriptscriptstyle M} : J^1M \to M$.

**Definition 2.1.** *(see e.g. [20]) Two sections $\phi$ and $\psi$ of $\pi_{\scriptscriptstyle R}$ are said to be equivalent at $t$ if $\phi(t) = \psi(t)$ and $T_t\phi\left(\frac{d}{dt}\right) = T_t\psi\left(\frac{d}{dt}\right)$ ($\frac{d}{dt}$ is the basis vector in $T_tI\!R$). The equivalence class of $\phi$ under this equivalence relation is called the first jet of $\phi$ at $t$, denoted by $j^1_t\phi$. The collection of all 1-jets of sections of $\pi_{\scriptscriptstyle R}$ is a differentiable manifold denoted by $J^1M$. The map $\pi_{\scriptscriptstyle M} : J^1M \to M, j^1_t\phi \mapsto \phi(t)$ gives $J^1M$ the structure of an affine bundle. There exists also a second fibration, $J^1M \to I\!R$ given by $j^1_t\phi \mapsto t$.*

Coordinates $(t, x^i, v^i)$ of a first jet $j^1_t\phi$ are found as follows: $t$ is the coordinate that is given by the fibration $J^1M \to I\!R$; $(x^i)$ are $\pi_R$-fibre coordinates of $\pi_M(j^1_t\phi)$; $v^i = \frac{d\phi^i}{dt}|_t$. A coordinate change on $M$ (leaving the 'time' $t$ unchanged) induces the following coordinate change on $J^1M$:

$$\hat{t} = t, \qquad \hat{x}^i = \hat{x}^i(t,x) \qquad \text{and} \qquad \hat{v}^i = \frac{\partial \hat{x}^i}{\partial t}(t,x) + \frac{\partial \hat{x}^i}{\partial x^j}(t,x)v^j.$$

An introduction into the geometry of jet bundles can, for example, be found in [20, 47, 81]. Here we will only repeat those definitions that are indispensable for a clear understanding of this chapter.

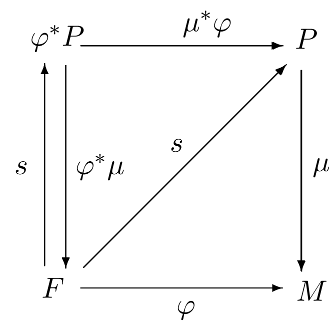
A time-dependent SODE is usually modelled as a vector field on $J^1M$. As in the framework of autonomous SODE's, we will show (in the next sections) how time-dependent ones define a non-linear connection on the bundle $\pi_M : J^1M \to M$. There are at least three known constructions in the literature of an associated linear connection. These were independently derived, from different perspectives, and do not make use of any other tool than the horizontal distribution coming from the given SODE. Therefore, they should somehow correspond to a generalised version of the concept of Berwald-type connection. Two of these linear connections were constructed on the tangent bundle $\tau_{J^1M} : T(J^1M) \to J^1M$, respectively by Massa and Pagani [57] and by Byrnes [9]. The third construction by Crampin *et al* [23] is a more direct one on the bundle $\pi_M^*\tau_M : \pi_M^*TM \to J^1M$ which generalises (1.4).

For completeness, we should mention that a construction of certain connections for a time-dependent framework, in particular a Cartan-type connection, can also be found in [67]. We shall, however, not go into the details of comparing our analysis with this work because the general setting is different. Indeed, 'space-time' in [67] is strictly separated: $M = I\!R \times Q$. As a consequence, the carrier space for the linear connections is $I\!R \times TQ$ (to which $J^1M$ is diffeomorphic, but not in a canonical way) and, unlike it was the case e.g. in [80], the constructions carried out in [67] have an intrinsic meaning only for a strict product bundle interpretation of $I\!R \times TQ$. In other words, it is as though one specific trivialisation of $J^1M$ is singled out and from then on coordinate transformations are not allowed to depend on time.

## 2.2 The pullback bundle $\pi_M^*\tau_M$

Before arriving at the three constructions, we will explain in detail the framework needed to study time-dependent SODEs. So far, we have encountered two different pullback bundles: the bundle $\tau_M^*\tau_M : \tau_M^*TM \to TM$ in the autonomous set-up and $\pi_M^*\tau_M : \pi_M^*TM \to J^1M$ for the time-dependent case. Many more pullback bundles will show up further on, so we shall spend a moment explaining the details of the structure of such bundles.

Let $\mu : P \to M$ be a fibre bundle and $\varphi : F \to M$ be a smooth map. The pullback of $\mu : P \to M$ by the map $\varphi$ is the bundle $\varphi^*\mu : \varphi^*P \to F$ where the total space $\varphi^*P$ is a submanifold of $F \times P$, consisting of those elements $(f,p)$, such that $\varphi(f) = \mu(p)$. The projection $\varphi^*\mu$ maps any such couple onto its first component $f$. The fibre over $f \in F$ is the set of points $p \in P$ such that $\mu(p) = \varphi(f)$, and therefore this fibre is identifiable with $P_{\varphi(f)}$. In particular, when also $\varphi$ is a bundle projection, we can look in the same way at the pullback of $\varphi : F \to M$ by $\mu$. Evidently, the total space $\mu^*F$ of this bundle will be equal to $\varphi^*P$ (and sometimes we will use also the notation $F \times_M P$ for this manifold), but the projection $\mu^*\varphi$ will map a couple $(f,p) \in \mu^*F$ now on the second element $p$.




A section of $\varphi^*\mu$ is a map $s : F \to \varphi^*P$ satisfying $\varphi^*\mu \circ s = id_F$. Such a map can equivalently be regarded as going from $F$ to $P$, in such a way that $\mu \circ s = \varphi$ (see the diagram). We therefore often call $s$ a section of $\mu$ along $\varphi$ (more specifically a vector field along $\varphi$, in case $P = TM$). Sections of $\mu$ itself can be regarded as sections along $\varphi$ by composition with $\varphi$. We refer to these as *basic sections*.

Let us now come back to the pullback bundle $\pi_M^*\tau_M$ of interest. The module of sections of this bundle will be denoted with $\mathcal{X}(\pi_M)$. It contains a canonically defined section, the *total time derivative operator* $\mathbf{T}$, which, regarded as a map $\mathbf{T} : J^1M \to TM$, can be defined as

$$\mathbf{T}(j^1_t\phi) = T_t\phi\left(\frac{d}{dt}\right).$$

Here $\phi$ is a section of $\pi_{\mathbb{R}}$, $j^1_t\phi$ its first jet at $t$ and $\frac{d}{dt}$ a basis vector in $T_t\mathbb{R}$. One can verify that this definition is independent of the choice of the representative $\phi$ of the equivalence class $j^1_t\phi$. The canonical vector field $\mathbf{T}$ has the following local coordinate representation

$$\mathbf{T} = \frac{\partial}{\partial t} + v^i\frac{\partial}{\partial x^i}.$$

*One-forms along* $\pi_M$ are sections of the pullback bundle $\pi_M^*\tau_M^*$, where $\tau_M^*$ denotes the cotangent bundle. We will use the notation $\bigwedge^1(\pi_M)$ for the set of one-forms along $\pi_M$. Then, $\mathbf{T}$ has a counterpart in $\bigwedge^1(\pi_M)$: the one-form $dt$ on $M$ is globally defined and can be interpreted as the basic one-form $dt \circ \pi_M$ along $\pi_M$. The natural pairing between vector fields and one-forms on $M$ easily extends to vector fields and one-forms along $\pi_M$. Sections that are annihilated by $dt$ are of the form $\overline{X} = \overline{X}^i(t,x,v)\frac{\partial}{\partial x^i}$. We will denote the set of all such sections by $\overline{\mathcal{X}}(\pi_M)$. If we write $VM$ for the kernel of the map $T\pi_{\mathbb{R}}$, then elements of $\overline{\mathcal{X}}(\pi_M)$ can be seen as sections of the bundle $\overline{\pi}_M : VM \to M$. $\overline{\mathcal{X}}(\pi_M)$ forms a direct sum complement with the span of $\mathbf{T}$ in $\mathcal{X}(\pi_M)$:

$$\mathcal{X}(\pi_M) \equiv \overline{\mathcal{X}}(\pi_M) \oplus \langle\mathbf{T}\rangle. \tag{2.1}$$

Throughout this chapter we shall write $X = \overline{X} + \langle X, dt\rangle\mathbf{T}$ for the decomposition of vector fields in $\mathcal{X}(\pi_M)$. In coordinates, if $X = X^0(t,x,v)\frac{\partial}{\partial t} + X^i(t,x,v)\frac{\partial}{\partial x^i}$, then

$$\overline{X} = (X^i - X^0v^i)\frac{\partial}{\partial x^i} \quad \text{and} \quad \langle X, dt\rangle = X^0.$$

## 2.3 The canonical short exact sequence and non-linear connections

Next to the pullback bundle $\pi_M^*\tau_M : \pi_M^*TM \to J^1M$, also the tangent bundle $\tau_{J^1M} : T(J^1M) \to J^1M$ of $J^1M$ will be important. The kernel of the tangent map of $\pi_M$ forms a submanifold in $TJ^1M$, *the vertical subbundle* $VJ^1M$. There exists a canonical short exact sequence

$$0 \to VJ^1M \to TJ^1M \to \pi_M^*TM \to 0, \tag{2.2}$$

where the second arrow is the natural injection of vertical vectors into $TJ^1M$ and the third is the map $j : TJ^1M \to \pi_M^* TM : \xi_w \mapsto (w, T\pi_M(\xi_w))$.

**Definition 2.2.** *A (non-linear) connection on $\pi_M$ is a (right) splitting ${}^H$ : $\pi_M^* TM \to TJ^1M$ of the sequence (2.2).*

Whenever we have a non-linear connection, the image of ${}^H$, *the horizontal subbundle* $HJ^1M$, will be a direct sum complement of the $VJ^1M$,

$$TJ^1M \equiv HJ^1M \oplus VJ^1M. \tag{2.3}$$

There exists an easy way to construct vertical vectors, starting from a pair in $\pi_M^* VM$: if $(w, \mathsf{v}) \in \pi_M^* VM$, then we can define an element $\mathsf{v}_w^V$ in $T_w J^1M$ by fixing its action on functions $f \in C^\infty(J^1M)$:

$$\mathsf{v}_w^V f = \frac{d}{dt}\big(f(w + t\mathsf{v})\big)_{t=0}.$$

The extension of this construction to vector fields along $\pi_M$, i.e. to a map ${}^V$ : $\overline{\mathcal{X}}(\pi_M) \to \mathcal{X}(J^1M)$, is called *the vertical lift.* Using the decomposition (2.1), it is easy to extend ${}^V$ to all vector fields along $\pi_M$: simply use $C^\infty(J^1M)$-linearity and put $\mathbf{T}^V = 0$. We will denote the set of vertical vector fields on $J^1M$ by $\overline{\mathcal{X}}(\pi_M)^V$ since the vertical lift is an isomorphism between $\overline{\mathcal{X}}(\pi_M)$ and the vertical vector fields.

All manifolds in the short exact sequence (2.2) are fibred over $J^1M$ and therefore there exists an equivalent short exact sequence at the level of their modules of sections

$$0 \to \overline{\mathcal{X}}(\pi_M)^V \to \mathcal{X}(J^1M) \to \mathcal{X}(\pi_M) \to 0. \tag{2.4}$$

Any splitting of (2.2) induces a splitting $\mathcal{X}(\pi_M) \to \mathcal{X}(J^1M)$ of (2.4), also denoted by ${}^H$. In the following, we will refer to ${}^H$ as the *horizontal lift.* An element in the image of ${}^H$ is a *horizontal vector field.* Analogously to (2.3) there will be a decomposition

$$\mathcal{X}(J^1M) \equiv \mathcal{X}(\pi_M)^H \oplus \overline{\mathcal{X}}(\pi_M)^V, \tag{2.5}$$

meaning that for a general $\xi \in \mathcal{X}(J^1M)$, there exists an element $\xi_H \in \mathcal{X}(\pi_M)$ and an element $\overline{\xi}_V \in \overline{\mathcal{X}}(\pi_M)$ such that

$$\xi = {\xi_H}^H + {\overline{\xi}_V}^V.$$

The horizontal distribution is locally spanned by vector fields

$$H_0 = \frac{\partial}{\partial t} - \Gamma_0^j(t,x,v)\frac{\partial}{\partial v^j}, \quad H_i = \frac{\partial}{\partial x^i} - \Gamma_i^j(t,x,v)\frac{\partial}{\partial v^j}. \tag{2.6}$$

The functions $\Gamma_0^j$ and $\Gamma_i^j$ are called the *connection coefficients*. We will also use the shorthand notations $V_i$ for the vector fields $\frac{\partial}{\partial v^i}$ which span the vertical distribution. For $\xi = \xi^0(t,x,v)\frac{\partial}{\partial t} + \xi^i(t,x,v)\frac{\partial}{\partial x^i} + \eta^i(t,x,v)\frac{\partial}{\partial v^i}$,

$$\xi_H = \xi^0\frac{\partial}{\partial t} + \xi^i\frac{\partial}{\partial x^i} \quad \text{and} \quad \overline{\xi}_V = \left(\xi^0\Gamma_0^i + \xi^k\Gamma_k^i + \eta^i\right)\frac{\partial}{\partial x^i}. \tag{2.7}$$

As in the autonomous case, we will use $P_H$ and $P_V$ for respectively the horizontal and vertical projection operators.

*The vertical endomorphism* $S$ is a (1,1) tensor field on $J^1M$. Its definition is easy when we assume that a connection is given:

$$S(\overline{X}^V) = 0 \quad \text{and} \quad S(Y^H) = \overline{Y}^V, \qquad \overline{X} \in \overline{\mathcal{X}}(\pi_M), Y \in \mathcal{X}(\pi_M)$$

with $\overline{Y} = Y - \langle Y, dt\rangle\mathbf{T}$. One can check that this definition is independent of the choice of the connection. In coordinates, $S$ is given by

$$S = \theta^i \otimes \frac{\partial}{\partial v^i}, \qquad \theta^i = dx^i - v^i dt. \tag{2.8}$$

In contrast with the vertical endomorphism, the (degenerate) *almost complex structure* $J$ *on* $J^1M$ is determined by the horizontal distribution according to the following defining relations:

$$J(\overline{X}^H) = \overline{X}^V, \qquad J(\overline{X}^V) = -\overline{X}^H, \qquad J(\mathbf{T}^H) = 0, \tag{2.9}$$

Finally, let $M$ be the degenerate *almost product structure* determined by

$$M(\overline{X}^H) = \overline{X}^V, \qquad M(\overline{X}^V) = \overline{X}^H, \qquad M(\mathbf{T}^H) = 0. \tag{2.10}$$

The role that $\mathbf{T}$ plays in $\mathcal{X}(\pi_M)$ has an analogue in $\mathcal{X}(J^1M)$, not in the form of one single vector field, but in the appearance of a whole subclass of vector fields on $J^1M$, the so-called SODEs .

**Definition 2.3.** *A* SODE $\Gamma$ *is a vector field on* $J^1M$ *such that* $S(\Gamma) = 0$ *and* $\langle \Gamma, dt\rangle = 1$.

The abbreviation SODE stands for Second-order Ordinary Differential Equation field and is justified by the fact that a SODE $\Gamma$ locally takes the form

$$\Gamma = \frac{\partial}{\partial t} + v^i \frac{\partial}{\partial x^i} + f^i(t,x,v)\frac{\partial}{\partial v^i}, \tag{2.11}$$

from which it is clear that its integral curves are solutions of the following set of time-dependent second-order ordinary differential equations

$$\ddot{x}^i = f^i(t,x,\dot{x}).$$

When a non-linear connection on $\pi_M$ is given, $\mathbf{T}^H$ is a SODE, often called the *associated semi-spray* of the given connection,

$$\mathbf{T}^H = \frac{\partial}{\partial t} + v^i \frac{\partial}{\partial x^i} - \left(\Gamma^j_0 + v^i \Gamma^j_i\right)\frac{\partial}{\partial v^j}. \tag{2.12}$$

The decomposition (2.1) of the vector fields along the projection then further decomposes expression (2.5) into:

$$\mathcal{X}(J^1M) \equiv \overline{\mathcal{X}}(\pi_M)^H \oplus \overline{\mathcal{X}}(\pi_M)^V \oplus \langle \mathbf{T}^H \rangle. \tag{2.13}$$

Needless to say, horizontal vector fields on $J^1M$ may have components both in the first and third set of the decomposition (2.13): for a general $\xi \in \mathcal{X}(J^1M)$, we may write now

$$\xi \;=\; \overline{\xi}_H{}^H + \overline{\xi}_V{}^V + \langle \xi, dt\rangle \mathbf{T}^H, \tag{2.14}$$

with both $\overline{\xi}_H, \overline{\xi}_V \in \overline{\mathcal{X}}(\pi_M)$. Here

$$\overline{\xi}_H = (\xi^i - v^i\xi^0)\frac{\partial}{\partial x^i} \quad \text{and} \quad \langle \xi, dt\rangle = \xi^0$$

while $\overline{\xi}_V$ remains the same as in (2.7). An important remark in this respect is the following: the generalisation from an autonomous framework to a time-dependent one in a way has two faces; some formulas tend to carry over in a natural way by thinking of the first decomposition in (2.14), as though one would formally copy the decomposition (1.1) with one extra dimension in the horizontal component; other features, however, tend to be better understood if one thinks of $\overline{\mathcal{X}}(\pi_M)$ as the analogue of $\mathcal{X}(\tau_M)$ and thus assigns a separate role to the one-dimensional distribution spanned by $\mathbf{T}^H$. Most of the technicalities in what follows (if not all) are related to this dichotomy.

We are in particular interested in the case where the connection is the one canonically associated to a given SODE $\Gamma$:

**Proposition 2.4.** *[26] For any* SODE $\Gamma$*, the projector*

$$P_H = \frac{1}{2}\Big(I - \mathcal{L}_\Gamma S + dt \otimes \Gamma\Big). \tag{2.15}$$

*defines a non-linear connection on* $\pi_M$*. We will often refer to this construction as the '*SODE *connection generated by* $\Gamma$*'.*

The corresponding connection coefficients in (2.6) are given by

$$\Gamma^j_i = -\tfrac{1}{2}\frac{\partial f^j}{\partial v^i}, \quad \Gamma^j_0 = -f^j - v^k\Gamma^j_k. \tag{2.16}$$

Note that an advantage of the time-dependent set-up is that $\mathbf{T}^H$ then coincides with the given $\Gamma$, a feature which is not in general true for the autonomous framework. Beware, however, that if one starts from a general horizontal distribution and looks at the SODE $\Gamma_0 = \mathbf{T}^H$, the original connection need not coincide with the SODE connection of $\Gamma_0$.

**Proposition 2.5.** *Non-linear connections which come from a* SODE *are characterised by the property that their torsion* $[P_H, S]$ *is zero.*

The computation of the Nijenhuis bracket $[P_H, S]$ is easy to carry out in a basis of local vector fields adapted to the decomposition (2.13). In fact, one finds that only two types of components are not trivially zero. We list them here for later use:

$$[P_H, S](\overline{X}^H, \overline{Y}^H) = [\overline{X}^H, \overline{Y}^V]_V{}^V - [\overline{Y}^H, \overline{X}^V]_V{}^V - [\overline{X}^H, \overline{Y}^H]_H{}^V, \tag{2.17}$$

$$[P_H, S](\mathbf{T}^H, \overline{X}^H) = [\mathbf{T}^H, \overline{X}^V]_V{}^V - [\mathbf{T}^H, \overline{X}^H]_H{}^V. \tag{2.18}$$

To end this section, we mention that the Nijenhuis bracket $\frac{1}{2}[P_H, P_H]$ is called the *curvature of the non-linear connection* . Here, the non-trivial terms are

$$\tfrac{1}{2}[P_H, P_H](\overline{X}^H, \overline{Y}^H) = [\overline{X}^H, \overline{Y}^H]_V{}^V, \tag{2.19}$$

$$\tfrac{1}{2}[P_H, P_H](\mathbf{T}^H, \overline{X}^H) = [\mathbf{T}^H, \overline{X}^H]_V{}^V. \tag{2.20}$$

## 2.4 Time-dependent Lagrangian mechanics

We next say a few words about time-dependent Lagrangian mechanics. A Lagrangian is a function $L : J^1M \to \mathbb{R}$. To any Lagrangian $L$ we can associate two forms: the *Poincaré-Cartan 1-form* $\theta_L = S(dL) + Ldt$ and its exterior derivative $\omega_L = d\theta_L$, the *Poincaré-Cartan 2-form.* A Lagrangian is said to be *regular* if the matrix

$$(g_{ij}) = (\frac{\partial^2 L}{\partial v^i \partial v^j}) \tag{2.21}$$

is everywhere non-singular.

**Proposition 2.6.** *[26] For any regular Lagrangian $L$, there exists a unique vector field $\Gamma$ on $J^1M$ satisfying*

$$i_\Gamma \omega_L = 0 \qquad \textit{and} \qquad \langle \Gamma, dt \rangle = 1.$$

*$\Gamma$ is a* SODE*, called the Euler-Lagrange vector field, and its integral curves are the solutions of the Euler-Lagrange equations*

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\right) - \frac{\partial L}{\partial x^i} = 0.$$

Next to this SODE, we can also characterise a symmetric (0,2) tensor field along $\pi_M$ by means of the Poincaré-Cartan 2-form.

$$g_L(X, Y) = \omega_L(X^V, Y^H). \tag{2.22}$$

This definition is independent of the choice of the connection. Indeed, since the difference between two horizontal lifts is a vertical lift and since $\omega_L(X^V, Y^V) = 0$, $g_L$ is uniquely defined. It is symmetric because for e.g. the SODE connection of the Euler-Lagrange vector field $\Gamma$, $\omega_L(\overline{X}^V, \overline{Y}^H) = -\omega_L(\overline{X}^H, \overline{Y}^V)$ and $\omega_L(\Gamma, \xi) = 0$. In particular, $g_L(\mathbf{T}, Y) = 0 = g_L(X, \mathbf{T})$. If the Lagrangian is regular, then the restriction of $g$ to vector fields in $\overline{\mathcal{X}}(\pi_M)$ is non-singular. In coordinates, $g_L = g_{ij}(t, x, v)\theta^i \otimes \theta^j$ with $g_{ij}$ as in (2.21).

The components $f^i(t, x, v)$ of the Euler-Lagrange vector field $\Gamma$ then become

$$f^i = g^{ij}\left(\frac{\partial L}{\partial x^j} - v^k \frac{\partial^2 L}{\partial x^k \partial v^j} - \frac{\partial^2 L}{\partial t \partial v^j}\right). \tag{2.23}$$

## 2.5 A general scheme for lifting linear connections from $\pi_M^* \tau_M$ to $\tau_{J^1M}$ and vice versa

Our primary objective is to explore a general scheme within which the three constructions of [57], [23] and [9] can be compared and related to each other:

as expressed in the first section, they should in some sense be equivalent and represent the 'time-dependent generalisation' of the connection of Berwald type associated to the given horizontal distribution on $\pi_M$. Their difference no doubt will come from a certain freedom or indeterminacy in 'fixing the time-component' of the connection.

There are no points in our analysis which could exclusively be dealt with in the case of a SODE connection of the form (2.15), so we will generally not specify the horizontal distribution even though that will require sometimes reformulating previous work of other authors in such a more general context. In the following we will be mainly interested in linear connections on $\pi_M^*\tau_M : \pi_M^*TM \to J^1M$ (which we will denote with D) and linear connections on $\tau_{J^1M} : T(J^1M) \to J^1M$ (for which we reserve the notation $\nabla$).

Let D be a linear connection on $\pi_M^*\tau_M$ which induces a linear connection on the subbundle $\pi_M^*\overline{\pi}_M$, i.e. which satisfies the assumption

$$\mathrm{D}_\xi(\overline{\mathcal{X}}(\pi_M)) \subset \overline{\mathcal{X}}(\pi_M) \quad \forall \xi \in \mathcal{X}(J^1M). \tag{2.24}$$

This is the only restriction which is required if we want to think of situations coming from an analogue in the tangent bundle set-up. Using a given horizontal distribution on $J^1M$ and having the autonomous doubling model (1.5) in mind, we define an associated class of linear connections on $\tau_{J^1M}$ by putting

$$\nabla_\xi \overline{X}^H = (\mathrm{D}_\xi \overline{X})^H, \qquad \nabla_\xi \overline{X}^V = (\mathrm{D}_\xi \overline{X})^V, \qquad \nabla_\xi \mathbf{T}^H = K(\xi), \tag{2.25}$$

where $K$ is any type (1,1) tensor field on $J^1M$. It is easy to verify that for $\eta \in \mathcal{X}(J^1M)$, the operation

$$\nabla_\xi \eta = (\mathrm{D}_\xi \overline{\eta}_H)^H + (\mathrm{D}_\xi \overline{\eta}_V)^V + \xi(\langle \eta, dt \rangle)\mathbf{T}^H + \langle \eta, dt \rangle K(\xi) \tag{2.26}$$

defines a linear connection indeed, for any choice of $K$. All elements of such a class have the following easy to establish properties. Firstly,

$$\nabla_\xi(\overline{\mathcal{X}}(\pi_M)^H) \subset \overline{\mathcal{X}}(\pi_M)^H, \quad \nabla_\xi(\overline{\mathcal{X}}(\pi_M)^V) \subset \overline{\mathcal{X}}(\pi_M)^V \;\; \forall \xi \in \mathcal{X}(J^1M). \tag{2.27}$$

Secondly, if $J$ is the almost complex structure on $J^1M$, determined by (2.9), then we have

$$J(\nabla_\xi \overline{X}^H) = \nabla_\xi \overline{X}^V, \qquad J(\nabla_\xi \overline{X}^V) = -\nabla_\xi \overline{X}^H, \tag{2.28}$$

or equivalently

$$\nabla_\xi J|_{\overline{\mathcal{X}}(J^1M)} = 0,$$

where $\overline{\mathcal{X}}(J^1M) \equiv \overline{\mathcal{X}}(\pi_M)^H \oplus \overline{\mathcal{X}}(\pi_M)^V$.

Conversely, let $\nabla$ be any linear connection on $\tau_{J^1M}$ having the properties (2.27,2.28), then we define an associated class of linear connections on $\pi_M^*\tau_M$ by putting:

$$\mathrm{D}_\xi \overline{X} = (\nabla_\xi \overline{X}^H)_H = (\nabla_\xi \overline{X}^V)_V, \qquad \mathrm{D}_\xi \mathbf{T} = L(\xi), \tag{2.29}$$

where $L$ is any $C^\infty(J^1M)$-linear map from $\mathcal{X}(J^1M)$ to $\mathcal{X}(\pi_M)$. Indeed, for any $X \in \mathcal{X}(\pi_M)$, the relation

$$\mathrm{D}_\xi X = (\nabla_\xi \overline{X}^H)_H + \xi(\langle X, dt\rangle)\mathbf{T} + \langle X, dt\rangle L(\xi) \tag{2.30}$$

is a linear connection on $\pi_M^*\tau_M$ for any tensorial $L$. Obviously, each element of the class will have the property (2.24). If we take such an element and raise it to the bigger space of linear connections on $\tau_{J^1M}$ again according to the first procedure, we will obtain for every choice of $K$ an element of the same class as the $\nabla$ we started from. The type of linear connections on $\tau_{J^1M}$ we encounter in this construction are the ones we wish to call *connections of Finsler type*.

**Definition 2.7.** *A pair $(P_H, \nabla)$ consisting of a horizontal distribution on $\pi_M$ (represented by its horizontal projector) and a linear connection on $\tau_{J^1M}$ is said to be of Finsler type if we have the properties (2.27,2.28).*

Essentially, connections of Finsler type come from a class of linear connections D on $\pi_M^*\tau_M$ with property (2.24) and we will sometimes figuratively term the couple $(P_H, \mathrm{D})$ as being of Finsler type as well.

In order to obtain some equivalent characterisations of Finsler-type connections, we first prove two simple lemmas. A preliminary notational convention is in order here: various types of identity tensors (operating on different vector fields) will play a role in the sequel. The identity operator for $\mathcal{X}(\pi_M)$ will be denoted by $I$. We write its natural decomposition as

$$I = \overline{I} + dt \otimes \mathbf{T}, \qquad \text{with} \quad \overline{I} = \theta^i \otimes \frac{\partial}{\partial x^i}. \tag{2.31}$$

Likewise, $I_{J^1M}$ is the identity on $\mathcal{X}(J^1M)$ and $\overline{I}_{J^1M}$ is that part of $I_{J^1M}$ which vanishes on $\mathbf{T}^H$.

**Lemma 2.8.** $S \circ J + J \circ S = -\overline{I}_{J^1M}$.

PROOF: From $S(X^H) = X^V$ and the defining relations of $J$, it follows that

$$S(J(X^V)) = -X^V = -\overline{I}_{J^1M}(X^V), \quad S(J(X^H)) = S(J(\mathbf{T}^H)) = 0,$$

and

$$J(S(X^V)) = J(S(\mathbf{T}^H)) = 0, \quad J(S(X^H)) = -X^H = -\overline{I}_{J^1M}(X^H).$$

The result then readily follows. □

Let now $P_{\overline{H}}$ be the 'strong horizontal projector' defined by $P_{\overline{H}}(\overline{X}^H) = \overline{X}^H$, $P_{\overline{H}}(\overline{X}^V) = 0$, $P_{\overline{H}}(\mathbf{T}^H) = 0$, and let $M$ be the degenerate almost product structure as in (2.10).

**Lemma 2.9.** $J \circ P_{\overline{H}} - P_{\overline{H}} \circ J = M$.

PROOF: The simple proof is similar to the one of Lemma 2.8. □

Note that one can also obtain the relation $J \circ P_{\overline{H}} + P_{\overline{H}} \circ J = J$.

**Proposition 2.10.** *The following are equivalent characterisations of connections of Finsler type:*

$$\text{(2.27) and (2.28)} \iff \nabla_\xi P_{\overline{H}}|_{\overline{\mathcal{X}}(J^1M)} = 0 \;\text{ and }\; \nabla_\xi J|_{\overline{\mathcal{X}}(J^1M)} = 0. \quad (2.32)$$

$$\left\{ \begin{array}{l} \nabla_\xi P_{\overline{H}}|_{\overline{\mathcal{X}}(J^1M)} = 0 \\ \nabla_\xi J|_{\overline{\mathcal{X}}(J^1M)} = 0 \end{array} \right. \iff \left\{ \begin{array}{l} \nabla_\xi P_{\overline{H}}|_{\overline{\mathcal{X}}(J^1M)} = 0 \\ \nabla_\xi S|_{\overline{\mathcal{X}}(J^1M)} = 0 \end{array} \right. \quad (2.33)$$

PROOF: Making use of the information in (2.27), one easily finds from taking a covariant derivative of the defining relations of $P_{\overline{H}}$ that (2.27) implies $\nabla_\xi P_{\overline{H}}|_{\overline{\mathcal{X}}(J^1M)} = 0$. Conversely, this invariance implies that $P_{\overline{H}}(\nabla_\xi \overline{X}^H) = \nabla_\xi \overline{X}^H$ and $P_{\overline{H}}(\nabla_\xi \overline{X}^V) = 0$. The first of these says that $\nabla_\xi \overline{X}^H \in \overline{\mathcal{X}}(\pi_M)^H$, whereas the second only ensures that $\nabla_\xi \overline{X}^V \in \overline{\mathcal{X}}(\pi_M)^V \oplus \langle \mathbf{T}^H \rangle$ in a direct way. Indirectly however, using also $\nabla_\xi \overline{X}^V = J(\nabla_\xi \overline{X}^H)$ and Lemma 2.9, we find that $0 = J(\nabla_\xi \overline{X}^H) - M(\nabla_\xi \overline{X}^H) = \nabla_\xi \overline{X}^V - M(\nabla_\xi \overline{X}^H)$, which ensures that $\nabla_\xi \overline{X}^V$ belongs to $\overline{\mathcal{X}}(\pi_M)^V$ anyway.

Secondly, from $S(\overline{X}^V) = 0$ and (2.27), it follows that $\nabla_\xi S(\overline{X}^V) = 0$. From $S(\overline{X}^H) = \overline{X}^V$ and the second relation in (2.28), it follows that $\nabla_\xi S(\overline{X}^H) - S(J(\nabla_\xi \overline{X}^V)) = \nabla_\xi \overline{X}^V$. Using Lemma 2.8 and the information that $\nabla_\xi \overline{X}^V$ is vertical (from (2.27) again), it also follows that $\nabla_\xi S(\overline{X}^H) = 0$. This means that (2.27) and (2.28) imply $\nabla_\xi S|_{\overline{\mathcal{X}}(J^1M)} = 0$. For the converse, note first that $\nabla_\xi P_{\overline{H}}|_{\overline{\mathcal{X}}(J^1M)} = 0$ implies that $\nabla_\xi \overline{X}^H \in \overline{\mathcal{X}}(\pi_M)^H$. Next, from $\nabla_\xi S|_{\overline{\mathcal{X}}(J^1M)} = 0$ we find with the help of Lemma 2.8 again that

$$\begin{aligned} J(\nabla_\xi \overline{X}^V) &= J(\nabla_\xi(S(\overline{X}^H))) = J(S(\nabla_\xi \overline{X}^H)) \\ &= (-\overline{I}_{J^1M} - S \circ J)(\nabla_\xi \overline{X}^H) = -\nabla_\xi \overline{X}^H. \end{aligned}$$

Applying $J$ to this relation, we obtain that also $J(\nabla_\xi \overline{X}^H) = \nabla_\xi \overline{X}^V$, so that $\nabla_\xi J|_{\overline{\mathcal{X}}(J^1M)} = 0$ indeed. This completes the proof. □

## 2.6 The class of Berwald-type connections

The motivation for introducing the equivalence classes of linear connections of the previous section is that we want to frame the three existing constructions in the literature of a linear connection associated to a given SODE within one common scheme: that of a class of Berwald-type connections. The philosophy here is that one has to understand first all aspects lying at the origin of the difference between these constructions, before one can decide upon an optimal selection. Now, the Berwald-type connection for the autonomous framework, at least in its appearance on the pullback bundle $\tau_M^*\tau_M$ in [21], is defined by (1.4). Within the present temporary scheme of equivalence classes of connections, we thus arrive at the following definition of the *class of Berwald-type connections*.

**Definition 2.11.** *A linear connection* $\mathrm{D}$ *on* $\pi_M^*\tau_M$ *with the property (2.24) belongs to the class of Berwald-type connections with respect to a given horizontal distribution, if it satisfies*

$$\mathrm{D}_\xi \overline{X} = [P_H(\xi), \overline{X}^V]_V + [P_V(\xi), \overline{X}^H]_H, \tag{2.34}$$

*for all* $\xi \in \mathcal{X}(J^1M)$ *and* $\overline{X} \in \overline{\mathcal{X}}(\pi_M)$. *A Finsler pair* $(P_H, \nabla)$ *on* $J^1M$ *is said to be of Berwald type if it is derived via (2.25) from a connection on* $\pi_M^*\tau_M$ *with the property (2.34).*

It is of some interest to look at the effect of the various assumptions so far discussed on the *torsion*

$$T(\xi, \eta) = \nabla_\xi \eta - \nabla_\eta \xi - [\xi, \eta]$$

of a pair $(P_H, \nabla)$ on $\tau_{J^1M}$. With the aid of the decomposition in horizontal and vertical parts, all components of $T$ can be traced back to tensor fields acting on $\mathcal{X}(\pi_M)$. We introduce notations similar to those in (1.3) for these tensor fields and list them in the first table below. Remark that, unlike the tensors in (1.3), the torsions in the table are tensors along $\pi_M$. The effect of assuming we have a connection of Finsler-type is that (2.29) can be invoked to express some covariant derivatives in terms of a D on $\pi_M^*\tau_M$ (see the 'Finsler' column in the second table ). If in addition we have a connection of Berwald type, further simplifications occur through the definition (2.34). For completeness: the component $T(\overline{X}^V, \overline{Y}^V)_H$ of the torsion becomes trivially zero as soon as the assumption (2.27) is satisfied and is therefore not listed.

| Definition |
|---|
| $\mathcal{A}(\overline{X}, \overline{Y}) = T(\overline{X}^H, \overline{Y}^H)_H$ |
| $\mathcal{R}(\overline{X}, \overline{Y}) = T(\overline{X}^H, \overline{Y}^H)_V$ |
| $\mathcal{B}(\overline{X}, \overline{Y}) = T(\overline{X}^H, \overline{Y}^V)_H$ |
| $\mathcal{P}(\overline{X}, \overline{Y}) = T(\overline{X}^H, \overline{Y}^V)_V$ |
| $\mathcal{S}(\overline{X}, \overline{Y}) = T(\overline{X}^V, \overline{Y}^V)_V$ |
| $\mathcal{A}_{\mathbf{T}}(\overline{X}) = T(\mathbf{T}^H, \overline{X}^H)_H$ |
| $\mathcal{R}_{\mathbf{T}}(\overline{X}) = T(\mathbf{T}^H, \overline{X}^H)_V$ |
| $\mathcal{B}_{\mathbf{T}}(\overline{X}) = T(\mathbf{T}^H, \overline{X}^V)_H$ |
| $\mathcal{P}_{\mathbf{T}}(\overline{X}) = T(\mathbf{T}^H, \overline{X}^V)_V$ |

| | Finsler | Berwald |
|---|---|---|
| $\mathcal{A}$ | $\mathrm{D}_{\overline{X}^H}\overline{Y} - \mathrm{D}_{\overline{Y}^H}\overline{X} - [\overline{X}^H, \overline{Y}^H]_H$ | $[\overline{X}^H, \overline{Y}^V]_V - [\overline{Y}^H, \overline{X}^V]_V - [\overline{X}^H, \overline{Y}^H]_H$ |
| $\mathcal{R}$ | $-[\overline{X}^H, \overline{Y}^H]_V$ | $-[\overline{X}^H, \overline{Y}^H]_V$ |
| $\mathcal{B}$ | $-\mathrm{D}_{\overline{Y}^V}\overline{X} - [\overline{X}^H, \overline{Y}^V]_H$ | 0 |
| $\mathcal{P}$ | $\mathrm{D}_{\overline{X}^H}\overline{Y} - [\overline{X}^H, \overline{Y}^V]_V$ | 0 |
| $\mathcal{S}$ | $\mathrm{D}_{\overline{X}^V}\overline{Y} - \mathrm{D}_{\overline{Y}^V}\overline{X} - [\overline{X}^V, \overline{Y}^V]_V$ | 0 |
| $\mathcal{A}_{\mathbf{T}}$ | $\mathrm{D}_{\mathbf{T}^H}\overline{X} - (\nabla_{\overline{X}^H}\mathbf{T}^H)_H - [\mathbf{T}^H, \overline{X}^H]_H$ | $[\mathbf{T}^H, \overline{X}^V]_V - (\nabla_{\overline{X}^H}\mathbf{T}^H)_H - [\mathbf{T}^H, \overline{X}^H]_H$ |
| $\mathcal{R}_{\mathbf{T}}$ | $-(\nabla_{\overline{X}^H}\mathbf{T}^H)_V - [\mathbf{T}^H, \overline{X}^H]_V$ | $-(\nabla_{\overline{X}^H}\mathbf{T}^H)_V - [\mathbf{T}^H, \overline{X}^H]_V$ |
| $\mathcal{B}_{\mathbf{T}}$ | $-(\nabla_{\overline{X}^V}\mathbf{T}^H)_H - [\mathbf{T}^H, \overline{X}^V]_H$ | $-(\nabla_{\overline{X}^V}\mathbf{T}^H)_H - [\mathbf{T}^H, \overline{X}^V]_H$ |
| $\mathcal{P}_{\mathbf{T}}$ | $\mathrm{D}_{\mathbf{T}^H}\overline{X} - (\nabla_{\overline{X}^V}\mathbf{T}^H)_V - [\mathbf{T}^H, \overline{X}^V]_V$ | $-(\nabla_{\overline{X}^V}\mathbf{T}^H)_V$ |

Perhaps one of the lines in the table requires an extra word. The $\mathcal{S}$-tensor, in the case of a Berwald-type connection, gives rise to the expression $\mathcal{S}(\overline{X},\overline{Y}) = [\overline{X}^V, \overline{Y}^H]_H - [\overline{Y}^V, \overline{X}^H]_H - [\overline{X}^V, \overline{Y}^V]_V$. This is manifestly zero when the arguments are basic vector fields, because a bracket such as $[\overline{X}^V, \overline{Y}^H]$ then is vertical, and therefore $\mathcal{S}$ is zero for all arguments. Note further that if the Berwald-type connection is given as a D on $\pi_M^*\tau_M$, the covariant derivatives of $\mathbf{T}^H$ in the right column of the last table are determined by our choice of the tensor $K$ in (2.25).

A case of particular interest is the one where the given horizontal distribution comes from a SODE connection (by means of Proposition 2.4). Indeed, in such a case we see from (2.17,2.18) that there are two further simplifications in the torsion for a connection of Berwald type: $\mathcal{A}(\overline{X},\overline{Y}) = 0$ and $\mathcal{A}_{\mathbf{T}}(\overline{X})$ reduces to $-(\nabla_{\overline{X}^H}\mathbf{T}^H)_H$.

## 2.7 Three Berwald-type connections

Let us now discuss the three available constructions referred to above for a linear connection associated to a SODE and verify whether they are of Berwald type indeed. As a preliminary remark, we should say that all of them were originally constructed with respect to the horizontal distribution associated to a given SODE $\Gamma$, but $\Gamma$ sometimes only enters the picture by the fact that it is $\mathbf{T}^H$. We will try to make our presentation somewhat more general by adapting the original construction to allow for any non-linear connection (or horizontal distribution) on $\pi_M$ as the starting point, although that will not be equally successful in all three cases.

The simplest construction to explain is the one by Crampin *et al* [23]. Essentially, it takes the direct construction formula (1.4), as first introduced in [52] for autonomous SODE's, as model and tries to carry it over to the time-dependent framework to construct a linear connection on $\pi_M^*\tau_M$. One then immediately observes that a correction term is needed for having $\mathrm{D}_\xi$ satisfying the derivation property. The defining relation of a linear connection, valid with respect to any given horizontal distribution, thus becomes

$$\mathrm{D}_\xi X = [P_H(\xi), X^V]_V + [P_V(\xi), X^H]_H + P_H(\xi)(\langle X, dt\rangle)\mathbf{T}. \tag{2.35}$$

Obviously, the requirement (2.34) is satisfied, so we are in the class of Berwald connections. It further follows that $\mathrm{D}_\xi \mathbf{T} = \overline{\xi}_V$. To say something about torsion in this case, we need to make a choice for the tensor $K$ in (2.25). It looks natural here to maintain the spirit in which the defining relation (2.35) was conceived by simply taking over the formula which raises the linear connection D to one on $\tau_{J^1M}$ from the autonomous framework. That is to say, we put (as in [23])

$$\nabla_\xi \eta = (\mathrm{D}_\xi \eta_H)^H + (\mathrm{D}_\xi \overline{\eta}_V)^V. \tag{2.36}$$

It is obvious then that the first two relations in (2.25) are satisfied and that the tensor $K$ is defined by

$$K(\xi) = \nabla_\xi \mathbf{T}^H = (\mathrm{D}_\xi \mathbf{T})^H = \overline{\xi}_V{}^H. \tag{2.37}$$

As a result, we have $\mathcal{P}_\mathbf{T} = 0$ and also $\mathcal{B}_\mathbf{T} = 0$ (since $\left(\nabla_{\overline{X}^V}\mathbf{T}^H\right)_H = \overline{X} = -[\mathbf{T}^H, \overline{X}^V]_H$), while $\mathcal{A}_\mathbf{T}$ and $\mathcal{R}_\mathbf{T}$ reduce to $\mathcal{A}_\mathbf{T} = [\mathbf{T}^H, \overline{X}^V]_V - [\mathbf{T}^H, \overline{X}^H]_H$, $\mathcal{R}_\mathbf{T} = -[\mathbf{T}^H, \overline{X}^H]_V$. If, in addition, the non-linear connection comes from a SODE $\Gamma$, all torsion tensors which can vanish become zero, except for $\mathcal{R}$ and $\mathcal{R}_\mathbf{T}$ (which are rather related to the curvature (2.19,2.20) of the non-linear connection).

Massa and Pagani [57] have constructed a linear connection on $\tau_{J^1M}$. Their way of building up the theory is somewhat harder to fit within our present approach, because a full horizontal distribution only becomes part of the data at the final stage of the argumentation, where a given SODE $\Gamma$ is singled out. Briefly, their construction starts as follows. First of all, among all possibly existing linear connections on $\tau_{J^1M}$, Massa and Pagani consider only those which preserve the 1-form $dt$, the canonical vertical endomorphism $S$ and (constant) parallel transport along the fibres. Explicitly, the as yet undetermined covariant derivative will have the properties: $\nabla_\xi dt = 0$, $\nabla_\xi S = 0$ and $\nabla_{\overline{X}^V}\overline{Y}^V = 0$, for every $\overline{X} \in \overline{\mathcal{X}}(\pi_M)$ and every *basic* $\overline{Y} \in \overline{\mathcal{X}}(\pi_M)$. With these assumptions, it is possible to construct two projection operators, which are assumed to be completely complementary at the subsequent stage, and are given by

$$P_{\overline{H}}(\eta) := T(\Gamma, S(\eta)), \qquad Q(\eta) := S(T(\Gamma,\eta)) + \langle \eta, dt\rangle \Gamma, \tag{2.38}$$

where $T$ is the torsion tensor of the linear connection to be constructed and $\Gamma$ is an arbitrary SODE. It is shown that $P_{\overline{H}}(\eta)$ and $Q(\eta)$ do not depend on the choice of $\Gamma$. Note further that the image of $Q$ contains all vertical vectors and all possible SODE's (which makes sense because the difference between two SODE's is vertical). After adopting some further restrictions (to which we come back later), a theorem is proved concerning existence and uniqueness of a linear connection which leaves a pre-selected SODE invariant. That SODE in fact, when added to the image of $P_{\overline{H}}$ completes the horizontal distribution to which the constructed linear connection can be thought of as being associated.

We explain now how this scheme can be slightly modified when an arbitrary horizontal distribution is given from the outset. Then, in particular, we have the SODE $\mathbf{T}^H$ at our disposal, which we can use to define the operators $P_{\overline{H}}$ and $Q$. In other words, we put

$$P_{\overline{H}}(\eta) := T(\mathbf{T}^H, S(\eta)), \qquad Q(\eta) := S(T(\mathbf{T}^H,\eta)) + \langle \eta, dt\rangle \mathbf{T}^H. \tag{2.39}$$

With $P_{\overline{H}} + Q = I_{J^1M}$ as part of the assumptions, we then have $P_H(\eta) = P_{\overline{H}}(\eta) + \langle \eta, dt\rangle \mathbf{T}^H$, and $P_V(\eta) = Q(\eta) - \langle \eta, dt\rangle \mathbf{T}^H$. The somewhat delicate point hereby is that, since the horizontal distribution is given, the defining relation for $P_{\overline{H}}$ here has to be regarded as an implicit restriction, via the torsion, on the class of admissible linear connections we want to consider.

Continuing now, in this modified picture, the line of reasoning of Massa and Pagani, assume that the class of potential $\nabla$'s is further restricted by requiring that they satisfy $\nabla_\xi P_{\overline{H}} = 0$ and have a curvature tensor *curv* which vanishes on any pair of SODE's, or equivalently satisfies $curv(\Gamma, \overline{X}^V) = 0$ for each SODE $\Gamma$ and all $\overline{X}$. One can prove as a minor modification of Theorem 2.2 in [57] that, with all hypotheses so far imposed, any admissible linear connection $\nabla$ is now completely determined if we know $\nabla_\Gamma \mathbf{T}^H$ for an arbitrary SODE $\Gamma$. The final point is to agree to fix this remaining freedom by requiring that $\nabla_{\mathbf{T}^H}\mathbf{T}^H = 0$ and $\nabla_{\partial/\partial v^i}\mathbf{T}^H = 0$ from which it follows as in [57] that actually $\nabla_\xi \mathbf{T}^H = 0$, $\forall\xi$. We thus have arrived (in a perhaps rather roundabout way) at a prescription for a uniquely defined linear connection on $\tau_{J^1M}$, corresponding to any pre-assigned horizontal distribution.

A point to be observed, however, is that this construction contains a hidden restriction which comes from the fact that the two explicit formulas for $P_{\overline{H}}$ and $Q$ in (2.39) are assumed to yield complementary projectors. Indeed, from the defining relation of $P_{\overline{H}}$, taking the later requirement $\nabla_\xi \mathbf{T}^H = 0$ into account, we have

$$\overline{X}^H = T(\mathbf{T}^H, \overline{X}^V) = \nabla_{\mathbf{T}^H}\overline{X}^V - [\mathbf{T}^H, \overline{X}^V],$$

from which it follows that

$$\nabla_{\mathbf{T}^H}\overline{X}^V = {[\mathbf{T}^H, \overline{X}^V]_V}^V.$$

On the other hand, we have

$$0 = Q(\overline{X}^H) = S(T(\mathbf{T}^H, \overline{X}^H)) = S(\nabla_{\mathbf{T}^H}\overline{X}^H) - S([\mathbf{T}^H, \overline{X}^H]).$$

Using the invariance of $S$, this implies that

$$\nabla_{\mathbf{T}^H}\overline{X}^V = {[\mathbf{T}^H, \overline{X}^H]_H}^V.$$

Compatibility of the two expressions for $\nabla_{\mathbf{T}^H}\overline{X}^V$ thus requires that $[P_H, S](\mathbf{T}^H, \overline{X}^H) = 0$, which is one of the conditions for having a SODE connection. In coordinates, if $\Gamma^l_0, \Gamma^l_k$ denote the connection coefficients of the given horizontal distribution, this condition reads,

$$v^j V_k(\Gamma^i_j) - \Gamma^i_k + V_k(\Gamma^i_0) = 0, \tag{2.40}$$

where $V_i$ is shorthand for $\partial/\partial v^i$. It can be verified that this is the only compatibility requirement coming from (2.39).

What remains to be verified now is whether such a connection belongs to the Berwald class. Obviously, from the assumptions $\nabla_\xi P_{\overline{H}} = 0$ and $\nabla_\xi S = 0$, we will have a connection of Finsler type and the question is whether the requirements of Definition 2.11 hold. It easily follows from the defining relation for $P_{\overline{H}}$ in (2.39) and the property $\nabla_\xi \mathbf{T}^H = 0$ that (2.34) holds for $\xi = \mathbf{T}^H$. But it turns out that there is an obstruction for the rest of the property to hold true. To see this, let $\overline{X}, \overline{Y}$ be basic vector fields in $\overline{\mathcal{X}}(\pi_M)$. From one of the first assumptions, we have $\nabla_{\overline{X}^V}\overline{Y}^V = 0$, from which it follows via (2.29) that also $\nabla_{\overline{X}^V}\overline{Y}^H = 0$. Next, using the $\nabla$-invariance of $\mathbf{T}^H$, one of the defining relations of the torsion tensor gives

$$\begin{aligned}[\mathbf{T}^H, \overline{X}^V] &= \nabla_{\mathbf{T}^H}\overline{X}^V - T(\mathbf{T}^H, \overline{X}^V), \\ &= \nabla_{\mathbf{T}^H}\overline{X}^V - \overline{X}^H, \end{aligned} \tag{2.41}$$

where we have used the definition (2.39) for $P_{\overline{H}}$ in the last line. The curvature requirement $curv\,(\mathbf{T}^H, \overline{X}^V) = 0$ subsequently implies that

$$\nabla_{\overline{X}^H} = \nabla_{\overline{X}^V}\nabla_{\mathbf{T}^H} - \nabla_{\mathbf{T}^H}\nabla_{\overline{X}^V} + \nabla_{\nabla_{\mathbf{T}^H}\overline{X}^V}. \tag{2.42}$$

Applying (2.42) to $\overline{Y}^V$, the last two terms vanish because $\overline{Y}$ is basic and $\nabla_{\mathbf{T}^H}\overline{X}^V$ is vertical in view of the properties (2.27). To compute the remaining vector field $\nabla_{\overline{X}^V}\nabla_{\mathbf{T}^H}\overline{Y}^V$ we proceed in coordinates. Using (2.41), one easily verifies that

$$\nabla_{\mathbf{T}^H}V_j = (V_j(\Gamma^l_0) + v^k V_j(\Gamma^l_k))V_l.$$

With $\overline{X}^V = X^i V_i$, $\overline{Y}^V = Y^i V_i$ ($X^i, Y^i$ basic), we further have

$$\nabla_{\overline{X}^V}\nabla_{\mathbf{T}^H}\overline{Y}^V = \overline{X}^H(Y^i)V_i + Y^i X^j \nabla_{V_j}\nabla_{\mathbf{T}^H}V_i.$$

It then easily follows that

$$\nabla_{\overline{X}^H}\overline{Y}^V = [\overline{X}^H, \overline{Y}^V] + X^j Y^i \left(V_i V_j(\Gamma^l_0) + v^k V_i V_j(\Gamma^l_k)\right) V_l.$$

As we have seen in the table of torsion components, however, $T(\overline{X}^H, \overline{Y}^V) = 0$ is a necessary requirement for a connection to be of Berwald type and this would require here that

$$V_i V_j(\Gamma^l_0) + v^k V_i V_j(\Gamma^l_k) = 0. \tag{2.43}$$

It is easy to see through its two components $\mathcal{B}$ and $\mathcal{P}$ that the vanishing of this torsion is also sufficient for having the Berwald condition (2.34). The final point to observe is that (2.40) and (2.43) imply that

$$V_k(\Gamma^i_l) - V_l(\Gamma^i_k) = 0, \tag{2.44}$$

which is the coordinate expression for having $[P_H, S](\overline{X}^H, \overline{Y}^H) = 0$. We reach the rather striking conclusion that our attempt to generalise the construction of Massa and Pagani to arbitrary horizontal distributions only gives rise to a connection of Berwald type if that distribution is actually a SODE connection (which then brings us back to the actual construction in [57]).

Limiting ourselves then to the SODE case, the main difference between this linear connection and the one of Crampin *et al* comes from the fact that here $\nabla_\xi \mathbf{T}^H = 0$ for all $\xi$. The effect on the torsion is merely that $\mathcal{B}_{\mathbf{T}}$ is no longer zero. Instead we have $\mathcal{B}_{\mathbf{T}} = -\overline{I}$.

Let us come now to the third construction, which was independently set up by Byrnes [9]. Again, the original construction was carried out starting from a SODE connection, but we can easily generalise it here to the case of an arbitrary horizontal distribution. Indeed, the main idea of the construction of Byrnes was simply the following: (i) define the covariant derivatives of vector fields in $\overline{\mathcal{X}}(J^1M)$ by looking at the formula (1.4) for $\mathrm{D}_\xi X$ on the pullback bundle in the autonomous framework and taking horizontal and vertical lifts as appropriate; (ii) put $\nabla_\Gamma \Gamma = 0$; (iii) select the remaining derivatives of $\Gamma$ in such a way that all torsion components which can be zero effectively vanish. By the nature of the construction, therefore, this is bound to give a connection belonging to the Berwald class. Transferred to the context of a general horizontal distribution, this idea becomes: (i) define $\nabla_\xi \overline{X}^H$ and $\nabla_\xi \overline{X}^V$ via (2.25) with $\mathrm{D}_\xi \overline{X}$ given by (2.34); (ii) put $\nabla_{\mathbf{T}^H} \mathbf{T}^H = 0$; (iii) define $\nabla_{\overline{X}^V} \mathbf{T}^H$ and $\nabla_{\overline{X}^H} \mathbf{T}^H$ in such a way that the last four torsion components in the above table all vanish. This of course means then that the tensor $K$ in (2.25) is constructed in a rather ad hoc manner.

Going back to the special case of a SODE connection, the only difference with our analysis of the first construction is that now also $\mathcal{R}_{\mathbf{T}} = 0$. Since the vertical part of the bracket $[\Gamma, \overline{X}^H]$ is determined by the so-called Jacobi endomorphism $\Phi$, which is essentially the time-component of the curvature of the non-linear connection (see e.g. (2.20)), we could say here that the construction of Byrnes boils down to choosing the tensor $K$ in (2.25) as:

$$K(\xi) = \overline{\xi}_V{}^H - \Phi(\overline{\xi}_H)^V. \tag{2.45}$$

Observe that from this point of view, i.e. if one regards the $\nabla$ under consideration as being constructed from a D on $\pi_M^* \tau_M$, the selection of $K$ that was made in the construction of Massa and Pagani was simply $K = 0$.

We have now completed our programme of defining the class of Berwald-type connections in a sufficiently general way to be able to accommodate the constructions of Crampin *et al*, Massa and Pagani and Byrnes, and we have discovered the features which distinguish these constructions in that process. Can we, on the basis of these features, find reasons why one of these constructions should have preference over the others? If the ideal for a Berwald-type connection would be, as in the autonomous case, to have as much torsion zero as possible, then obviously the last construction would prevail. But it looks a lot less natural than the first one, for example, which is based on two direct formulas: (2.35) for the linear connection on $\pi_M^*\tau_M$ and (2.36) for its lift to a connection on $\tau_{J^1M}$. The construction of Massa and Pagani deviates even further from the idea of maximally vanishing torsion, but we will now argue that it has a different interesting feature which the others fail to produce. In previous sections, we have emphasised the importance of the natural decompositions (2.1) and (2.13) of the sections under consideration. Yet, when introducing Finsler-type connections, we required only part of that decomposition to be preserved by the covariant derivatives: see (2.24) for D and (2.27) for $\nabla$. It would seem to be a natural assumption also to expect that these operators in addition would have the property

$$\mathrm{D}_\xi(\langle \mathbf{T}\rangle) \subset \langle \mathbf{T}\rangle, \qquad \text{respectively} \qquad \nabla_\xi(\langle \mathbf{T}^H\rangle) \subset \langle \mathbf{T}^H\rangle.$$

In this respect, only the construction of Massa and Pagani would be satisfactory in view of the property $\nabla_\xi \mathbf{T}^H = 0$.

Going back to our definition of the class of Berwald-type connections, it is obvious that the selection of a particular representative of the class is a matter of making a choice for $\mathrm{D}_\xi \mathbf{T}$ (when it concerns a connection on $\pi_M^*\tau_M$) or for $\nabla_\xi \mathbf{T}^H$ (for a connection on $\tau_{J^1M}$). Clearly, there is much to say for giving preference to the simplest possible choice where these vector fields would both be zero. Note, however, that this would indirectly impose a restriction also on the freedom in lifting the connection (the choice of $K$ in (2.25)) or lowering it (the choice of $L$ in (2.29)). In the next section, therefore, we will explore some other interesting features of the theory, with an eye on discovering additional elements which can tell us whether there is a certain degree of optimality in choosing the simplest possible representative.

## 2.8 Further aspects of connections of Finsler and Berwald type

Recall that the only restriction so far considered for connections on $\pi_M^*\tau_M$ was the requirement (2.24). It can equivalently be expressed as

$$\mathrm{D}_\xi \overline{I}|_{\overline{\mathcal{X}}(\pi_M)} = 0.$$

If a horizontal distribution is given and we lift the linear connection to one on $\tau_{J^1M}$ via (2.25), we have seen from (2.32, 2.33) that an immediate consequence is:

$$\nabla_\xi P_{\overline{H}}|_{\overline{\mathcal{X}}(J^1M)} = \nabla_\xi J|_{\overline{\mathcal{X}}(J^1M)} = \nabla_\xi S|_{\overline{\mathcal{X}}(J^1M)} = 0.$$

This should not come as a surprise as all the tensor fields under consideration here can in fact be constructed out of $\overline{I}$ via appropriate lifting operations. To be precise, we have

$$P_{\overline{H}} = \overline{I}^{H;H}, \quad J = \overline{I}^{H;V} - \overline{I}^{V;H} \quad \text{and} \quad S = \overline{I}^{H;V}.$$

These lifts, introduced in [80], are defined as follows for a general type (1,1) tensor field $U$ along $\pi_M$:

$$\begin{aligned} &U^{H;H}(X^H) = U(X)^H, && U^{H;H}(\overline{X}^V) = 0,\\ &U^{H;V}(X^H) = U(X)^V, && U^{H;V}(\overline{X}^V) = 0,\\ &U^{V;H}(X^H) = 0, && U^{V;H}(\overline{X}^V) = U(\overline{X})^H,\\ &U^{V;V}(X^H) = 0, && U^{V;V}(\overline{X}^V) = U(\overline{X})^V. \end{aligned} \tag{2.46}$$

The interest of these operations is, as with the horizontal and vertical lifts of vector fields along $\pi_M$, that every type (1,1) tensor field $\mathcal{U}$ on $J^1M$ has a unique decomposition in the form

$$\mathcal{U} = U_1^{H;H} + U_2^{H;V} + U_3^{V;H} + U_4^{V;V}, \tag{2.47}$$

where the $U_i$ are tensor fields along $\pi_M$ which have the following characteristics: $U_1$ is general, $U_2(\mathcal{X}(\pi_M)) \subset \overline{\mathcal{X}}(\pi_M)$, $U_3(\mathbf{T}) = 0$ and $U_4$ has the properties of $U_2$ and $U_3$. For Finsler-type connections, covariant derivatives of a $\mathcal{U}$ on $J^1M$ should to some extent be computable from the covariant derivatives of the $U_i$ along $\pi_M$ which generate it. Ideally, of course, the latter should preserve the characteristic properties of each of the $U_i$.

If no further restrictions are imposed on the freedom in the procedures for raising or lowering the linear connection (see (2.25) and (2.29)), one can prove that for a $\mathcal{U}$ which maps $\overline{\mathcal{X}}(J^1M)$ into itself (the corresponding $U_i$ in (2.47) then map $\overline{\mathcal{X}}(\pi_M)$ into itself), $\nabla_\xi \mathcal{U}|_{\overline{\mathcal{X}}(J^1M)} = 0$ if and only if $\mathrm{D}_\xi U_i|_{\overline{\mathcal{X}}(\pi_M)} = 0$, $i = 1, \ldots, 4$. But we may hope to discover natural additional restrictions as soon as we attempt to extend the scope of such a statement beyond the action on $\overline{\mathcal{X}}(\pi_M)$. It turns out that the very first restriction which imposes itself in this respect is to have the following direct link between the covariant derivatives of $\mathbf{T}$ and $\mathbf{T}^H$:

$$\nabla_\xi \mathbf{T}^H = (\mathrm{D}_\xi \mathbf{T})^H. \tag{2.48}$$

Indeed, we can state the following result.

**Proposition 2.12.** *Under the assumption (2.48), we have for an arbitrary type (1,1) tensor field $\mathcal{U}$ on $J^1M$*

$$\nabla_\xi \mathcal{U}(\eta) = \mathrm{D}_\xi U_1(\eta_H)^H + \mathrm{D}_\xi U_2(\eta_H)^V + \mathrm{D}_\xi U_3(\overline{\eta}_V)^H + \mathrm{D}_\xi U_4(\overline{\eta}_V)^V. \tag{2.49}$$

Proof: The idea is to compute $\nabla_\xi \mathcal{U}(\eta)$ for arbitrary $\xi$ and $\eta$ and $\mathcal{U}$ in its decomposition (2.47). At the start, we only assume that $\nabla$ comes via (2.25) from some D with property (2.24); we want to find out which further restrictions (if any) impose themselves in a natural way for obtaining a closed form expression such as (2.49).

Let us start by looking in detail at the term $\nabla_\xi U_3^{V;H}(\eta)$, knowing that $U_3$ vanishes on $\mathbf{T}$. We have

$$\begin{aligned}
\nabla_\xi U_3^{V;H}(\eta) &= \nabla_\xi \left(U_3(\overline{\eta}_V)^H\right) - U_3^{V;H}\big((\mathrm{D}_\xi \overline{\eta}_H)^H + (\mathrm{D}_\xi \overline{\eta}_V)^V + \xi(\langle \eta, dt\rangle)\mathbf{T}^H \\
&\quad + \langle \eta, dt\rangle \nabla_\xi \mathbf{T}^H\big) \\
&= \nabla_\xi \left(\overline{U_3(\overline{\eta}_V)}^H + \langle U_3(\overline{\eta}_V), dt\rangle \mathbf{T}^H\right) - U_3(\mathrm{D}_\xi \overline{\eta}_V)^H \\
&\quad - \langle \eta, dt\rangle U_3^{V;H}(\nabla_\xi \mathbf{T}^H) \\
&= \left(\mathrm{D}_\xi (\overline{U_3(\overline{\eta}_V)})\right)^H + \xi\left(\langle U_3(\overline{\eta}_V), dt\rangle\right)\mathbf{T}^H + \langle U_3(\overline{\eta}_V), dt\rangle \nabla_\xi \mathbf{T}^H \\
&\quad - U_3(\mathrm{D}_\xi \overline{\eta}_V)^H - \langle \eta, dt\rangle U_3^{V;H}(\nabla_\xi \mathbf{T}^H) \\
&= (\mathrm{D}_\xi (U_3(\overline{\eta}_V)))^H - \langle U_3(\overline{\eta}_V), dt\rangle (\mathrm{D}_\xi \mathbf{T})^H + \langle U_3(\overline{\eta}_V), dt\rangle \nabla_\xi \mathbf{T}^H \\
&\quad - U_3(\mathrm{D}_\xi \overline{\eta}_V)^H - \langle \eta, dt\rangle U_3^{V;H}(\nabla_\xi \mathbf{T}^H) \\
&= (\mathrm{D}_\xi U_3(\overline{\eta}_V))^H + \langle U_3(\overline{\eta}_V), dt\rangle (\nabla_\xi \mathbf{T}^H - (\mathrm{D}_\xi \mathbf{T})^H) \\
&\quad - \langle \eta, dt\rangle U_3^{V;H}(\nabla_\xi \mathbf{T}^H).
\end{aligned}$$

Under the condition (2.48), this reduces to

$$\nabla_\xi U_3^{V;H}(\eta) = (\mathrm{D}_\xi U_3(\overline{\eta}_V))^H.$$

The computation for $U_4$ is quite similar. Since $U_4$ takes values in $\overline{\mathcal{X}}(\pi_M)$, there is in fact a further simplification, we find:

$$\nabla_\xi U_4^{V;V}(\eta) = (\mathrm{D}_\xi U_4(\overline{\eta}_V))^V - \langle\eta, dt\rangle U_4^{V;V}(\nabla_\xi \mathbf{T}^H),$$

from which it follows under the same assumption (2.48) that

$$\nabla_\xi U_4^{V;V}(\eta) = (\mathrm{D}_\xi U_4(\overline{\eta}_V))^V.$$

For $U_1$ we have, taking this time (2.48) already into account,

$$\begin{aligned}
\nabla_\xi U_1^{H;H}(\eta) &= \nabla_\xi \left(U_1(\eta_H)^H\right) - U_1^{H;H}\big((\mathrm{D}_\xi\overline{\eta}_H)^H + \xi(\langle\eta, dt\rangle)\mathbf{T}^H \\
&\quad +\langle\eta, dt\rangle\nabla_\xi\mathbf{T}^H\big) \\
&= \nabla_\xi \left(\overline{U_1(\eta_H)}^H + \langle U_1(\eta_H), dt\rangle\mathbf{T}^H\right) - U_1(\mathrm{D}_\xi\overline{\eta}_H)^H \\
&\quad - \xi(\langle\eta, dt\rangle)U_1(\mathbf{T})^H - \langle\eta, dt\rangle U_1(\mathrm{D}_\xi\mathbf{T})^H \\
&= \left(\mathrm{D}_\xi(\overline{U_1(\eta_H)})\right)^H \\
&\quad +\xi(\langle U_1(\eta_H), dt\rangle)\mathbf{T}^H + \langle U_1(\eta_H), dt\rangle\nabla_\xi\mathbf{T}^H - U_1(\mathrm{D}_\xi\eta_H)^H \\
&= (\mathrm{D}_\xi(U_1(\eta_H)))^H - U_1(\mathrm{D}_\xi\eta_H)^H \\
&= \mathrm{D}_\xi U_1(\eta_H)^H.
\end{aligned}$$

The computation for $U_2$ is similar, with an extra simplification again because $U_2$ takes values in $\overline{\mathcal{X}}(\pi_M)$. We find

$$\nabla_\xi U_2^{H;V}(\eta) = \mathrm{D}_\xi U_2(\eta_H)^V,$$

which completes the proof. □

Note finally that the statement preceding Proposition 2.12 can easily be proved from the above computations as well. Indeed, if each of the $U_i$ maps $\overline{\mathcal{X}}(\pi_M)$ into itself and we restrict ourselves to such vector field arguments, none of the terms which prompted the assumption (2.48) will occur.

The meaning of the extra condition (2.48) is the following. If D is the linear connection we start from, then the raising procedure (2.25) with $K(\xi) = (\mathrm{D}_\xi\mathbf{T})^H$ corresponds exactly to the quite natural expression (2.36). If $\nabla$ is the starting point, then the tensor $L$ in (2.29) must be chosen in such a way that $L(\xi)^H = \nabla_\xi\mathbf{T}^H$, which is possible only if $\nabla_\xi\mathbf{T}^H$ is horizontal. Clearly, this is not the case for (our generalised version of) the construction of Byrnes, which means that it is rather unnatural to pursue maximally zero torsion in the time-dependent set-up: one should not insist on having $\mathcal{R}_{\mathbf{T}} = 0$. This is hardly surprising as $\mathcal{R}_{\mathbf{T}}$, just as $\mathcal{R}$, itself is related to the curvature of the non-linear connection one starts from.

**Corollary 2.13.** *Under the assumption of Proposition 2.12, we have*

$$\nabla_\xi \mathcal{U} = 0 \quad \Leftrightarrow \quad \begin{cases} \mathrm{D}_\xi U_1(\eta_H) = 0, & \mathrm{D}_\xi U_2(\eta_H) = 0, \\ \mathrm{D}_\xi U_3(\overline{\eta}_V) = 0, & \mathrm{D}_\xi U_4(\overline{\eta}_V) = 0, \end{cases} \qquad \forall \eta_H, \overline{\eta}_V. \tag{2.50}$$

Proof: The proof is almost immediate. The only point to be careful about is that for the vertical parts in (2.49) the immediate conclusion is that the component in $\overline{\mathcal{X}}(\pi_M)$ of the corresponding vector field along $\pi_M$ must be zero. But $U_2$ and $U_4$ take their values in $\overline{\mathcal{X}}(\pi_M)$ and the property (2.24) of D then ensures that the same is true for their covariant derivatives. □

The final point to observe is that the above results do not necessarily imply that the special features of the tensor fields $U_i$ are preserved under covariant differentiation. One of the consequences then is that (2.50) in general is not sufficient to conclude that $\mathrm{D}_\xi U_i = 0$, $\forall i$. As a matter of fact, knowing that $U_3(\mathbf{T}) = 0$, we have $\mathrm{D}_\xi U_3(\mathbf{T}) = -U_3(\mathrm{D}_\xi \mathbf{T})$. It then follows that $\mathrm{D}_\xi U_3(\overline{\eta}_V) = 0$, $\forall \overline{\eta}_V$ implies $\mathrm{D}_\xi U_3 = 0$ if and only if

$$\mathrm{D}_\xi \mathbf{T} \in \langle \mathbf{T} \rangle. \tag{2.51}$$

The same is true for $U_4$. We thus have proved the following result.

**Corollary 2.14.** *If (2.48) holds together with (2.51), we have* $\nabla_\xi \mathcal{U} = 0$ *if and only if* $\mathrm{D}_\xi U_i = 0$, $i = 1, \ldots, 4$.

The linear connection (2.35) on $\pi_M^* \tau_M$ as constructed in [23] does not have the property (2.51). The above considerations will prompt us to an improvement of the construction (2.35) in the next section.

## 2.9 Is there an optimal Berwald-type connection?

We will now attempt to come to an 'optimal' choice of a representative of the class of Berwald-type connections associated to an arbitrary horizontal distribution on $J^1M$. Obviously, such a choice should combine all the good features we have encountered in discussing the different faces of the theory in the preceding sections. As we have seen, the essence of all such connections (as soon as they are of Finsler type) lies in a linear connection on $\pi_M^*\tau_M$. So, in the first place, we want an explicit construction formula for a linear connection on $\pi_M^*\tau_M$ which, unlike the explicit formula (2.35) of [23], does have the additional property (2.51) for preserving the natural decomposition (2.1). Secondly, we want to decide about an explicit rule for raising the linear connection to $\tau_{J^1M}$ which will then determine the 'optimal' Berwald-type connection there. Preferably, there should also be an explicit expression for the inverse of this rule.

As explained in Section 2.6, the idea of the direct construction formula (2.35) was simply to copy the known formula (1.4) from the autonomous framework and see what correction terms are needed to have the right derivation properties for a linear connection on $\pi_M^*\tau_M$. This way, one is guaranteed to arrive at a generalisation which will give back the original theory when restricting to objects which are time-independent. There is, however, another way in which such an idea can be carried out: it consists in "copying the formula from the autonomous theory" with $\overline{X}$ in the place of $X$ and then see what correction is needed to have a linear connection on $\pi_M^*\tau_M$ again. This way, one arrives at the following explicit formula:

$$\mathrm{D}_\xi X = [P_H(\xi), \overline{X}^V]_V + [P_V(\xi), \overline{X}^H]_H + \xi(\langle X, dt\rangle)\mathbf{T}. \tag{2.52}$$

It is immediately clear that this linear connection has the property (2.51) since it is in fact the simplest representative for which $\mathrm{D}_\xi\mathbf{T} = 0$ for all $\xi$.

There is little doubt about the choice of an optimal lifting procedure now. Indeed, the further aspects of Finsler-type connections explored in the preceding section have revealed that it is advantageous to have the property (2.48), which will imply here that also $\nabla_\xi\mathbf{T}^H = 0$. The raising procedure then is just the natural one (2.36). Looking at the table of torsion components of Section 2.6, our candidate for an optimal Berwald-type connection on $\tau_{J^1M}$ will have $\mathcal{B} = \mathcal{P} = \mathcal{S} = \mathcal{P}_\mathbf{T} = 0$ and $\mathcal{B}_\mathbf{T} = -\overline{I}$. If in particular the horizontal distribution comes from a Sode, we know that in addition $\mathcal{A} = 0$ and we will also have here $\mathcal{A}_\mathbf{T} = 0$. In the case of a Sode connection therefore, our optimal Berwald-type connection on $\tau_{J^1M}$ is just the linear connection constructed in [57].

There remains the question about an explicit formula for the inverse procedure of lowering a linear connection on $\tau_{J^1M}$ to one on $\pi_M^*\tau_M$. Such a formula of course must have the properties (2.29) and can simply be taken to be

$$\mathrm{D}_\xi X = (\nabla_\xi X^H)_H, \qquad \forall X \in \mathcal{X}(\pi_M). \tag{2.53}$$

As an aside, note that there is another explicit formula by which a $\nabla$ on $\tau_{J^1M}$ can be lowered to a D on $\pi_M^*\tau_M$, namely

$$\mathrm{D}_\xi X = (\nabla_\xi X^V)_V + \xi(\langle X, dt\rangle)\mathbf{T}, \qquad \forall X \in \mathcal{X}(\pi_M). \tag{2.54}$$

In the case of our optimal Berwald-type connection on $\tau_{J^1M}$, these two procedures give rise to the same D, thanks to the property $\nabla_\xi \mathbf{T}^H = 0$. By contrast, for example, if we were to start from the linear connection (2.35), raise it to $\tau_{J^1M}$ via (2.36) and subsequently come back to a linear connection on $\pi_M^*\tau_M$ via the procedure (2.54), we would not end up with the linear connection we started from, but rather with the linear connection (2.52).

Summarising what precedes, we come to the following formal definition.

**Definition 2.15.** *The optimal Berwald-type connection on $\pi_M^*\tau_M$, associated to a given horizontal distribution on $\pi_M$, is defined explicitly by (2.52). The corresponding Berwald-type connection on $\tau_{J^1M}$ is produced by (2.36).*

Of course, one should interpret the adjective 'optimal' here only in the sense 'optimal in view of the restrictions (2.48) and (2.51)'. We are conscious that it might be possible to produce different arguments to prefer another linear connection to our 'optimal one' (2.52). In the following we will use the symbol $\hat{\mathrm{D}}$ for the linear connection defined by (2.52).

## 2.10 Derived constructions when a metric tensor field along $\pi_M$ is available

Suppose now that we have an additional tool at our disposal, namely a symmetric type (0,2) tensor field $g$ along $\pi_M$, having the property $g(\mathbf{T}, .) = 0$ and being non-singular when restricted to $\overline{\mathcal{X}}(\pi_M)$. We would like then to generalise the concepts (1.6-1.7) of the autonomous framework to arrive in the end at suitable generalisations of connections of the type of Cartan, Chern–Rund and Hashiguchi. It should be emphasised at this point that the context in which we wish to achieve this is far more general than the case of geodesic sprays on a Finsler manifold: both the horizontal distribution we start from and the tensor field $g$ are (apart from the restrictions on $g$ mentioned above) completely arbitrary and need not have anything to do with each other.

The main point about a Cartan-type connection is that it should be fully metrical and that the other two should be horizontally or vertically metrical only. There is, however, not a unique way of achieving such properties, even though from now on we agree that the Berwald-type connection we start from is fixed by (2.52). As we learn for example from [67] (Chapter X, Theorem 2.4) (see also [65]), there is a lot of freedom still in pursuing the idea of constructing a metrical connection. One way to proceed here, for example, would be to define Cartan-type tensor fields $\mathcal{C}_V$ and $\mathcal{C}_H$ exactly as in equations (1.6,1.7), at least when all arguments are elements of $\overline{\mathcal{X}}(\pi_M)$. This may seem to be the most direct way to proceed. We prefer, however, to define $\mathcal{C}_V$ and $\mathcal{C}_H$ in this general context in a different way; it will lead to a metrical connection which is more closely related to the work of the Miron school on what they call "generalized Lagrange spaces" (cf. [2, 67, 65]).

**Definition 2.16.** *The vertical and horizontal Cartan tensor fields associated to the Berwald-type connection (2.52) and the metric tensor field $g$ along $\pi_M$, are type (1,2) tensor fields $\mathcal{C}_V$ and $\mathcal{C}_H$ along $\pi_M$, determined by the relations*

$$g(\mathcal{C}_V(\overline{X},\overline{Y}),\overline{Z}) = \hat{\mathrm{D}}_{\overline{X}^V} g(\overline{Y},\overline{Z}) + \hat{\mathrm{D}}_{\overline{Y}^V} g(\overline{X},\overline{Z}) - \hat{\mathrm{D}}_{\overline{Z}^V} g(\overline{X},\overline{Y}) \quad (2.55)$$

$$g(\mathcal{C}_H(X,\overline{Y}),\overline{Z}) = \hat{\mathrm{D}}_{X^H} g(\overline{Y},\overline{Z}) + \hat{\mathrm{D}}_{\overline{Y}^H} g(X,\overline{Z}) - \hat{\mathrm{D}}_{\overline{Z}^H} g(X,\overline{Y}) \quad (2.56)$$

*and by the following restrictions for fixing the remaining time-components:* $\mathcal{C}_V(\,.\,,\mathbf{T}) = \mathcal{C}_V(\mathbf{T},\,.\,) = 0,\ \mathcal{C}_H(\,.\,,\mathbf{T}) = 0.$

Thinking then of another linear connection on $\pi_M^*\tau_M$, D say, which differs from the Berwald-type connection by a tensor field $\delta$, i.e. such that $\mathrm{D}_\xi X - \hat{\mathrm{D}}_\xi X = \delta(\xi, X)$, we introduce type (1,2) tensor fields $\delta^V$ and $\delta^H$ along $\pi_M$, defined by

$$\delta^V(\overline{Z},X) = \delta(\overline{Z}^V,X), \qquad \delta^V(\mathbf{T},X) = 0, \quad (2.57)$$

$$\delta^H(Z,X) = \delta(Z^H,X). \quad (2.58)$$

Having optimalised the freedom in the class of Berwald-type connections by making $\hat{\mathrm{D}}_\xi\mathbf{T} = 0$, we will do the same for the derived connections related to $g$ which we will now discuss. That is to say, we choose to have also $\mathrm{D}_\xi\mathbf{T} = 0$, which implies that $\delta^V(Z,\mathbf{T}) = 0$ and $\delta^H(Z,\mathbf{T}) = 0$ (but $\delta^H(\mathbf{T},X)$ need not be zero). This selection makes the following definitions perfectly compatible with the properties of the tensor fields $\mathcal{C}_V$ and $\mathcal{C}_H$ introduced above.

**Definition 2.17.** *The Cartan-type connection $\overset{\mathrm{C}}{\mathrm{D}}$ on $\pi_M^*\tau_M$, associated to the given metric tensor field $g$ along $\pi_M$, deviates from the optimal Berwald-type* *connection by*

$$\delta^{V} = \tfrac{1}{2}\mathcal{C}_{V}, \qquad \delta^{H} = \tfrac{1}{2}\mathcal{C}_{H}. \tag{2.59}$$

*The Hashiguchi-type connection* $\overset{\mathrm{H}}{\mathrm{D}}$ *is likewise defined by*

$$\delta^{V} = \tfrac{1}{2}\mathcal{C}_{V}, \qquad \delta^{H} = 0. \tag{2.60}$$

*Finally, the connection of Chern–Rund type* $\overset{\mathrm{CR}}{\mathrm{D}}$ *is determined by*

$$\delta^{V} = 0, \qquad \delta^{H} = \tfrac{1}{2}\mathcal{C}_{H}. \tag{2.61}$$

**Proposition 2.18.** *The Cartan-type connection is metrical in the sense that* $\overset{\mathrm{C}}{\mathrm{D}}_{\xi} g = 0$, $\forall \xi$. *For the connection of Hashiguchi type we have* $\overset{\mathrm{H}}{\mathrm{D}}_{\overline{X}^{V}} g = 0$, *while for the connection of Chern–Rund type:* $\overset{\mathrm{CR}}{\mathrm{D}}_{X^{H}} g = 0$.

Proof: Let us see what the meaning is of, for example, the assumption $\delta^{H} = \frac{1}{2}\mathcal{C}_{H}$. We have

$$\begin{aligned}
\overset{\mathrm{C}}{\mathrm{D}}_{X^H} g\,(\overline{Y}, \overline{Z}) &= \overset{\mathrm{C}}{\mathrm{D}}_{X^H}(g(\overline{Y}, \overline{Z})) - g(\overset{\mathrm{C}}{\mathrm{D}}_{X^H}\overline{Y}, \overline{Z}) - g(\overline{Y}, \overset{\mathrm{C}}{\mathrm{D}}_{X^H}\overline{Z}) \\
&= \hat{\mathrm{D}}_{X^H}(g(\overline{Y}, \overline{Z})) - g(\hat{\mathrm{D}}_{X^H}\overline{Y}, \overline{Z}) - g(\overline{Y}, \hat{\mathrm{D}}_{X^H}\overline{Z}) \\
&\quad - g(\delta^{H}(X, \overline{Y}), \overline{Z}) - g(\overline{Y}, \delta^{H}(X, \overline{Z})) \\
&= \hat{\mathrm{D}}_{X^H} g\,(\overline{Y}, \overline{Z}) - \tfrac{1}{2}\left(g(\mathcal{C}_{H}(X, \overline{Y}), \overline{Z}) + g(\overline{Y}, \mathcal{C}_{H}(X, \overline{Z})\right),
\end{aligned}$$

and all terms on the right cancel out when the defining relation (2.56) is used to replace the terms involving $\mathcal{C}_{H}$. $\overset{\mathrm{C}}{\mathrm{D}}_{X^H} g$ further inherits the property of $g$ of vanishing whenever one of the arguments is $\mathbf{T}$, therefore $\overset{\mathrm{C}}{\mathrm{D}}_{X^H} g = 0$. The meaning of the assumption $\delta^{V} = \frac{1}{2}\mathcal{C}_{V}$ is similar. The statements of the proposition now immediately follow. □

Needless to say, as in the autonomous case, making the connection more metrical has the effect of having less of the torsion components equal to zero. Without going into the details here, it is worth mentioning that the advantage of taking (2.55,2.56) as defining relations for the tensors $\mathcal{C}_{V}$ and $\mathcal{C}_{H}$ (rather than a direct transcription of (1.6,1.7) which would only have the first term in the right-hand side) is that more of the torsion components are zero.

## 2.11 Coordinate expressions

We wish to make the different levels of generality and the different types of linear connections which have been considered in the previous section a bit more perceptible by presenting a survey now of the relevant coordinate expressions in each case. This will make it easier for the reader to compare our results with related features in, for example, the books of Miron and Anastasiei [67] and Antonelli et al [4], where the theory is often developed through coordinate calculations only.

At the first level, all that is given is an arbitrary horizontal distribution and we can simply express the corresponding Berwald-type connection $\hat{\mathrm{D}}$ from (2.52). If in addition a metric tensor field $g$ along $\pi_M$ is given, we list the coordinate expressions for the tensor fields $\mathcal{C}_V$ and $\mathcal{C}_H$ defined by (2.55,2.56) and the connection coefficients for the resulting Cartan-type connection. For a second stage, we look at the special interest case where the horizontal distribution comes from an arbitrary SODE $\Gamma$ on $J^1M$. Finally, we have a closer look at the particular case when both the SODE $\Gamma$ and the tensor field $g$ are determined by a regular Lagrangian function $L$.

Using $\mathbf{T}^H$ (as in (2.12)), $H_i$ (as in (2.6)) and $V_i = \frac{\partial}{\partial v^i}$ as local basis for vector fields on $J^1M$, a straightforward application of the defining relation (2.52) shows that the Berwald-type connection on $\pi_M^*\tau_M$ is determined by

$$\hat{\mathrm{D}}_{\mathbf{T}^H}\frac{\partial}{\partial x^j} = \kappa^k_j\frac{\partial}{\partial x^k}, \quad \hat{\mathrm{D}}_{H_i}\frac{\partial}{\partial x^j} = V_j(\Gamma^k_i)\frac{\partial}{\partial x^k}, \quad \hat{\mathrm{D}}_{V_i}\frac{\partial}{\partial x^j} = 0, \qquad (2.62)$$

where

$$\kappa^k_j = V_j(\Gamma^k_0) + v^l V_j(\Gamma^k_l),$$

and of course $\hat{\mathrm{D}}_{\mathbf{T}^H}\mathbf{T} = \hat{\mathrm{D}}_{H_i}\mathbf{T} = \hat{\mathrm{D}}_{V_i}\mathbf{T} = 0$. Since we will have, by construction, $\mathrm{D}_\xi\mathbf{T} = 0$ for all connections which follow, we will not repeat these zero-components below.

Assume next that a symmetric tensor field of the form $g = g_{ij}(t,x,v)\,\theta^i\otimes\theta^j$ is given. Then, it follows from (2.55) that the vertical Cartan tensor $\mathcal{C}_V$ is of the form $\mathcal{C}_V = {\mathcal{C}_V}^k_{ij}\,\theta^i\otimes\theta^j\otimes(\partial/\partial x^k)$, with

$${\mathcal{C}_V}^k_{ij} = g^{kl}\Big(V_i(g_{lj}) + V_j(g_{li}) - V_l(g_{ij})\Big). \qquad (2.63)$$

The horizontal Cartan tensor, on the other hand, has a non-zero $dt$-component; it is of the form

$$\mathcal{C}_H = {\mathcal{C}_H}^k_{ij}\,\theta^i\otimes\theta^j\otimes\frac{\partial}{\partial x^k} + {\mathcal{C}_H}^k_{0i}\,dt\otimes\theta^i\otimes\frac{\partial}{\partial x^k},$$

with

$$\mathcal{C}_{H0i}^{\;k} = -\kappa_i^k + g^{kl}\Big(\mathbf{T}^H(g_{li}) - \kappa_l^m g_{mi}\Big), \tag{2.64}$$

$$\begin{aligned}\mathcal{C}_{Hij}^{\;k} &= -\Big(V_i(\Gamma_j^k) + V_j(\Gamma_i^k)\Big) + g^{kl}\Big(H_i(g_{lj}) + H_j(g_{li}) - H_l(g_{ij})\Big) \\ &\quad + g^{kl}\Big(g_{im}(V_j(\Gamma_l^m) - V_l(\Gamma_j^m)) + g_{jm}(V_i(\Gamma_l^m) - V_l(\Gamma_i^m))\Big). \end{aligned} \tag{2.65}$$

As a result, the Cartan-type connection along $\pi_M$, the way it is intrinsically defined by (2.59), is determined locally by the following relations:

$$\overset{\mathrm{C}}{\mathrm{D}}_{\mathbf{T}^H}\frac{\partial}{\partial x^j} = \Big[\tfrac{1}{2}\kappa_j^k + \tfrac{1}{2}g^{kl}\left(\mathbf{T}^H(g_{lj}) - \kappa_l^m g_{mj}\right)\Big]\frac{\partial}{\partial x^k}, \tag{2.66}$$

$$\begin{aligned}\overset{\mathrm{C}}{\mathrm{D}}_{H_i}\frac{\partial}{\partial x^j} &= \Big[\tfrac{1}{2}\Big(V_j(\Gamma_i^k) - V_i(\Gamma_j^k)\Big) + \tfrac{1}{2}g^{kl}\Big(H_i(g_{lj}) + H_j(g_{li}) - H_l(g_{ij})\Big) \\ &\quad + g^{kl}\Big(g_{im}\big(V_j(\Gamma_l^m) - V_l(\Gamma_j^m)\big) + g_{jm}\big(V_i(\Gamma_l^m) - V_l(\Gamma_i^m)\big)\Big)\Big]\frac{\partial}{\partial x^k}\end{aligned} \tag{2.67}$$

$$\overset{\mathrm{C}}{\mathrm{D}}_{V_i}\frac{\partial}{\partial x^j} = \tfrac{1}{2}g^{kl}\Big(V_i(g_{lj}) + V_j(g_{li}) - V_l(g_{ij})\Big)\frac{\partial}{\partial x^k}. \tag{2.68}$$

We leave it as an exercise for the reader to write down in the same way the local determining equations for the connections of Hashiguchi and of Chern–Rund type, as defined by (2.60) and (2.61) respectively.

Coming now to the second stage, let the horizontal distribution be the one canonically associated to a given SODE. The connection coefficients are then given by (2.16). The two conditions which essentially determine whether a non-linear connection is a SODE-connection, have already been mentioned in coordinates (see (2.40) and (2.44)). They read: $\kappa_j^i = \Gamma_j^i$ and $V_k(\Gamma_j^i) = V_j(\Gamma_k^i)$. The first of these has an immediate effect on the coefficients in the equations for the associated Berwald-type connection, which now become:

$$\hat{\mathrm{D}}_\Gamma\frac{\partial}{\partial x^j} = \Gamma_j^k\frac{\partial}{\partial x^k}, \quad \hat{\mathrm{D}}_{H_i}\frac{\partial}{\partial x^j} = V_j(\Gamma_i^k)\frac{\partial}{\partial x^k}, \quad \hat{\mathrm{D}}_{V_i}\frac{\partial}{\partial x^j} = 0. \tag{2.69}$$

The second results in obvious cancelations in the horizontal covariant derivative of the Cartan connection (still for an arbitrary metric tensor field $g$ along $\pi_M$). We get:

$$\overset{\mathrm{C}}{\mathrm{D}}_\Gamma\frac{\partial}{\partial x^j} = \Big[\tfrac{1}{2}\Gamma_j^k + \tfrac{1}{2}g^{kl}\left(\Gamma(g_{lj}) - \Gamma_l^m g_{mj}\right)\Big]\frac{\partial}{\partial x^k}, \tag{2.70}$$

$$\overset{\mathrm{C}}{\mathrm{D}}_{H_i}\frac{\partial}{\partial x^j} = \tfrac{1}{2}g^{kl}\Big(H_i(g_{lj}) + H_j(g_{li}) - H_l(g_{ij})\Big)\frac{\partial}{\partial x^k}, \tag{2.71}$$

$$\overset{\mathrm{C}}{\mathrm{D}}_{V_i}\frac{\partial}{\partial x^j} = \tfrac{1}{2}g^{kl}\Big(V_i(g_{lj}) + V_j(g_{li}) - V_l(g_{ij})\Big)\frac{\partial}{\partial x^k}. \tag{2.72}$$

Obviously, the elegance of this result is that both the horizontal and vertical covariant derivative resemble the classical formula for the Levi-Civita connection.

Consider now finally the particular case of a Lagrangian system. That is to say, let $L(t,x,v)$ be a given regular Lagrangian function on $J^1M$ and suppose that the coefficients of $g$ are $g_{ij} = V_iV_j(L)$. Let further $\Gamma$ denote the SODE field governing the Euler-Lagrange equations, i.e. take the $f^i$ in (2.11) to be (2.23). The Berwald-type connection remains determined by (2.69), but we can express the relevant coefficients $\Gamma^k_j$ and $V_j(\Gamma^k_i)$ here in terms of the Lagrangian $L$. One can verify that:

$$\Gamma^k_j = \tfrac{1}{2}g^{kl}\left(\Gamma(g_{lj}) + \frac{\partial^2 L}{\partial v^l \partial x^j} - \frac{\partial^2 L}{\partial x^l \partial v^j}\right), \tag{2.73}$$

$$\begin{aligned} V_j(\Gamma^k_i) = {} & \tfrac{1}{2}g^{kl}\Big(\Gamma V_j(g_{li}) + H_j(g_{li}) + H_i(g_{lj}) - H_l(g_{ij}) \\ & - \Gamma^m_l V_m(g_{ij}) - \Gamma^m_j V_m(g_{li}) - \Gamma^m_i V_m(g_{lj})\Big). \end{aligned} \tag{2.74}$$

Turning then to the Cartan-type connection for this case, the following simplifications of the previous situation can be verified. First of all, we obviously have $V_j(g_{li}) - V_l(g_{ij}) = 0$. Furthermore, the coefficients $g_{ij}$ will satisfy the property $\Gamma(g_{lj}) = \Gamma^m_l g_{mj} + \Gamma^m_j g_{ml}$ (in fact, this is one of the so-called Helmholtz conditions for the existence of a Lagrangian, see later in section 3.2). From this it easily follows that the right-hand side in (2.70) is equal to $\Gamma^k_j$ (i.e. is the same as for the Berwald connection, see also the final remark in section 3.2). As a result, the Cartan-type connection for the Lagrangian case is determined by

$$\overset{\mathrm{C}}{\mathrm{D}}_\Gamma \frac{\partial}{\partial x^j} = \Gamma^k_j \frac{\partial}{\partial x^k}, \tag{2.75}$$

$$\overset{\mathrm{C}}{\mathrm{D}}_{H_i} \frac{\partial}{\partial x^j} = \tfrac{1}{2}g^{kl}\Big(H_i(g_{lj}) + H_j(g_{li}) - H_l(g_{ij})\Big)\frac{\partial}{\partial x^k}, \tag{2.76}$$

$$\overset{\mathrm{C}}{\mathrm{D}}_{V_i} \frac{\partial}{\partial x^j} = \tfrac{1}{2}g^{kl}V_i(g_{lj})\frac{\partial}{\partial x^k}. \tag{2.77}$$

To finish this summary of coordinate expressions, let us repeat that one should be a little cautious in comparing our expressions with those in [67] for time-dependent Lagrangians. The point is that the set-up is different: due to a strict separation between time and space variables in [67], some of the concepts developed in that work loose there intrinsic meaning within the jet bundle approach which we have adopted.

## 2.12 An aside: From covariant derivatives to exterior derivatives and the classification of derivations

It is somehow intriguing that the construction of horizontal and vertical covariant derivative operators, the way they were derived from the classification theory of derivations of forms along $\pi_M$ in [80], gave rise exactly to the same, less optimal, construction of the linear connection (2.35) of Crampin *et al.* The branching point in the theory of derivations in [80] was a freedom in selecting a natural vertical exterior derivative. By way of application of the newly acquired insights, therefore, we will discuss here the reverse process, namely the way different choices for the linear connection on $\pi_M^*\tau_M : \pi_M^*TM \to J^1M$ affect the classification theory of derivations of forms along $\pi_M$.

A classification of derivations of scalar and vector-valued forms along $\pi_M$, in the line of the standard work of Frölicher and Nijenhuis [32], makes use of a vertical and horizontal exterior derivative. For the horizontal derivative one needs a horizontal distribution, while the vertical derivative is canonically available from the intrinsic structure of $J^1M$. Yet, not surprisingly, there is not just one canonically defined vertical exterior derivative: one encounters a certain freedom in fixing the time-component. Scalar differential forms along $\pi_M$ can be identified with semi-basic forms on $J^1M$ and there is a natural derivation of degree 1 on $J^1M$ which preserves semi-basic forms, namely (in the notations of [32]) $d_S = [i_S, d]$. To maintain the analogy with the autonomous theory, the authors in [80] decided to model their vertical exterior derivative $d^V$ on $d_S$, even though this derivation does not have the coboundary property $(d^V)^2 = 0$. The authors were well aware of the availability of another vertical derivation which does have that property. But from the point of view of setting up the theory of derivations, this other one comes somehow in the second place as it can be derived from $d^V$: it is the derivation $d^V_{\overline{I}} = [i_{\overline{I}}, d^V]$. Much later in the story of classifying derivations, one encounters vertical and horizontal covariant derivatives which appear to coincide with the ones coming from the linear connection (2.35) in [23].

The purpose of this final section is to approach this matter from the other end. That is to say, we wish to explore to what extent the optimal choice of a Berwald-type connection adds something to the debate about the best possible choice of a vertical exterior derivative.

Let us first discuss some generalities about the way to construct an exterior derivative from a covariant derivative. Suppose a covariant derivative

$\mathrm{D}^\star$ on $\pi_M^*\tau_M$ is given. We can extend it, by duality, to a (self-dual) degree 0 derivation on tensor fields along $\pi_M$ (the present discussion, by the way, applies just as well to covariant derivatives on a general manifold). Putting $[X,Y]_\star = \mathrm{D}^\star_X Y - \mathrm{D}^\star_Y X$, we have a bilinear (over $\mathbb{R}$) skew-symmetric operator on $\mathcal{X}(\pi_M)$ which satisfies a Leibniz rule, namely $[FX,Y]_\star = F\,[X,Y]_\star - (\mathrm{D}^\star_Y F)X$, but which need not have the Jacobi identity property. Any other bracket operator with these properties differs from the first one by a vector-valued 2-form (torsion form) along $\pi_M$. In other words, given $\mathrm{D}^\star$, the most general skew-symmetric bracket operator with the above Leibniz property is of the form

$$[X,Y]_\star = \mathrm{D}^\star_X Y - \mathrm{D}^\star_Y X + T^\star(X,Y), \tag{2.78}$$

where $T^\star$ is any element of $V^2(\pi_M)$ (the $C^\infty(J^1M)$-module of vector-valued 2-forms along $\pi_M$).

Let now $\omega$ be a scalar $k$-form along $\pi_M$ (notation: $\omega \in \bigwedge^k(\pi_M)$) or a vector-valued $k$-form (then $\omega \in V^k(\pi_M)$).

**Proposition 2.19.** *The operator $d^\star$, defined by*

$$\begin{aligned} d^\star\omega(X_0,\dots,X_k) &= \sum_{i=0}^{k}(-1)^i \mathrm{D}^\star_{X_i}(\omega(X_0,\dots,\hat{X}_i,\dots,X_k)) \\ &\quad + \sum_{0\le i<j\le k}(-1)^{i+j}\omega([X_i,X_j]_\star, X_0,\dots,\hat{X}_i,\dots,\hat{X}_j,\dots,X_k), \end{aligned} \tag{2.79}$$

*is a derivation of degree 1 on $\bigwedge(\pi_M)$ and $V(\pi_M)$.*

Proof: From the defining relation, it follows that the action of $d^\star$ on functions $F$ on $J^1M$, 1-forms $\alpha$ and vector fields $X$ along $\pi_M$ is given by: $d^\star F(X) = \mathrm{D}^\star_X F$, $d^\star\alpha(X,Y) = \mathrm{D}^\star_X(\alpha(Y)) - \mathrm{D}^\star_Y(\alpha(X)) - \alpha([X,Y]_\star)$, $d^\star X(Y) = \mathrm{D}^\star_Y X$. It is easy to verify that this restricted action has the necessary properties for a derivation, i.e. we have $d^\star(FG) = Fd^\star G + Gd^\star F$, $d^\star(F\alpha) = d^\star F\wedge\alpha + Fd^\star\alpha$ and $d^\star(FX) = F\,d^\star X + d^\star F\wedge X$. It follows that there is a unique derivation $\check{d}^\star$ which coincides with $d^\star$ when restricted to functions, 1-forms and vector fields. Defining $\check{d}^\star_X = [i_X, \check{d}^\star]$ as usual, one can create another self-dual degree zero derivation $(\check{d}^\star_X)^*$ which is obtained from $\check{d}^\star_X|_{\bigwedge^1(\pi_M)}$ by imposing the duality rule

$$\langle(\check{d}^\star_X)^*Y,\alpha\rangle = \check{d}^\star_X(\langle Y,\alpha\rangle) - \langle Y, \check{d}^\star_X\alpha\rangle$$

$\forall X, Y \in \mathcal{X}(\pi_M)$ and $\alpha \in \bigwedge^1(\pi_M)$. It was proved in [54] (see Prop. 3.3) that $\check{d}^\star$ then has the property

$$\begin{aligned}\check{d}^\star\omega(X_0,\ldots,X_k) &= \sum_{i=0}^{k}(-1)^i\check{d}^\star_{X_i}(\omega(X_0,\ldots,\hat{X}_i,\ldots,X_k)) \\ &+ \sum_{0\le i<j\le k}(-1)^{i+j}\omega((\check{d}^\star_{X_i})^*(X_j),X_0,\ldots,\hat{X}_i,\ldots,\hat{X}_j,\ldots X_k).\end{aligned}$$

One easily computes, however, that $(\check{d}^\star_X)^*Y = [X,Y]_\star$. Comparison of the above result with the defining relation (2.79) then shows that $\check{d}^\star \equiv d^\star$ and the result follows. □

We come back to the actual situation now, where we have two explicitly defined covariant derivatives on $\pi_M^*\tau_M$ which are of Berwald type. We write the original one (2.35) as D and the newly introduced one (2.52) as $\hat{\mathrm{D}}$. We make a further notational convention which has the advantage of focussing entirely on operations which involve only tensorial objects along the projection: if $\xi \in \mathcal{X}(J^1M)$ is itself the horizontal or vertical lift of some vector field $Y \in \mathcal{X}(\pi_M)$, we shall write

$$\mathrm{D}^H_Y X \text{ for } \mathrm{D}_{Y^H}X, \text{ and likewise } \mathrm{D}^V_Y X \text{ for } \mathrm{D}_{Y^V}X. \tag{2.80}$$

Such notations will make it easier to relate the discussion to the calculus of derivations developed in [53, 54, 80]. Similar notations are used for $\hat{\mathrm{D}}$.

The difference between the two Berwald-type connections is given by $\mathrm{D}_\xi X - \hat{\mathrm{D}}_\xi X = \langle X, dt\rangle\overline{\xi}_V$. This translates into the following relations between horizontal and vertical derivatives:

$$\mathrm{D}^H_Y X = \hat{\mathrm{D}}^H_Y X, \tag{2.81}$$

$$\mathrm{D}^V_Y X = \hat{\mathrm{D}}^V_Y X + \langle X, dt\rangle\overline{Y}. \tag{2.82}$$

The idea is now to let the $^\star$ of the above general considerations play the role of $^H$ and $^V$. Clearly, if we make the same choice of torsion forms for the brackets coming from both connections and subsequently use (2.79) to construct exterior derivatives, we will obtain the same horizontal exterior derivatives $d^H$ and $\hat{d}^H$, but the vertical exterior derivatives will be different.

From the general classification results of self-dual derivations in [80], we know that the difference between $\mathrm{D}^V_Y$ and $\hat{\mathrm{D}}^V_Y$ is a so-called derivation of type $\mu_*$. Such a derivation is of algebraic type and consists of two parts.

For a derivation of degree $r$, for example, we write $\mu_Q = a_Q - i_Q$, where $Q$ is a type (1,1) tensor-valued $r$-form along $\pi_M$; $a_Q$ vanishes on scalar forms, while $i_Q$ vanishes on vector fields. Specifically now, we derive from (2.82) that

$$\mathrm{D}^V_Y = \hat{\mathrm{D}}^V_Y + \mu_Q, \qquad \text{with} \quad Q = dt \otimes \overline{Y}. \tag{2.83}$$

It follows that for the dual action on 1-forms

$$\mathrm{D}^V_Y \alpha = \hat{\mathrm{D}}^V_Y \alpha - i_Q \alpha = \hat{\mathrm{D}}^V_Y \alpha - \langle \overline{Y}, \alpha \rangle dt.$$

Since the vertical bracket in [80] had no torsion, we take $\hat{T}^V = 0$ as well. We then have

$$\begin{aligned} d^V \alpha(X,Y) &= \mathrm{D}^V_X \alpha(Y) - \mathrm{D}^V_Y \alpha(X) \\ &= \hat{\mathrm{D}}^V_X \alpha(Y) - \hat{\mathrm{D}}^V_Y \alpha(X) - \langle \overline{X}, \alpha \rangle \langle Y, dt \rangle + \langle \overline{Y}, \alpha \rangle \langle X, dt \rangle, \end{aligned}$$

so that

$$d^V \alpha = \hat{d}^V \alpha + i_{dt \wedge I} \alpha.$$

Similarly, for the action on vector fields we find

$$d^V X = \hat{d}^V X + a_{\overline{I} \otimes dt} X, \quad \text{where} \quad a_{\overline{I} \otimes dt} X = \langle X, dt \rangle \, \overline{I}. \tag{2.84}$$

In conclusion, the difference between the two vertical derivatives is expressed by

$$\hat{d}^V = d^V - i_{dt \wedge I} - a_{\overline{I} \otimes dt} = d^V_{\overline{I}} - a_{\overline{I} \otimes dt}. \tag{2.85}$$

It may come a bit as a surprise that $\hat{d}^V$ is not the same as $d^V_{\overline{I}}$. One can verify, however, that just like $d^V_{\overline{I}}$, $\hat{d}^V$ has the coboundary property $\hat{d}^{V\,2} = 0$. Indeed, on scalar forms this is obvious since the $a_*$-term then does not contribute. To complete the argument, since both terms on the right-hand side of (2.85) manifestly vanish on $\partial/\partial x^i$, it then suffices to check that $\hat{d}^{V\,2}\mathbf{T} = 0$. This follows easily from the fact that $d^V_{\overline{I}} \mathbf{T} = a_{\overline{I} \otimes dt} \mathbf{T} = \overline{I}$ and thus $\hat{d}^V \mathbf{T} = 0$. In coordinates, the action of the new $\hat{d}^V$ on the local basis of 1-forms and vector fields is given by

$$\hat{d}^V \theta^i = 0, \quad \hat{d}^V dt = 0, \quad \hat{d}^V \frac{\partial}{\partial x^i} = 0, \quad \hat{d}^V \mathbf{T} = 0.$$

It would perhaps be worthwhile to enter more deeply into the question of the effect of selecting $\hat{d}^V$ as the fundamental vertical exterior derivative on the classification theory of derivations along $\pi_M$. This, of course, is beyond the scope of this chapter. In a sense, one expects that the influence of such a change will be minor as long as one deals with forms acting on $\overline{\mathcal{X}}(\pi_M)$. We finish our discussion by deriving a couple of properties which express this expectation in more precise terms.

Let $\omega$ be an element of $\bigwedge^k(\pi_M)$. Applying the definition of derivations of type $i_*$ (cf. [53]), we find that

$$i_{dt\wedge I}\omega\,(\overline{X}_1,\ldots,\overline{X}_{k+1}) = \frac{1}{2!(k-1)!}\sum_{\sigma\in S_{k+1}}(\operatorname{sign}\sigma)\,\omega\left((dt\wedge I)(\overline{X}_{\sigma(1)},\overline{X}_{\sigma(2)}),\overline{X}_{\sigma(3)},\ldots,\overline{X}_{\sigma(k+1)}\right) = 0.$$

Since derivations of type $a_*$ act trivially on scalar forms, we can conclude from this that

$$\hat{d}^V\omega\,(\overline{X}_1,\ldots,\overline{X}_{k+1}) = d^V\omega\,(\overline{X}_1,\ldots,\overline{X}_{k+1}).$$

Secondly, for $L \in V^k(\pi_M)$ we have

$$a_{\overline{I}\otimes dt}L\,(\overline{X}_1,\ldots,\overline{X}_{k+1}) = \frac{1}{k!}\sum_{\sigma\in S_{k+1}}(\operatorname{sign}\sigma)\,\overline{X}_{\sigma(1)}\left(\langle L(\overline{X}_{\sigma(2)},\ldots,\overline{X}_{\sigma(k+1)}),dt\rangle\right),$$

from which it follows that if $L$ takes values in $\overline{\mathcal{X}}(\pi_M)$, the actions of $\hat{d}^V$ and $d^V$ coincide when the resulting forms are restricted to $\overline{\mathcal{X}}(\pi_M)$ again.

## Chapter 3

# *The use of Berwald-type connections in geometric mechanics*

### 3.1 The Jacobi endomorphism and the mixed curvature

So far, we did not mention the areas of application of our Berwald-type connections in mechanics. In this chapter, we will say a few words about the contexts in which Berwald-type connections actually are important intrinsic tools.

In particular the vertical and horizontal covariant derivative operators (2.80) which come with the optimal Berwald-type connection (2.52) associated to a SODE $\Gamma$ play an important role in the characterisation of a variety of qualitative features of that SODE. A third key object in the theory of SODEs is the so-called Jacobi endomorphism which we have already encountered in Section 2.7. In our present approach, it can also be viewed as the $C^\infty(J^1M)$-linear map $\Phi : \overline{\mathcal{X}}(\pi_M) \to \overline{\mathcal{X}}(\pi_M)$ given by

$$\Phi = -\mathcal{R}_{\mathbf{T}},$$

where $\mathcal{R}_{\mathbf{T}} = -[\Gamma, \overline{X}^H]_V$ is a torsion component of the optimal Berwald-type connection associated to the SODE $\Gamma$ (see the second table in Section 2.6). It is further easy to extend $\Phi$ in such a way that it can be applied to all vector fields along $\pi_M$, simply by putting $\Phi(\mathbf{T}) = 0$. Very often, it is appropriate to choose a local basis of vector fields along $\pi_M$, which is adapted to the structure of $\Phi$ (for example a basis of eigenvector fields of $\Phi$). The covariant derivative operators then allow to replace analytical computations by intrinsic, geometrical ones.

$\Phi$ completely determines the curvature of the SODE connection (2.15), as well as the torsion of the associated optimal Berwald-type connection. Indeed, the only torsion term left to be determined is $\mathcal{R}$ (see again the second table of Section 2.6). It is clearly related to expression (2.19) of the curvature of the SODE connection and one can compute that

$$-3\mathcal{R}(\overline{X},\overline{Y}) = (\mathrm{D}_{\overline{X}^V}\Phi)(\overline{Y}) - (\mathrm{D}_{\overline{Y}^V}\Phi)(\overline{X}) \tag{3.1}$$

(see [80, 23]). Remark that, in the right-hand side of the above expression, it doesn't really matter which linear connection D we choose, as long as it is contained in the class of Berwald-type connections. Given a SODE $\Gamma$, once we have obtained $\Phi$ and $\mathcal{R}$, the curvature

$$curv(\xi,\eta)X = \mathrm{D}_\xi \mathrm{D}_\eta X - \mathrm{D}_\eta \mathrm{D}_\xi X - \mathrm{D}_{[\xi,\eta]}X.$$

of a Berwald-type connection D is almost completely determined. One can compute (see again [23]), for any linear connection D in the class of Berwald-type connections (associated to $\Gamma$), the following curvature components:

$$\begin{array}{ll} curv(\overline{X}^V,\overline{Y}^V)\overline{Z} = 0, & curv(\Gamma,\overline{Y}^V)\overline{Z} = 0, \\ curv(\overline{X}^H,\overline{Y}^H)\overline{Z} = Rie(\overline{X},\overline{Y})\overline{Z}, & curv(\Gamma,\overline{Y}^H)\overline{Z} = -(\mathrm{D}_{\overline{Z}^V}\Phi)(\overline{Y}) + \mathcal{R}(\overline{Y},\overline{Z}), \end{array}$$

where $Rie(\overline{X},\overline{Y})\overline{Z} = \mathrm{D}_{\overline{Z}^V}\mathcal{R}(\overline{X},\overline{Y})$ and $\mathcal{R}$ is as in (3.1). Evidently for the optimal Berwald-type connection (2.52), we find that $curv(\xi,\eta)\mathbf{T} = 0$, for all $\xi,\eta \in \mathcal{X}(J^1M)$, while for the linear connection (2.35) of [23] one finds

$$\begin{array}{ll} curv(\overline{X}^V,\overline{Y}^V)\mathbf{T} = 0, & curv(\Gamma,\overline{Y}^V)\mathbf{T} = 0, \\ curv(\overline{X}^H,\overline{Y}^H)\mathbf{T} = \mathcal{R}(\overline{X},\overline{Y}), & curv(\Gamma,\overline{Y}^H)\mathbf{T} = -\Phi(\overline{Y}). \end{array}$$

We are left to determine only two more curvature components. First, $curv(\overline{X}^V,\overline{Y}^V)\mathbf{T} = 0$, for both the optimal Berwald-type connection and the linear connection of [23]. Finally, the last curvature component is the 'mixed curvature'. It is, unlike all the other non-trivial ones, not related to $\Phi$. Therefore it defines a fourth object of interest, which has been denoted by $\theta$ in [23]:

$$\theta(\overline{X},\overline{Y})\overline{Z} = curv(\overline{X}^V,\overline{Y}^H)\overline{Z} = \mathrm{D}_{\overline{X}^V}\mathrm{D}_{\overline{Y}^H}\overline{Z} - \mathrm{D}_{\overline{Y}^H}\mathrm{D}_{\overline{X}^V}\overline{Z} - \mathrm{D}_{[\overline{X}^V,\overline{Y}^H]}\overline{Z}.$$

Remark again that, in the expression for $\theta$, we only made use of the restriction of the Berwald-type connection to vector fields in $\overline{\mathcal{X}}(\pi_M)$, and thus $\theta$ will be the same for all Berwald-type connections.

We now briefly sketch three applications in which the covariant derivative operators $\mathrm{D}^V$ and $\mathrm{D}^H$, the Jacobi endomorphism $\Phi$ and the mixed curvature $\theta$ play a significant role.

## 3.2 The inverse problem of the calculus of variations

A first application is the so-called inverse problem of the calculus of variations. This concerns the question whether, for a given SODE $\ddot{x}^i - f^i = 0$, a symmetric type (0,2) tensor field $g$ along $\pi_M$ exists (having the property $g(\mathbf{T}, .) = 0$ and being non-singular when restricted to $\overline{\mathcal{X}}(\pi_M)$) such that the equivalent system $g_{ij}(\ddot{x}^j - f^j) = 0$ is a set of Euler-Lagrange equations.

A geometrical study of this problem, making use of operators that are associated to Berwald-type connections (associated to the given SODE), has been initiated in [27]. One can start the discussion with an arbitrary D in the class of Berwald-type connections (i.e. a linear connection satisfying (2.34)). The vertical covariant derivative, as defined in (2.80) can be extended to a (self)-dual degree 0 derivation on tensor fields along $\pi_M$, by putting

$$\mathrm{D}^V_X F = X^V F, \quad \mathrm{D}^V_X Y = \mathrm{D}_{X^V} Y, \quad (\mathrm{D}^V_X \alpha)(Y) = \mathrm{D}^V_X\big(\alpha(Y)\big) - \alpha(\mathrm{D}^V_X Y)$$

($X, Y \in \mathcal{X}(\pi_M)$, $\alpha \in \bigwedge^1(\pi_M)$) and subsequently, for a general $(k,l)$-tensor field $A$ along $\pi_M$

$$\begin{aligned}(\mathrm{D}^V_X A)(\alpha_1, \ldots, \alpha_k, X_1, \ldots, X_l) \;=\;& X^V\big(A(\alpha_1, \ldots, \alpha_k, X_1, \ldots, X_l)\big)\\ &-\sum_{i=1}^{k} A(\alpha_1, \ldots, \mathrm{D}^V_X\alpha_i, \ldots, \alpha_k, X_1, \ldots, X_l)\\ &-\sum_{j=1}^{l} A(\alpha_1, \ldots, \alpha_k, X_1, \ldots, \mathrm{D}^V_X X_j, \ldots, X_l).\end{aligned}$$

It is then possible to define a 'covariant differential': this is an operator $\mathrm{D}^V$ that maps a $(k,l)$ tensor field $A$ along $\pi_M$ onto the $(k, l+1)$ tensor field $\mathrm{D}^V A$ along $\pi_M$ in such a way that

$$(\mathrm{D}^V A)(\alpha_1, \ldots, \alpha_k, X, X_1, \ldots, X_l) := (\mathrm{D}^V_X A)(\alpha_1, \ldots, \alpha_k, X_1, \ldots, X_l).$$

The conditions for the existence of a metric tensor $g$ with the above mentioned properties are known as the *Helmholtz conditions*. They can be expressed in the following coordinate free way:

$$\begin{cases} \mathrm{D}_\Gamma g(\overline{X}, \overline{Y}) = 0, \\ \mathrm{D}^V_{\overline{X}} g(\overline{Y}, \overline{Z}) = \mathrm{D}^V_{\overline{Y}} g(\overline{X}, \overline{Z}) \qquad \text{and} \\ g(\Phi(\overline{X}), \overline{Y}) = g(\Phi(\overline{Y}), \overline{X}). \end{cases} \tag{3.2}$$

Here, $\overline{X}$, $\overline{Y}$ and $\overline{Z}$ are arbitrary vector fields in $\overline{\mathcal{X}}(\pi_M)$ and D is a connection of Berwald-type, associated to the SODE $\Gamma$. It was shown in [75] that the calculus originating from Berwald-type connections makes it possible to give an extensive geometrical treatment of the integrability conditions of these equations for $g$. Such integrability conditions involve, among others also, the horizontal covariant derivatives of $g$

$$\mathrm{D}^H_{\overline{X}} g(\overline{Y}, \overline{Z}) = \mathrm{D}^H_{\overline{Y}} g(\overline{X}, \overline{Z}) \tag{3.3}$$

and conditions in which the mixed curvature $\theta$ makes its appearance. Further developments were made in [25, 79].

We will now show some small advantages when the linear connection D under consideration is chosen to be the optimal Berwald-type connection $\hat{\mathrm{D}}$, defined by expression (2.52). The Helmholtz conditions above are then equivalent with

$$\begin{cases} \hat{\mathrm{D}}_\Gamma g = 0, \\ \hat{\mathrm{D}}^V_X g(Y, Z) = \hat{\mathrm{D}}^V_Y g(X, Z) \qquad \text{and} \\ g(\Phi(X), Y) = g(\Phi(Y), X). \end{cases} \tag{3.4}$$

for *all* vector fields $X, Y, Z \in \mathcal{X}(\pi_M)$. We will check e.g. that for the second relation, for all $\overline{X}, \overline{Z} \in \overline{\mathcal{X}}(\pi_M)$,

$$\hat{\mathrm{D}}^V_{\overline{X}} g(\mathbf{T}, \overline{Z}) = \overline{X}^V \big(g(\mathbf{T}, \overline{Z})\big) - g(\hat{\mathrm{D}}^V_{\overline{X}}\mathbf{T}, \overline{Z}) - g(\mathbf{T}, \hat{\mathrm{D}}^V_{\overline{X}}\overline{Z}) = 0.$$

In this calculation, next to the requirement $g(\mathbf{T}, .) = 0$, the fact that $\hat{\mathrm{D}}^V_X \mathbf{T}$ vanishes is essential. Obviously also $\hat{\mathrm{D}}^V_{\mathbf{T}} g(\overline{X}, \overline{Z}) = 0$ since $\mathbf{T}^V = 0$. Of course, the use of the optimal Berwald-type connection in this context has only an 'aesthetical' advantage and will not lead to any new developments in the field.

Let us suppose now that $\Gamma$ is indeed an Euler-Lagrange vector field, w.r.t. a regular Lagrangian $L$ and that we want to characterise the corresponding metric $g_L$. Already in Section 2.4 (expression (2.22)) we have shown that a coordinate-free expression for $g_L$ can be found by means of the Poincaré-Cartan two-form $\omega_L$. However, there exists also a fairly simple definition for $g_L$ that makes use of a Berwald-type connection D. It is easy to see that

$$\begin{aligned} \mathrm{D}^V \mathrm{D}^V L(X, Y) &= \mathrm{D}^V_X(\mathrm{D}^V L)(Y) = \mathrm{D}^V_X \mathrm{D}^V_Y L - \mathrm{D}^V_{\mathrm{D}^V_X Y} L \\ &= X^V Y^V L - (\mathrm{D}^V_X Y)^V L. \end{aligned} \tag{3.5}$$

If we take *basic* vector fields $\overline{X}$ and $\overline{Y}$ in $\overline{\mathcal{X}}(\pi_M)$, then for any Berwald-type connection $\mathrm{D}^V_{\overline{X}}\overline{Y} = 0$, and we find

$$\mathrm{D}^V\mathrm{D}^V L(\overline{X},\overline{Y}) = \overline{X}^V\overline{Y}^V L,$$

which, in coordinates, clearly gives the Hessian of $L$ and is symmetric under an interchange of $\overline{X}$ and $\overline{Y}$. Let us continue by putting $\mathbf{T}$ and a basic $\overline{Y}$ in the arguments of (3.5), then

$$\mathrm{D}^V\mathrm{D}^V L(\mathbf{T},\overline{Y}) = 0, \qquad \text{and} \qquad \mathrm{D}^V\mathrm{D}^V L(\overline{Y},\mathbf{T}) = -(\mathrm{D}^V_{\overline{X}}\mathbf{T})^V L.$$

Apparently, only for the optimal Berwald-type connection $\hat{\mathrm{D}}$, we can conclude that the tensor field defined by $g_L = \hat{\mathrm{D}}^V\hat{\mathrm{D}}^V L$ is symmetric for *all* vector fields in $\mathcal{X}(\pi_M)$, and that it satisfies the property $g(\mathbf{T},.) = 0$.

Let us finally come back now to the Cartan-type connection $\overset{\mathrm{C}}{\mathrm{D}}$ (2.59) for the metric $g_L = \hat{\mathrm{D}}^V\hat{\mathrm{D}}^V L$. The second and third term on the right-hand side of the defining relations (2.55,2.56) for $\mathcal{C}_V$ and $\mathcal{C}_H$ cancel each other in view of the Helmholtz conditions (3.2) and the derived integrability condition (3.3) satisfied by the tensor field $g_L$. Moreover, since in such a case also $\hat{\mathrm{D}}_\Gamma g = 0$, we will have $\mathcal{C}_H(\mathbf{T},\,.\,) = 0$. The effect of this last property is that all four connections (Berwald, Cartan, Hashiguchi and Chern–Rund) then share the same 'dynamical covariant derivative operator' $\hat{\mathrm{D}}_\Gamma$. This feature was emphasised (for autonomous systems) in Crampin's discussion of the second variation formula [22], because the dynamical covariant derivative and the Jacobi endomorphism is all one needs in such an analysis. Note, however, that if we are not in the Lagrangian case, the Berwald $\hat{\mathrm{D}}_\Gamma$ and the Cartan $\overset{\mathrm{C}}{\mathrm{D}}_\Gamma$ may be different; in fact, the necessary and sufficient condition for them to be identical is that $\hat{\mathrm{D}}_{\mathbf{T}^H} g = 0$.

## 3.3 Linearisable and separable second-order differential equations

In this section we very briefly mention two results from SODE-theory where one makes use of Berwald type connections. A SODE is said to be *linearisable* if, by an appropriate coordinate change, it can be cast in the form

$$\begin{cases} \dot{x}^i = v^i, \\ \dot{v}^i = A^i_j(t)v^j + B^i_j(t)x^j + a^i(t). \end{cases}$$

It turns out (see [23]) that necessary and sufficient conditions for the linearisability of a given SODE are that

(i) $\theta = 0$,

(ii) $\mathrm{D}_{\overline{X}^H}\Phi(\overline{Y}) = \mathrm{D}_{\overline{X}^V}\Phi(\overline{Y}) = 0$ for all $\overline{X}, \overline{Y} \in \overline{\mathcal{X}}(\pi_M)$.

Again, only the restriction of D to vector fields in $\overline{\mathcal{X}}(\pi_M)$ was required. Clearly the advantage of the approach using Berwald-type connections is that the conditions for linearisability are put in an intrinsic form and can thus be tested on the given data in any coordinates (and for any connection in the class of Berwald-type connections).

A much more difficult problem to characterise is *separability.* A SODE is said to be separable if, locally, the associated system of differential equations admits a full decoupling into separate one-dimensional second-order equations, i.e. if, by an appropriate coordinate change, it can be cast in the following form

$$\begin{cases} \dot{x}^i = v^i, \\ \dot{v}^i = f^i(t, x^i, v^i), \end{cases}$$

where the functions $f^i$ depend on $t$ and on the corresponding coordinates $x^i$ and $v^i$ only.

Before restating a theorem from [12], we need to explain a few definitions. A (1,1) tensor field $U$ along $\pi_M$ is said to be *diagonalisable* if:

(i) for each $v \in J^1_m M$, the linear map $U(v)|_{V_m M} : V_m M \to V_m M$ is diagonalisable (in the sense that the real Jordan normal form of $(U^i_j(m))$ is diagonal);

(ii) there (locally) exist smooth functions $\mu_\alpha$ such that $\mu_\alpha(v)$ is an eigenvalue of $U(v)|_{V_m M}$;

(iii) the rank of $\mu_\alpha \overline{I} - U$ is constant.

If $U$ is a (1,1) tensor field along $\pi_M$ of type $U_4$ in expression (2.47) (meaning that $U$ should be locally of the form $U^i_j \theta^j \otimes \frac{\partial}{\partial x^i}$ ), then we will use the notation $A^V_U$ for the object defined by

$$\begin{aligned} A^V_U(\overline{X}, \overline{Y}) &= [\mathrm{D}^V_{\overline{X}} U, U](\overline{Y}) && \text{for all } \overline{X}, \overline{Y} \in \overline{\mathcal{X}}(\pi_M) \text{ and} \\ A^V_U(Z, \mathbf{T}) &= A^V_U(\mathbf{T}, Z) = 0 && \text{for all } Z \in \mathcal{X}(\pi_M) \end{aligned}$$

(here the bracket $[\cdot,\cdot]$ stands for the commutator). We now need only one more definition: to each vector field $X$ with $\langle X, dt\rangle = 1$ on $M$, we associate a (1,1)-tensor field $\mathbf{t}_X$ in the following way. Look at the decomposition (2.3) of the natural prolongation of $X$ to a vector field $X^{(1)}$ on $J^1M$

$$X^{(1)} = X_1^H + \overline{X}_2^V.$$

Then $X_1 = X \circ \pi$ and $\overline{X}_2$ is a vector field in $\overline{\mathcal{X}}(\pi_M)$. We put

$$\mathbf{t}_X = d^V \overline{X}_2.$$

Again we should remark that the action of the derivation $d^V$, associated to a Berwald-type connection, on elements of $\overline{\mathcal{X}}(\pi_M)$ is the same for all connections of the class (see e.g. expression (2.84)).

Finally, we can formulate the main result of [12]. It has been proved in [12] that a SODE $\Gamma$ is separable if and only if the following conditions hold

(i) $\Phi$ is diagonalisable,

(ii) $A^V_\Phi = 0$,

(iii) $[\mathrm{D}_\Gamma\Phi, \Phi] = 0$,

(iv) $\mathcal{R} = 0$,

(v) for each degenerate eigenvalue $\mu_\alpha$ of $\Phi$ there exists a submanifold $\mathcal{U}_\alpha$ of $M$, fibred over (a connected subset of) $I\!R$, and a vector field $X$ on $\mathcal{U}_\alpha$ with $\langle X, dt\rangle = 1$, such that the corresponding $\mathbf{t}_X$ is

  (a) diagonalisable, with all eigenvalues nondegenerate,
  (b) satisfies $A^V_{\mathbf{t}_X} = 0$,
  (c) satisfies $[D_\Gamma \mathbf{t}_X, \mathbf{t}_X] = 0$.

# Chapter 4

# *Lagrangian mechanics on Lie algebroids*

## 4.1 Weinstein's Lagrangian systems on Lie algebroids

In the last years, there has been a lot of interest in the study of dynamical systems which have a Lie algebroid as carrying space (see e.g. [13, 14, 16, 45, 48, 49, 50, 72, 91]). The potential relevance of Lie algebroids for applications in physics and other fields of applied mathematics has gradually become more evident. In particular, a contribution by A. Weinstein [91] has revealed the role Lie algebroids play in modeling certain problems in mechanics. The topic of that paper we want to comment upon here, is that of dynamical systems which were called 'Lagrangian systems' on Lie algebroids. This concept certainly defines an interesting generalisation of Lagrangian systems as known from classical mechanics, if only because of the more general class of differential equations it involves while preserving a great deal of the very rich geometrical structure of Lagrangian (and Hamiltonian) mechanics. First, we will explain briefly the basic set-up of Weinstein's paper.

**Definition 4.1.** *A Lie algebroid is a vector bundle $\tau : \mathsf{V} \to M$, which comes equipped with*

- *a bracket operation on the set of sections of $\tau$, $[\cdot,\cdot] : \mathrm{Sec}(\tau) \times \mathrm{Sec}(\tau) \to \mathrm{Sec}(\tau)$,*
- *a linear bundle map $\varrho : \mathsf{V} \to TM$ (and its extension $\varrho : \mathrm{Sec}(\tau) \to \mathcal{X}(M)$), called the anchor map,*

*which are related in such a way that*

1. *$[\cdot,\cdot]$ is the bracket of a real Lie algebra structure.*

2. *there is a certain compatibility between bracket and the anchor map. To be precise, we have*

$$[\mathsf{s}, f\mathsf{r}] = f[\mathsf{s}, \mathsf{r}] + \varrho(\mathsf{s})(f)\,\mathsf{r}, \tag{4.1}$$

*for all* $f \in C^\infty(M)$.

*As a consequence, the anchor map will establish a Lie algebra homomorphism between* $\mathrm{Sec}(\tau)$ *and the real Lie algebra of vector fields on* $M$, *i.e.*

$$[\varrho(\mathsf{s}), \varrho(\mathsf{r})] = \varrho([\mathsf{s}, \mathsf{r}]) \tag{4.2}$$

*for all* $\mathsf{s}, \mathsf{r} \in \mathrm{Sec}(\tau)$.

The property (4.2) follows from the Jacobi identity of the Lie algebroid bracket and property (4.1). A standard reference for the theory of Lie algebroids (and their relation to groupoids) is the book by K. Mackenzie [46], while a more recent work is for example [10]. We will only mention here that, for any Lie algebroid, it is possible to develop an exterior calculus of forms. Here, 'forms' are made up of exterior powers of sections of the dual bundle $\tau^* : \mathsf{V}^* \to M$. An exterior derivative is defined, for example in Ch. IV, p. 198 in [46]. The defining properties of the Lie algebroid then exactly translate in the fact that the exterior derivative satisfies the coboundary condition $d^2 = 0$.

Let us denote by $x^I$ the coordinates on $M$ and by $\mathsf{v}^a$ fibre coordinates on $\mathsf{V}$. Then locally the anchor map $\varrho : \mathsf{V} \to TM$ is of the form

$$(x^I, \mathsf{v}^a) \mapsto \varrho^I_a(x)\mathsf{v}^a \frac{\partial}{\partial x^I}.$$

Suppose that $\{\mathsf{e}_a\}$ is a basis for $\mathrm{Sec}(\tau)$. Then, the algebroid bracket is uniquely determined by its action on the base sections:

$$[\mathsf{e}_a, \mathsf{e}_b] = \mathsf{C}^{\,c}_{ab}(x)\mathsf{e}_c. \tag{4.3}$$

The functions $\mathsf{C}^{\,a}_{bc} \in C^\infty(M)$ are called the *structure functions of the Lie algebroid*. They are skew-symmetric in their lower indices and related to the anchor map by means of the following compatibility conditions:

$$\varrho^I_a \frac{\partial \varrho^J_b}{\partial x^I} - \varrho^I_b \frac{\partial \varrho^J_a}{\partial x^I} = \mathsf{C}^{\,c}_{ab}\varrho^J_c, \tag{4.4}$$

$$\sum_{a,b,c} \left( \varrho^I_a \frac{\partial \mathsf{C}^{\,d}_{bc}}{\partial x^I} + \mathsf{C}^{\,d}_{ae}\mathsf{C}^{\,e}_{bc} \right) = 0. \tag{4.5}$$

The first expression merely states that $\varrho$ is a Lie algebra homomorphism, the second is a coordinate version of the Jacobi identity. Next to Lie algebroids, we will also discuss a second structure.

**Definition 4.2.** *A Poisson structure on a manifold $M$ is a real Lie algebra bracket $\{\cdot,\cdot\}$ on the ring of smooth functions $C^\infty(M)$, which further satisfies the Leibniz identity:*

$$\{f, gh\} = g\{f,h\} + \{f,g\}h.$$

An important theorem by T. Courant [17] establishes the relation between Lie algebroids and linear Poisson brackets on the dual bundle $\mathsf{V}^*$.

**Proposition 4.3.** *A vector bundle $\tau : \mathsf{V} \to M$ is a Lie algebroid if and only if there exists a Poisson structure on the total manifold $\mathsf{V}^*$ of the dual bundle $\tau^* : \mathsf{V}^* \to M$ with respect to which the fibre linear functions on $\mathsf{V}^*$ form a Lie subalgebra.*

Here, we will only recall the construction of this Poisson bracket from the Lie algebroid bracket. In the case under consideration, due to the Leibniz property, it will be enough to specify the Poisson bracket on linear and constant functions. Due to the identification $\mathsf{V}_m \simeq (\mathsf{V}_m^*)^*$, it is easy to see that a section $\mathsf{s}$ of the vector bundle $\tau$ can be regarded as a linear function $\hat{\mathsf{s}}$ on the dual bundle $\tau^*$, and vice versa. Starting from a Lie algebroid structure on $\tau$, the Poisson structure on $\tau^*$ can be defined by

$$\{f,g\} = 0, \qquad \{\hat{\mathsf{s}}, \hat{\mathsf{r}}\} = -\widehat{[\mathsf{s},\mathsf{r}]} \qquad \text{and} \qquad \{f,\mathsf{s}\} = \varrho(\mathsf{s})f. \tag{4.6}$$

Here $f, g \in C^\infty(M)$ should be viewed as functions on $\mathsf{V}^*$, constant on the fibres. We have adopted a different sign convention then in [17]. The Leibniz property of the bracket $\{\cdot,\cdot\}$ follows, for example, from (4.1). Evidently, the Jacobi identity of $\{\cdot,\cdot\}$ for linear functions follows directly from the one of $[\cdot,\cdot]$. The Jacobi identity for a mixture of linear and constant functions is satisfied due to the fact that $\varrho$ is supposed to be a Lie algebra homomorphism. Using the structure functions and the anchor map as above, we find

$$\{x^I, x^J\} = 0, \qquad \{\mu_a, \mu_b\} = \mathsf{C}_{ab}^{\;\;c}\mu_c \qquad \text{and} \qquad \{x^I, \mu_a\} = \varrho_a^I,$$

where $\mu_a = \hat{\mathsf{e}}_a$ is the linear function on $\mathsf{V}^*$ associated to $\mathsf{e}_a$.

Weinstein makes use of the above correspondence between Lie algebroids and Poisson structures to define Lagrangian systems. Suppose that $L(x,\mathsf{v})$

is a smooth function on $\mathsf{V}$. The *Legendre map* is the fibre derivative of $L$, $\mathcal{F}_L : \mathsf{V} \to \mathsf{V}^*, (x^I, \mathsf{v}^a) \mapsto (x^I, \frac{\partial L}{\partial \mathsf{v}^a}(x, \mathsf{v}))$. $L$ is said to be a *regular Lagrangian* if $\mathcal{F}_L$ is a local diffeomorphism, i.e. if the Hessian $\frac{\partial^2 L}{\partial \mathsf{v}^a \partial \mathsf{v}^b}$ has maximal rank at any point. For any regular Lagrangian, it is possible to pull-back the Poisson structure on $\mathsf{V}^*$ (coming from the Lie algebroid on $\tau$ by means of (4.6)) to a Poisson structure on $\mathsf{V}$. This bracket can be characterised by the relations

$$\{x^I, x^J\} = 0, \quad \{\frac{\partial L}{\partial \mathsf{v}^a}, \frac{\partial L}{\partial \mathsf{v}^b}\} = -\mathsf{C}_{ab}^{\ c} \frac{\partial L}{\partial \mathsf{v}^c} \quad \text{and} \quad \{x^I, \frac{\partial L}{\partial \mathsf{v}^a}\} = \varrho_a^I. \tag{4.7}$$

The *energy* $E$ of a Lagrangian $L$ on $\mathsf{V}$ is defined by $E(\mathsf{v}) = \langle \mathcal{F}_L(\mathsf{v}), \mathsf{v} \rangle - L(\mathsf{v})$, or

$$E(\mathsf{v}) = \mathsf{v}^a \frac{\partial L}{\partial \mathsf{v}^a} - L.$$

We can now compute the Hamilton equations for the energy $E$ w.r.t. the Poisson bracket (4.7). The first equation is

$$\begin{aligned} \frac{dx^I}{dt} &= \{x^I, E\} = \{x^I, \mathsf{v}^a\} \frac{\partial L}{\partial \mathsf{v}^a} + \mathsf{v}^a \{x^I, \frac{\partial L}{\partial \mathsf{v}^a}\} - \{x^I, L\} \\ &= \mathsf{v}^a \{x^I, \frac{\partial L}{\partial \mathsf{v}^a}\} = \mathsf{v}^a \varrho_a^I. \end{aligned}$$

Secondly, instead of calculating $\frac{d\mathsf{v}^a}{dt} = \{\mathsf{v}^a, E\}$, one usually computes directly the time derivative of $\frac{\partial L}{\partial \mathsf{v}^a}$:

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L}{\partial \mathsf{v}^a}\right) &= \{\frac{\partial L}{\partial \mathsf{v}^a}, E\} = \{\frac{\partial L}{\partial \mathsf{v}^a}, \mathsf{v}^b\} \frac{\partial L}{\partial \mathsf{v}^b} + \mathsf{v}^b \{\frac{\partial L}{\partial \mathsf{v}^a}, \frac{\partial L}{\partial \mathsf{v}^b}\} - \{\frac{\partial L}{\partial \mathsf{v}^a}, L\} \\ &= \mathsf{v}^b \{\frac{\partial L}{\partial \mathsf{v}^a}, \frac{\partial L}{\partial \mathsf{v}^b}\} - \{\frac{\partial L}{\partial \mathsf{v}^a}, x^I\} \frac{\partial L}{\partial x^I} = -\mathsf{v}^b \mathsf{C}_{ab}^{\ c} \frac{\partial L}{\partial \mathsf{v}^c} + \varrho_a^I \frac{\partial L}{\partial x^I}. \end{aligned}$$

To conclude, the local coordinate expression of Weinstein's 'Lagrangian systems' on a Lie algebroid reads:

$$\begin{aligned} \dot{x}^I &= \varrho_a^I(x) \mathsf{v}^a, \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \mathsf{v}^a}\right) &= \varrho_a^I \frac{\partial L}{\partial x^I} - \mathsf{C}_{ab}^{\ c} \mathsf{v}^b \frac{\partial L}{\partial \mathsf{v}^c}, \end{aligned} \tag{4.8}$$

where the functions $\varrho_a^I$ and $\mathsf{C}_{ab}^{\ c}$ satisfy (4.4). Applications for such model equations can be found e.g. in the theory of systems with symmetries on principal fibre bundles and in rigid body dynamics. Note that, more generally, equations of the form

$$\begin{aligned} \dot{x}^I &= \varrho^I_a(x)\mathsf{v}^a, \\ \dot{\mathsf{v}}^a &= f^a(x,\mathsf{v}), \end{aligned} \tag{4.9}$$

were called "second-order equations on a Lie algebroid" by Weinstein. They are indeed, to some extent, the analogues of second-order dynamics on a tangent bundle. It is clear, however, that these equations truly are second-order differential equations only when the base manifold and the fibres have the same dimension and $\varrho$ is injective. We will therefore rather call them 'pseudo-second-order ordinary differential equations', pseudo-SODEs for short. Weinstein also raised the question whether there would be a geometrical way of defining equations of the form (4.8), much in the line of the geometrical construction of classical Lagrange equations, which makes use of the intrinsic structures living on a tangent bundle. Martínez has resolved this issue [48] by showing that one can prolong a Lie algebroid in such a way that the newly obtained space has all the features of tangent bundle geometry, which are important for the geometrical construction of Lagrangian systems. That is to say, the prolonged Lie algebroid carries a Liouville-type section and a vertical endomorphism which enables the definition of a Poincaré-Cartan type 1-form, associated to a function $L$; the available exterior derivative then does the rest for arriving at an analogue of the symplectic structure from which Lagrangian equations can be derived.

## 4.2 Time-dependent version of Lagrangian systems on Lie algebroids

In this dissertation, we first wish to set the stage for an appropriate generalisation of this theory to non-autonomous systems of differential equations. Let us suppose first that the base manifold $M$ is fibred over $\mathbb{R}$. Under a time-dependent coordinate change $\hat{x}^i = \hat{x}^i(t,x)$, the induced transformations for jet coordinates are

$$\dot{\hat{x}}^i = \frac{\partial \hat{x}^i}{\partial x^j}(t,x)\dot{x}^j + \frac{\partial \hat{x}^i}{\partial t}(t,x).$$

Therefore, for example at the level of pseudo-second-order equations, if we wish the structure of the equations to be invariant under transformations $\hat{x}^i = x^i(t,x)$, the right generalisation of (4.9) to the tin-dependent setting will not be just a matter of allowing the functions $\varrho^i_\alpha$ and $f^\alpha$ in (4.9) to depend on time. We believe that the right generalisation rather should produce equations of the form:

$$\begin{aligned} \dot{x}^i &= \lambda^i_\alpha(t,x)y^\alpha + \lambda^i(t,x) \\ \dot{y}^\alpha &= f^\alpha(t,x,y). \end{aligned} \tag{4.10}$$

Of course, one can also transform the coordinates $y^\alpha$. It is easy to see that the structure of the first equation remains preserved when affine coordinate changes

$$y'^\alpha = A^\alpha_\beta(t,x)y^\beta + B^\alpha(t,x)$$

are allowed. To conclude, we believe that our generalisation should be based on an *affine bundle* $\pi : E \to M$ which is 'anchored' in the first jet bundle by means of an *affine bundle map* $\lambda : E \to J^1M$.

As for Lagrange-type equations, our only concern at the moment is to have an idea of what a time-dependent generalisation of (4.8) should look like. Now, there is a way of developing a kind of formal calculus of variations approach which leads to equations of the form (4.8), and in which the first set of equations are treated as constraints. We will show in the next section of this chapter that if such an approach is adopted when the Lagrangian is allowed to depend on time and the constraints are as in the first equation of (4.10), one obtains equations of the form

$$\begin{aligned} \dot{x}^i &= \lambda^i_\alpha(t,x)y^\alpha + \lambda^i(t,x), \\ \frac{d}{dt}\left(\frac{\partial L}{\partial y^\alpha}\right) &= \lambda^i_\alpha \frac{\partial L}{\partial x^i} - (C^\gamma_{\alpha\beta}y^\beta - C^\gamma_{0\alpha})\frac{\partial L}{\partial y^\gamma} \end{aligned} \tag{4.11}$$

where the functions $\lambda^i_\alpha$, $\lambda^i$, $C^\gamma_{\alpha\beta}$, $C^\gamma_{0\alpha}$ satisfy the relations

$$\lambda^i_\alpha \frac{\partial \lambda^j_\beta}{\partial x^i} - \lambda^i_\beta \frac{\partial \lambda^j_\alpha}{\partial x^i} = \lambda^j_\gamma C^\gamma_{\alpha\beta}, \tag{4.12}$$

$$\frac{\partial \lambda^j_\beta}{\partial t} + \lambda^i \frac{\partial \lambda^j_\beta}{\partial x^i} - \lambda^i_\beta \frac{\partial \lambda^j}{\partial x^i} = \lambda^j_\alpha C^\alpha_{0\beta}. \tag{4.13}$$

Thus, in the next chapters, we want to address the question of explaining the nature of the conditions (4.12) and (4.13), which presumably should again have something to do with a Lie algebroid structure. Inspired by these analytical considerations, we will introduce the notion of a Lie algebroid structure on an affine bundle $\pi : E \to M \to \mathbb{R}$.

Note however, that one can extend the theory to more general affine bundles $\pi : E \to M$, whose base manifold need not be fibred over $I\!R$. In fact, the final step in our programme will be to introduce the notion of a Lie algebroid structure on *any* affine bundle, that is to say: without any reference to 'time' as a special coordinate. Therefore, in a slightly more general set-up, the affine bundles will be 'anchor mapped' into tangent bundles, rather than first jet bundles, i.e. the anchor map will be an affine bundle map $\rho : E \to TM$. For example, at the level of pseudo-SODEs, the dynamical systems under investigation will be of the form

$$\begin{aligned} \dot{x}^I &= \rho^I_\alpha(x) y^\alpha + \rho^I_0(x) \\ \dot{y}^\alpha &= f^\alpha(x,y). \end{aligned} \tag{4.14}$$

where $x^I$ are coordinates on $M$. In the special case that the manifold $M$ is fibred over $I\!R$ we can think of time as the zeroth coordinate and write $x^I = (t, x^i)$. We recover the situation that we discussed earlier in the special case that $\rho$ is of the form $\iota_{J^1M} \circ \lambda$, with $\lambda : E \to J^1M$ as above and $\iota_{J^1M} : J^1M \to TM$ the natural injection of $J^1M$ into $TM$, and thus $\rho^0_0 = 0$, $\rho^0_\alpha = 0$, $\rho^i_0 = \lambda^i$ and $\rho^i_\alpha = \lambda^i_\alpha$. In the following, we will refer to this special situation within the general theory of affine Lie algebroids as the 'time-dependent case'.

## 4.3 'Rudimentary' calculus of variations

We sketch how one can relate the 'Lagrange' equations (4.11) of the time-dependent case with a calculus of variations problem for curves $t \mapsto \gamma(t) = (t, x^i(t), y^\alpha(t))$ in $I\!R^{n+k+1}$ say. Assume we have a given functional

$$\mathcal{J}(\gamma) = \int_{t_1}^{t_2} L(t, x(t), y(t)) dt,$$

and want to find its extremals, within arbitrary one-parameter families of curves which satisfy the constraints

$$\dot{x}^i = \lambda^i_\alpha(t,x) y^\alpha + \lambda^i(t,x). \tag{4.15}$$

We will proceed in a very formal way, without worrying too much about the mathematical complications which come from constraints depending on velocities. Formally, taking variations of the constraint equations, we get:

$$\delta \dot{x}^i = \left( \frac{\partial \lambda^i_\alpha}{\partial x^j} y^\alpha + \frac{\partial \lambda^i}{\partial x^j} \right) \delta x^j + \lambda^i_\alpha \delta y^\alpha.$$

Multiplying these by Lagrange multipliers $p_i$ and adding the result to the variation of the functional, one obtains (after an integration by parts on the term $p_i\delta\dot{x}^i$)

$$\int_{t_1}^{t_2}\left[\left(\frac{\partial L}{\partial x^j}-\dot{p}_j-p_i\Big(\frac{\partial\lambda^i_\alpha}{\partial x^j}y^\alpha+\frac{\partial\lambda^i}{\partial x^j}\Big)\right)\delta x^j+\left(\frac{\partial L}{\partial y^\alpha}-p_i\lambda^i_\alpha\right)\delta y^\alpha\right]dt=0$$

It is tacitly assumed that all variations $\delta x^i$ and $\delta y^\alpha$ vanish at the endpoints, thereby skipping over the mathematical complications which come from the differential equations they have to satisfy. The traditional argument then is that one can choose the multipliers $p_i$ in such a way that the coefficients of $\delta x^j$ vanish, leaving only terms in $\delta y^\alpha$, which are arbitrary, so that those coefficients must vanish in view of the fundamental lemma of the calculus of variations. We thus get the equations

$$\begin{aligned}\dot{p}_j &= \frac{\partial L}{\partial x^j}-p_i\Big(\frac{\partial\lambda^i_\alpha}{\partial x^j}y^\alpha+\frac{\partial\lambda^i}{\partial x^j}\Big),\\ \frac{\partial L}{\partial y^\alpha} &= p_j\lambda^j_\alpha.\end{aligned}$$

The next step one would like to take is to eliminate the $p_i$. Taking the total time derivative of the second equations and using the first to substitute for $\dot{p}_i$, one is left with a number of terms containing $p_j$, which will vanish only if they combine in such a way that they pick up a factor $\lambda^j_\alpha$. Therefore, an interesting situation is the case that there exist functions $C^\gamma_{\alpha\beta}(t,x)$ and $C^\alpha_{0\beta}(t,x)$ such that:

$$\begin{aligned}\lambda^i_\alpha\frac{\partial\lambda^j_\beta}{\partial x^i}-\lambda^i_\beta\frac{\partial\lambda^j_\alpha}{\partial x^i}&=\lambda^j_\gamma C^\gamma_{\alpha\beta},\\ \frac{\partial\lambda^j_\beta}{\partial t}+\lambda^i\frac{\partial\lambda^j_\beta}{\partial x^i}-\lambda^i_\beta\frac{\partial\lambda^j}{\partial x^i}&=\lambda^j_\alpha C^\alpha_{0\beta}.\end{aligned}\tag{4.16}$$

These relations clearly generalise the conditions (4.4). The equations which result from the elimination of the $p_i$ then are

$$\frac{d}{dt}\left(\frac{\partial L}{\partial y^\alpha}\right)=\lambda^i_\alpha\frac{\partial L}{\partial x^i}-(C^\gamma_{\alpha\beta}y^\beta-C^\gamma_{0\alpha})\frac{\partial L}{\partial y^\gamma}\,,\tag{4.17}$$

and they of course have to be supplemented by the constraints (4.15).

The point of this formal exercise is the following: if one carries out the same procedure in an autonomous framework, one arrives exactly at the equations

(4.8), with functions $\varrho_a^i(x)$ and $\mathsf{C}_{ab}^{\,c}(x)$ satisfying the relations (4.4); therefore, we can feel confident that the more general equations we derived in this section are indeed the ones we are looking for. Our programme thus becomes: identify now an appropriate geometrical framework for generalisation of the classical notion of a Lie algebroid, which gives rise to compatibility conditions of the type we have just encountered, and within which time-dependent Lagrange equations of the form (4.17) can be accommodated.

Note in passing that we did not encounter Jacobi-type conditions in our formal analysis, which means that it may even make sense to relax the axioms of a Lie algebroid if the purpose merely would be to describe differential equations of the form (4.17), but this is a path we do not wish to explore.

# Chapter 5

# *Lie algebra structure on an affine space*

## 5.1 Immersion of an affine space in a vector space

Before arriving at the definition of a Lie algebroid structure on an affine bundle, we need to say a few words about affine spaces. To avoid confusion in terminology, we start this section with a few elementary definitions. For a larger overview, see e.g. [20] or [81].

**Definition 5.1.** *A space $A$ is an affine space modelled on a vector space $\overline{A}$ if there is a map $- : A \times A \to \overline{A}$, $(a,b) \mapsto b-a$ satisfying the following two rules:*

1. *for all $a,b,c \in A$, $(b-a)+(c-b) = c-a$;*
2. *for every $a \in A$, the map $b \mapsto b-a$ is a one-to-one map from $A$ onto $\overline{A}$.*

*The dimension of $A$ is by definition the dimension of $\overline{A}$.*

Of course, $+$ stands here for the additive operation in the vector space $\overline{A}$. We will use the same symbol also for the affine action of $\overline{A}$ on $A$, for example if $a \in A$ and $\overline{a} \in \overline{A}$ are given, then $b = a + \overline{a}$ stands for the unique point in $A$ such that $b - a = \overline{a}$.

Let $A$ and $B$ now be two affine spaces, modelled respectively on the vector spaces $\overline{A}$ and $\overline{B}$. A map $\varphi : A \to B$ is called *affine* if there exists a linear map $\overline{\varphi} : \overline{A} \to \overline{B}$ such that

$$\varphi(a+\overline{a}) = \varphi(a) + \overline{\varphi}(\overline{a}) \tag{5.1}$$

for all $a \in A$ and $\overline{a} \in \overline{A}$. In particular, if $B = I\!R = \overline{B}$, then we say that $\varphi$ is an affine function. The set of all affine functions on $A$ is denoted with

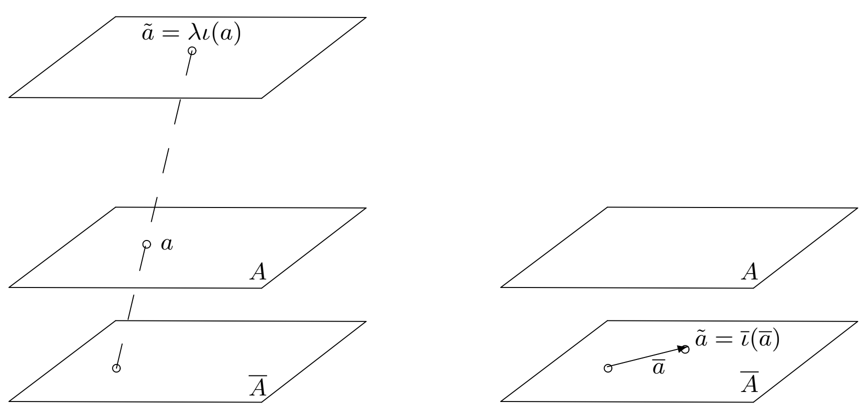
$\mathrm{Aff}(A, I\!R)$. Since the set of all linear functions on $\overline{A}$ is nothing but the dual of $\overline{A}$, in the following written as $\overline{A}^*$, an affine function $\varphi : A \to I\!R$ satisfies (5.1) for some $\overline{\varphi} \in \overline{A}^*$. The set of affine functions can be given a vector space structure: for affine functions $\varphi_1, \varphi_2$, let $\varphi_1 + \varphi_2$ be the function defined by $(\varphi_1 + \varphi_2)(a) := \varphi_1(a) + \varphi_2(a)$. Evidently, the function $\varphi_1 + \varphi_2$ is affine again, modelled on $\overline{\varphi}_1 + \overline{\varphi}_2 \in \overline{A}^*$. The vector space $A^\dagger = \mathrm{Aff}(A, I\!R)$ is called the *extended dual of* $A$. We consider now the bi-dual $\tilde{A}$ of $A$, in the sense $\tilde{A} = (A^\dagger)^*$. It is well known that, in the case of a vector space $\mathsf{V}$, the bi-dual $\tilde{\mathsf{V}} = (\mathsf{V}^*)^*$ is isomorphic to $\mathsf{V}$ again. In the case of an affine space, the bi-dual $\tilde{A}$ includes 'a copy' of $A$, as is shown in the following statement.

**Proposition 5.2.** *The map* $\iota \colon A \to \tilde{A}$ *given by* $\iota(a)(\varphi) = \varphi(a)$ *is an injective affine map, whose associated vector map is* $\overline{\iota} \colon \overline{A} \to \tilde{A}$ *given by* $\overline{\iota}(\overline{a})(\varphi) = \overline{\varphi}(\overline{a})$

Proof: If $a \in A$ and $\overline{a} \in \overline{A}$, then for all $\varphi \in A^\dagger$,

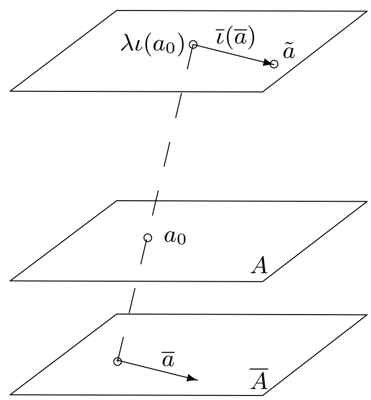
$$\iota(a + \overline{a})(\varphi) = \varphi(a + \overline{a}) = \varphi(a) + \overline{\varphi}(\overline{a}) = \iota(a)(\varphi) + \overline{\iota}(\overline{a})(\varphi),$$

from which it follows that $\iota$ is an affine map whose associated linear map is $\overline{\iota}$. Now, to prove that $\iota$ is injective, it suffices to prove that $\overline{\iota}$ is injective, which is obvious since if $\overline{a}$ is an element in the kernel of $\overline{\iota}$ then $\overline{\iota}(\overline{a})(\overline{\varphi}) = \overline{\varphi}(\overline{a}) = 0$ for all $\overline{\varphi} \in \overline{A}^*$, hence $\overline{a} = 0$. $\square$




The vector space $\tilde{A}$ is foliated by hyperplanes parallel to the image of $\overline{\iota}$. There are two ways to pin down an arbitrary point $\tilde{a} \in \tilde{A}$. It is easy to see that every vector $\tilde{a} \in \tilde{A}$ is either of the form $\tilde{a} = \overline{\iota}(\overline{a})$ for some $\overline{a} \in \overline{A}$ or of the form $\tilde{a} = \lambda\iota(a)$ for some $\lambda \in I\!R \setminus \{0\}$ and $a \in A$. Moreover, $\lambda$ and $a$, or $\overline{a}$ are uniquely determined by $\tilde{a}$ (see the above pictures). The image of the map $\iota$ consists of the points for which $\lambda = 1$.

There exists also a second way to look at an arbitrary $\tilde{a} \in \tilde{A}$. If we fix one specific $a_0 \in A$, there exists a unique point $\lambda\iota(a_0)$ in the hyperplane through $\tilde{a}$ that maps onto $a_0$ in $A$ ($\lambda$ can be zero). The difference between $\tilde{a}$ and $\lambda\iota(a_0)$ further fixes a vector $\overline{a} \in \overline{A}$ (which can possibly be the zero vector). Thus, for a fixed $a_0 \in A$, every $\tilde{a}$ is of the form $\lambda\iota(a_0) + \overline{\iota}(\overline{a})$ for some $\lambda \in I\!R$ and $\overline{a} \in \overline{A}$.




To understand this description in more detail, we will prove that we have an exact sequence of vector spaces $0 \longrightarrow \overline{A} \stackrel{\overline{\iota}}{\longrightarrow} \tilde{A} \longrightarrow I\!R \longrightarrow 0$. To this end we consider the dual sequence.

**Proposition 5.3.** *Let $l\colon I\!R \to A^\dagger$ be the map that associates to $\lambda \in I\!R$ the constant function $\lambda$ on $A$. Let $k\colon A^\dagger \to \overline{A}^*$ denote the map that associates to every affine function on $A$ the corresponding linear function on $\overline{A}$. Then, the sequence of vector spaces*

$$0 \longrightarrow I\!R \stackrel{l}{\longrightarrow} A^\dagger \stackrel{k}{\longrightarrow} \overline{A}^* \longrightarrow 0$$

*is exact.*

PROOF: Indeed, it is clear that $l$ is injective, $k$ is surjective and $k \circ l = 0$, so that $\mathrm{Im}(l) \subset \mathrm{Ker}(k)$. If $\varphi \in A^\dagger$ is in the kernel of $k$, that is, if the linear part of $\varphi$ vanishes, then for every pair of points $a$ and $b = a + \overline{a}$ we have that

$$\varphi(b) = \varphi(a + \overline{a}) = \varphi(a) + \overline{\varphi}(\overline{a}) = \varphi(a),$$

and thus $\varphi$ is constant, and hence contained in the image of $l$. We can conclude that $\mathrm{Im}(l) = \mathrm{Ker}(k)$. $\square$

The dual map of $k$ is $\bar{\imath}$, since for $\overline{a} \in \overline{A}$ we have

$$\langle k(\varphi), \overline{a}\rangle = \langle \overline{\varphi}, \overline{a}\rangle = \langle \varphi, \bar{\imath}(\overline{a})\rangle$$

The dual map $\jmath$ of the map $l$ is given by $\jmath(\alpha\iota(a) + \bar{\imath}(\overline{a})) = \alpha$. Indeed, for every $\lambda \in I\!R$ we have

$$\jmath(\tilde{a})\lambda = \langle \tilde{a}, l(\lambda)\rangle = \langle \alpha\iota(a) + \bar{\imath}(\overline{a}), l(\lambda)\rangle = \alpha\langle \iota(a), l(\lambda)\rangle + \langle \bar{\imath}(\overline{a}), l(\lambda)\rangle = \alpha\lambda$$

It follows that

**Corollary 5.4.** *If $A$ is finite dimensional, then the sequence*

$$0 \longrightarrow \overline{A} \overset{\bar{\imath}}{\longrightarrow} \tilde{A} \overset{\jmath}{\longrightarrow} I\!R \longrightarrow 0$$

*is exact.*

Note that in this way we can clearly identify the image of $\overline{A}$ as the hyperplane of $\tilde{A}$ determined by the equation $\jmath(\tilde{a}) = 0$, and the image of $A$ as the hyperplane of $\tilde{A}$ with equation $\jmath(\tilde{a}) = 1$, in other words

$$\bar{\imath}(\overline{A}) = \jmath^{-1}(0) \qquad \text{and} \qquad \iota(A) = \jmath^{-1}(1).$$

Note in passing that, if we have an exact sequence $0 \longrightarrow V \overset{\alpha}{\longrightarrow} W \overset{\jmath}{\longrightarrow} I\!R \longrightarrow 0$, then we can define $A = \jmath^{-1}(1)$; it follows that $A$ is an affine space modelled on the vector space $V$ and $W$ is canonically isomorphic to $\tilde{A}$. The isomorphism is the dual map of $\Psi : W^* \to A^\dagger$, $\Psi(\phi)(a) = \phi(i(a))$, where $i : A \to W$ is the canonical inclusion.

## 5.2 Coordinates on an affine space

Suppose the dimension of $A$ is $N$. We now discuss the construction of a basis for $A^\dagger$. For that purpose, we need to chose an origin, i.e. a point $o$ in $A$, and a basis $\{\overline{e}_1, \dots \overline{e}_N\}$ for the vector space $\overline{A}$ on which $A$ is modelled. The set $(o, \{\overline{e}_1, \dots, \overline{e}_N\})$ is called *an affine frame on* $A$. In some situations we will also make use of the affine points

$$\hat{e}_\alpha = o + \overline{e}_\alpha, \qquad \alpha = 1 \dots N. \tag{5.2}$$

For any point $a \in A$, $a - o \in \overline{A}$, and there exist uniquely defined coordineates $a^\alpha$ such that

$$a - o = a^\alpha \overline{e}_\alpha.$$

The family of affine maps $\{e^0, e^1, \ldots, e^N\}$ given by

$$e^0(a) = 1 \qquad e^\alpha(a) = a^\alpha, \tag{5.3}$$

is a basis for $A^\dagger$. If $\varphi \in A^\dagger$, and we put $\varphi_0 = \varphi(o)$ and $\varphi_\alpha = \overline{\varphi}(\overline{e}_\alpha)$, then $\varphi = \varphi_0 e^0 + \varphi_\alpha e^\alpha$.

It is to be noticed that, contrary to $e^1, \ldots, e^N$, the map $e^0$ does not depend on the frame we have chosen for $A$. In fact, $e^0$ coincides with the map $\jmath$.

Let now $\{e_0, e_1, \ldots, e_N\}$ denote the basis of $\tilde{A}$ dual to $\{e^0, e^1, \ldots, e^N\}$. Then the image of the canonical immersion is given by

$$\iota(o) = e_0 \qquad \overline{\iota}(\overline{e}_\alpha) = e_\alpha \tag{5.4}$$

from which it follows that for $a = o + a^\alpha \overline{e}_\alpha$, we have $\iota(a) = e_0 + a^\alpha e_\alpha$. In particular, for the points of the form (5.2), we find $\iota(\hat{e}_\alpha) = e_0 + e_\alpha$. If we denote by $(y^0, y^1, \ldots, y^N)$ the coordinate system on $\tilde{A}$ associated to the basis $\{e_0, \ldots, e_N\}$, then the equation of the image of the map $\iota$ is $y^0 = 1$, while the equation of the image of $\overline{\iota}$ is $y^0 = 0$.

Coordinates in $\tilde{A}^* = A^\dagger$ associated to the basis above will be denoted by $(\mu_0, \mu_1, \ldots, \mu_N)$, that is $\mu_\alpha(\varphi) = \langle \overline{e}_\alpha, \varphi \rangle$ for every $\varphi \in A^\dagger$.

## 5.3 Exterior algebra on an affine space

A $k$-form on a vector space $\overline{A}$ is a skew-symmetric and multi-linear map $\overline{A} \times \cdots \times \overline{A} \to I\!R$. Such a form can equivalently be seen to be made up from the exterior powers of elements of the dual space $\overline{A}^*$. In this section, we would like to find a notion of forms on the affine space $A$. One can start by taking elements of $A^\dagger$ as 1-forms and look at its exterior powers. We will be mainly interested in the action of such elements on elements of $A$. This brings some subtleties into the picture which need to be investigated in sufficient detail. Now, to begin with, let $a \in A$. Expression (5.1) can be put in slightly different terms: a $\varphi \in A^\dagger$ is such that there exist a $\varphi_0 \in A^\dagger$ and a $\overline{\varphi} \in \overline{A}^*$, such that for all $a \in A$:

$$\varphi(a) = \varphi_0(a_0) + \overline{\varphi}(\overline{a}), \tag{5.5}$$

where $a_0$ is any element and the corresponding $\overline{a}$ can be obtained from $a = a_0 + \overline{a}$. The two composing elements $\varphi_0$ (which in fact is simply $\varphi$ itself here) and $\overline{\varphi}$ do not depend on the choice of $a_0$. With elements of $A^\dagger$ as our notion of 1-forms on $A$, there is of course no linearity with respect to multiplication by real numbers. We can now come in a similar way to the following concept of $k$-forms on $A$.

**Definition 5.5.** *A $k$-form on an affine space $A$ is a map $\omega\colon \underbrace{A \times \cdots \times A}_{k} \to$ $I\!R$ for which there exists a $k$-form $\overline{\omega}$ on the associated vector space $\overline{A}$, and a map $\omega_0\colon A \times \underbrace{\overline{A} \times \cdots \times \overline{A}}_{k-1} \to I\!R$ with the following properties*

1. *$\omega_0$ is skew-symmetric and linear in its $k-1$ vector arguments.*
2. *For every $a \in A$ and for every $\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_k \in \overline{A}$, we have*

$$\omega_0(a + \overline{a}_1, \overline{a}_2, \ldots, \overline{a}_k) = \omega_0(a, \overline{a}_2, \ldots, \overline{a}_k) + \overline{\omega}(\overline{a}_1, \overline{a}_2, \ldots, \overline{a}_k). \quad (5.6)$$

3. *For every $a_1, \ldots, a_k \in A$, if we choose an arbitrary $a_0 \in A$ and put $a_i = a_0 + \overline{a}_{0i}$, we have*

$$\omega(a_1, \ldots, a_k) = \sum_{i=1}^{k} (-1)^{i-1} \omega_0(a_0, \overline{a}_{01}, \ldots, \widehat{\overline{a}_{0i}}, \ldots, \overline{a}_{0k}) + \overline{\omega}(\overline{a}_{01}, \overline{a}_{02}, \ldots, \overline{a}_{0k}). \quad (5.7)$$

We will denote the set of $k$-forms on $A$ by $\Omega^k(A^\dagger)$. There are a number of properties to be checked to make sure that this definition makes sense. First of all, one can verify that with two different choices of reference point, $a_0$ and $a_0'$ for example, related through $a_0 = a_0' + \overline{b}$, it follows from the second requirement that, if $a_i = a_0' + \overline{a}_{0i}'$,

$$\begin{aligned} &\sum_{i=1}^{k} (-1)^{i-1} \omega_0(a_0, \overline{a}_{01}, \ldots, \widehat{\overline{a}_{0i}}, \ldots, \overline{a}_{0k}) + \overline{\omega}(\overline{a}_{01}, \ldots, \overline{a}_{0k}) \\ &\qquad = \sum_{i=1}^{k} (-1)^{i-1} \omega_0(a_0', \overline{a}_{01}', \ldots, \widehat{\overline{a}_{0i}'}, \ldots, \overline{a}_{0k}') + \overline{\omega}(\overline{a}_{01}', \ldots, \overline{a}_{0k}'). \end{aligned}$$

Secondly, the two elements $\omega_0$ and $\overline{\omega}$ which make up $\omega$ are unique. Indeed, assuming there would be a second couple $\omega_0'$ and $\overline{\omega}'$ making up the same $\omega$, it follows by choosing $a_0 = a_1$ (such that $\overline{a}_{01} = 0$) in the third requirement that $\omega_0 = \omega_0'$, after which it is clear that also $\overline{\omega} = \overline{\omega}'$. Note finally that the definition implies that $\omega$ itself is skew-symmetric in all its arguments.

The meaning of our forms will become more transparent if we can establish a relation with ordinary forms on the vector space $\tilde{A}$. We next show that a $k$-form on $A$ is just the pull-back by the canonical immersion of a $k$-form on $\tilde{A}$, in other words, $\omega$ is a $k$-form if we have

$$\omega(a_1, \ldots, a_k) = \tilde{\omega}(\iota(a_1), \ldots, \iota(a_k)), \tag{5.8}$$

for some ordinary exterior $k$-form $\tilde{\omega}$ on $\tilde{A}$.

**Proposition 5.6.** *If $\tilde{\omega}$ is a $k$-form on $\tilde{A}$ then $\omega = \iota^*\tilde{\omega}$ is a $k$-form on the affine space $A$. Conversely, given a $k$-form on the affine space $A$, there exists a unique $k$-form $\tilde{\omega}$ on $\tilde{A}$ such that $\omega = \iota^*\tilde{\omega}$.*

PROOF: For a given $k$-form $\tilde{\omega}$ on $\tilde{A}$ we define the maps

$$\begin{aligned}
&\omega(a_1, \ldots, a_k) = \tilde{\omega}(\iota(a_1), \ldots, \iota(a_k)),\\
&\omega_0(a, \overline{a}_2, \ldots, \overline{a}_k) = \tilde{\omega}(\iota(a), \overline{\iota}(\overline{a}_2) \ldots, \overline{\iota}(\overline{a}_k)),\\
&\overline{\omega}(\overline{a}_1 \ldots, \overline{a}_k) = \tilde{\omega}(\overline{\iota}(\overline{a}_1) \ldots, \overline{\iota}(\overline{a}_k)).
\end{aligned}$$

Then, conditions *1* and *2* in the definition above are trivially satisfied. Moreover, if we fix $a_0 \in A$ and write $a_i = a_0 + \overline{a}_{0i}$, then by skew-symmetry of $\tilde{\omega}$ we have

$$\begin{aligned}
\omega(a_1, \ldots, a_k) &= \tilde{\omega}(\iota(a_1), \ldots, \iota(a_k))\\
&= \tilde{\omega}(\iota(a_0) + \overline{\iota}(\overline{a}_{01}), \ldots, \iota(a_0) + \overline{\iota}(\overline{a}_{0k}))\\
&= \sum_{j=1}^{k} (-1)^{j+1} \tilde{\omega}(\iota(a_0), \overline{\iota}(\overline{a}_{01}), \ldots, \widehat{\overline{\iota}(\overline{a}_{0j})}, \ldots, \overline{\iota}(\overline{a}_{0k}))\\
&\qquad + \tilde{\omega}(\overline{\iota}(\overline{a}_{01}), \ldots, \overline{\iota}(\overline{a}_{0k}))\\
&= \sum_{j=1}^{k} (-1)^{j+1} \omega_0(a_0, \overline{a}_{01}, \ldots, \widehat{\overline{a}_{0j}}, \ldots, \overline{a}_{0k}) + \overline{\omega}(\overline{a}_{01}, \ldots, \overline{a}_{0k}),
\end{aligned}$$

which proves condition *3*.

Conversely, assume we are given a $k$-form $\omega$ on the affine space $A$ with its associated $\omega_0$ and $\overline{\omega}$. Fixing $a_0 \in A$, we know that every point $\tilde{a}_i \in \tilde{A}$ can be written in the form $\tilde{a}_i = \lambda_i \iota(a_0) + \bar{\iota}(\overline{a}_{0i})$ for $\lambda_i \in I\!R$ and $\overline{a}_{0i} \in \overline{A}$. We define the map $\tilde{\omega}$ by

$$\tilde{\omega}(\tilde{a}_1, \ldots, \tilde{a}_k) = \tilde{\omega}(\lambda_1 \iota(a_0) + \bar{\iota}(\overline{a}_{01}), \ldots, \lambda_k \iota(a_0) + \bar{\iota}(\overline{a}_{0k}))$$
$$= \sum_{j=1}^{k} (-1)^{j+1} \lambda_j \, \omega_0(a_0, \overline{a}_{01}, \ldots, \widehat{\overline{a}_{0j}}, \ldots, \overline{a}_{0k}) + \overline{\omega}(\overline{a}_{01}, \ldots, \overline{a}_{0k}).$$

By virtue of conditions *1* and *3* in Definition 5.5, it follows that $\tilde{\omega}$ is multilinear and skew-symmetric, i.e. it is a $k$-form on $\tilde{A}$. Moreover, $\tilde{\omega}(\tilde{a}_1, \ldots, \tilde{a}_k)$ is independent of the choice of the point $a_0$. Indeed, if we choose a different point $a_0' = a_0 + \overline{b}$, then $\tilde{a}_j = \lambda_j \iota(a_0') + \bar{\iota}(\overline{a}_{0j}')$ with $\overline{a}_{0j}' = \overline{a}_{0j} - \lambda_j \overline{b}$, and applying the definition above we get

$$\begin{aligned}
\tilde{\omega}(\tilde{a}_1, \ldots, \tilde{a}_k) &= \tilde{\omega}(\lambda_1 \iota(a_0') + \bar{\iota}(\overline{a}_{01}'), \ldots, \lambda_k \iota(a_0') + \bar{\iota}(\overline{a}_{0k}')) \\
&= \sum_{j=1}^{k} (-1)^{j+1} \lambda_j \, \omega_0(a_0', \overline{a}_{01}', \ldots, \widehat{\overline{a}_{0j}'}, \ldots, \overline{a}_{0k}') + \overline{\omega}(\overline{a}_{01}', \ldots, \overline{a}_{0k}') \\
&= \sum_{j=1}^{k} (-1)^{j+1} \lambda_j \, \omega_0(a_0 + \overline{b}, \overline{a}_{01} - \lambda_1 \overline{b}, \ldots, \widehat{\overline{a}_{0j}}, \ldots, \overline{a}_{0k} - \lambda_k \overline{b}) \\
&\qquad + \overline{\omega}(\overline{a}_{01} - \lambda_1 \overline{b}, \ldots, \overline{a}_{0k} - \lambda_k \overline{b}) \\
&= \sum_{j=1}^{k} (-1)^{j+1} \lambda_j \, \omega_0(a_0, \overline{a}_{01}, \ldots, \widehat{\overline{a}_{0j}}, \ldots, \overline{a}_{0k}) + \overline{\omega}(\overline{a}_{01}, \ldots, \overline{a}_{0k})
\end{aligned}$$

where we have used the properties of $\omega_0$ and $\overline{\omega}$.

The form $\tilde{\omega}$ is unique, since, if $\tilde{\theta}$ is a $k$-form on $\tilde{A}$ such that $\iota^* \tilde{\theta} = 0$, then it follows that the associated $\theta_0$ and $\overline{\theta}$ vanish from where we deduce that $\tilde{\theta} = 0$. □

Making use of the identification (5.8) there is no difficulty in defining operations such as the wedge product for exterior forms on an affine space $A$. It is of some interest, however, to investigate how one can do that without recourse to the vector space $\tilde{A}$ which contains $A$. The main point to understand then is the following: for $\omega = \varphi \wedge \psi$ say, how do $\varphi_0$, $\overline{\varphi}$ and $\psi_0$, $\overline{\psi}$ give rise in a consistent way to the defining components $\omega_0$ and $\overline{\omega}$ of $\omega$? For a start, if $\varphi$ and $\psi$ are 1-forms, we have

$$(\varphi \wedge \psi)(a_1, a_2) = \varphi(a_1)\psi(a_2) - \varphi(a_2)\psi(a_1),$$

which for every choice of a reference point $a_0$ gives rise to:

$$(\varphi \wedge \psi)(a_1, a_2) = \varphi(a_0)\overline{\psi}(\overline{a}_{02} - \overline{a}_{01}) - \psi(a_0)\overline{\varphi}(\overline{a}_{02} - \overline{a}_{01}) + (\overline{\varphi} \wedge \overline{\psi})(\overline{a}_{01}, \overline{a}_{02}). \tag{5.9}$$

It follows that $\overline{\varphi} \wedge \overline{\psi}$ is the 2-form on $\overline{A}$ corresponding to $\varphi \wedge \psi$, and

$$(\varphi \wedge \psi)_0(a, \bar{b}) = \varphi(a)\overline{\psi}(\bar{b}) - \psi(a)\overline{\varphi}(\bar{b}). \tag{5.10}$$

Similarly, for the wedge product of three 1-forms, we have

$$(\varphi \wedge \psi \wedge \chi)_0(a, \bar{b}_1, \bar{b}_2) = \Big(\varphi(a)(\overline{\psi} \wedge \overline{\chi}) + \psi(a)(\overline{\chi} \wedge \overline{\varphi}) + \chi(a)(\overline{\varphi} \wedge \overline{\psi})\Big)(\bar{b}_1, \bar{b}_2). \tag{5.11}$$

These examples suggest to formalise the representation of $k$-forms a bit further. As a preliminary remark, it may sometimes be of interest to extend the interpretation of the operator $\omega_0$ in such a way that its single affine argument need not necessarily be the first. This can simply be achieved by declaring $\omega_0$ to be skew-symmetric in all its arguments (but still $\mathbb{R}$-linear in its vector arguments only). More importantly, we shall take the sum of $\omega_0$-terms in (5.7) to define another operator, denoted by $\omega^0$, as follows:

$$\begin{aligned} \omega^0(a_1, \dots, a_k) &= \sum_{i=1}^{k} (-1)^{i-1} \omega_0(a_0, \overline{a}_{01}, \dots, \widehat{\overline{a}_{0i}}, \dots, \overline{a}_{0k}) \\ &= \sum_{i=1}^{k} \omega_0(\overline{a}_{01}, \dots, \underset{i}{\check{a}_0}, \dots, \overline{a}_{0k}), \end{aligned} \tag{5.12}$$

where the second expression takes the above remark into account and the symbol $\underset{i}{\check{a}_0}$ then indicates that $a_0$ has been inserted in the $i$-th argument. The other new convention we will adopt is to regard $\overline{\omega}$ also as acting on elements of $A$, by putting:

$$\overline{\omega}(a_1, \dots, a_k) = \overline{\omega}(\overline{a}_{01}, \dots, \overline{a}_{0k}). \tag{5.13}$$

This way, we can formally write

$$\omega = \omega^0 + \overline{\omega}, \tag{5.14}$$

whereby it is to be understood that the two composing terms $\omega^0$ and $\overline{\omega}$ are not $k$-forms on $A$ by themselves. In fact, to compute their value when acting on $k$ elements $a_i$, a reference point $a_0$ has to be chosen, but as argued above, the value of the sum $\omega^0 + \overline{\omega}$ in the end does not depend on that choice.

The rather formal looking decomposition (5.14) now greatly facilitates the representation of wedge products and will make the general coordinate representation of a form more transparent. For example, the result (5.11) means that

$$(\varphi \wedge \psi \wedge \chi)_0 = \varphi^0 \otimes (\overline{\psi} \wedge \overline{\chi}) + \psi^0 \otimes (\overline{\chi} \wedge \overline{\varphi}) + \chi^0 \otimes (\overline{\varphi} \wedge \overline{\psi}), \tag{5.15}$$

which then implies from (5.12) that

$$(\varphi \wedge \psi \wedge \chi)^0 = \varphi^0 \wedge \overline{\psi} \wedge \overline{\chi} + \overline{\varphi} \wedge \psi^0 \wedge \overline{\chi} + \overline{\varphi} \wedge \overline{\psi} \wedge \chi^0, \tag{5.16}$$

as expected. More generally, it follows directly from the defining formula for wedge products that for $\omega = \omega^0 + \overline{\omega}$ and $\nu = \nu^0 + \overline{\nu}$:

$$\omega \wedge \nu = \omega^0 \wedge \nu^0 + \omega^0 \wedge \overline{\nu} + \overline{\omega} \wedge \nu^0 + \overline{\omega} \wedge \overline{\nu}, \tag{5.17}$$

where the sum of the first three terms is $(\omega \wedge \nu)^0$.

Let us have a look at coordinate expressions. Suppose that, for a coordinatisation of $A$, we have chosen an origin $o$ and a local basis of vectors $\overline{e}_\alpha$. We have shown in Section 5.2 that any affine frame gives rise to an adapted basis (5.3) for $A^\dagger$. In this basis, any $\varphi \in A^\dagger$ has the local representation

$$\varphi = \varphi_0 e^0 + \varphi_\alpha e^\alpha, \tag{5.18}$$

where, in agreement with the general decomposition (5.14) and their definition in (5.3), $e^\alpha$ acts on elements of $A$ and $\varphi^0 = \varphi_0 e^0$. There is a slight abuse of notation in (5.18) since $\varphi_0$ could have a double meaning: in (5.18) it represents a real number, whereas it also could refer to the operator introduced in (5.5) and more generally in Definition 5.5. We will, however, seldom use the notation $\varphi_0$ in the latter sense when dealing with coordinate calculations, so that the meaning will always be clear from the context.

Let us now see how all the notations fit together when we start wedging 1-forms. For two 1-forms $\varphi = \varphi^0 + \overline{\varphi}$ and $\psi = \psi^0 + \overline{\psi}$ we find, for example from (5.10) and (5.12), that

$$(\varphi \wedge \psi)^0 = \varphi^0 \wedge \overline{\psi} - \psi^0 \wedge \overline{\varphi}. \tag{5.19}$$

This is in agreement with the general formula (5.17) since obviously $\varphi^0 \wedge \psi^0 = 0$. Expressing $\varphi$ and $\psi$ with respect to the basis $(e^0, e^\alpha)$, we find

$$\varphi \wedge \psi = (\varphi_0 \psi_\gamma - \psi_0 \varphi_\gamma)\, e^0 \wedge e^\gamma + \tfrac{1}{2}(\varphi_\gamma \psi_\delta - \varphi_\delta \psi_\gamma)\, e^\gamma \wedge e^\delta. \tag{5.20}$$

Similarly, for the wedge product of three 1-forms with local representations of the form (5.18), we obtain

$$\begin{aligned}\varphi \wedge \psi \wedge \chi \;&=\; \tfrac{1}{2}\Big(\varphi_0(\psi_\mu\chi_\nu - \psi_\nu\chi_\mu) + \psi_0(\chi_\mu\varphi_\nu - \chi_\nu\varphi_\mu) \\ &\quad + \chi_0(\varphi_\mu\psi_\nu - \varphi_\nu\psi_\mu)\Big)e^0 \wedge e^\mu \wedge e^\nu + \overline{\varphi} \wedge \overline{\psi} \wedge \overline{\chi}. \qquad (5.21)\end{aligned}$$

It should now be clear without going into further detail that a general $k$-form on $A$ locally has the following representation,

$$\omega = \frac{1}{(k-1)!}\,\omega_{0\mu_1\cdots\mu_{k-1}}\, e^0 \wedge e^{\mu_1} \wedge \cdots \wedge e^{\mu_{k-1}} + \frac{1}{k!}\,\omega_{\mu_1\cdots\mu_k}\, e^{\mu_1} \wedge \cdots \wedge e^{\mu_k}, \quad (5.22)$$

where the coefficients are real numbers, which are skew-symmetric in all their indices (including the zero for the first term). These coefficients can be computed from the action of $\omega$ on the affine elements $\hat{e}_\alpha = o + \overline{e}_\alpha$ we have

$$\omega_{0\mu_1\cdots\mu_{k-1}} = \omega(o, \hat{e}_{\mu_1}, \ldots, \hat{e}_{\mu_{k-1}}), \qquad (5.23)$$

$$\omega_{\mu_1\cdots\mu_k} = \omega(\hat{e}_{\mu_1}, \ldots, \hat{e}_{\mu_k}) - \sum_{i=1}^{k} \omega_{\mu_1\cdots\underset{i}{\check{0}}\cdots\mu_k}, \qquad (5.24)$$

where $\underset{i}{\check{0}}$ again means that the index $\mu_i$ has been replaced by 0. It is easy to see that the representation (5.22) is in perfect agreement with the results of Proposition 5.6.

## 5.4 Lie algebra structure on an affine space

In this section we present a definition of a Lie algebra over an affine space. We will suppose that the definition of a Lie algebra over a vector space is known (see e.g. [38]).

**Definition 5.7.** *Let $A$ be an affine space over a vector space $\overline{A}$. A Lie algebra structure on $A$ is given by*

- *a Lie algebra structure $[\cdot,\cdot]$ on $\overline{A}$, and*

- *an action by derivations of $A$ on $\overline{A}$, i.e. a map $D\colon A\times\overline{A}\to\overline{A}$, $(a,\overline{a})\mapsto D_a\overline{a}$ with the properties ($\lambda\in I\!\!R$)*

$$D_a(\lambda\overline{a}) = \lambda D_a\overline{a} \tag{5.25}$$
$$D_a(\overline{a}_1+\overline{a}_2) = D_a\overline{a}_1 + D_a\overline{a}_2 \tag{5.26}$$
$$D_a[\overline{a}_1,\overline{a}_2] = [D_a\overline{a}_1,\overline{a}_2] + [\overline{a}_1, D_a\overline{a}_2], \tag{5.27}$$

- *satisfying the compatibility property*

$$D_{a+\overline{a}_1}\overline{a}_2 = D_a\overline{a}_2 + [\overline{a}_1,\overline{a}_2]. \tag{5.28}$$

One of the conditions in this definition is redundant. In the first item it is sufficient to require that the bracket on $\overline{A}$ is $I\!\!R$-bilinear and skew-symmetric, since the Jacobi identity then follows from the requirements on $D_a$. Indeed, if we replace $a$ in (5.27) by $a+\overline{a}_3$, for an arbitrary $\overline{a}_3\in\overline{A}$, it follows from (5.28) that $\forall\overline{a}_1,\overline{a}_2,\overline{a}_3\in\overline{A}$:

$$[\overline{a}_3,[\overline{a}_1,\overline{a}_2]] = [[\overline{a}_3,\overline{a}_1],\overline{a}_2]] + [\overline{a}_1,[\overline{a}_3,\overline{a}_2]], \tag{5.29}$$

which is the Jacobi identity for $[\cdot,\cdot]$.

In the following we will also use a bracket notation $[a,\overline{a}]$ for $D_a\overline{a}$. The conditions in the definition above then read

$$[a,\lambda\overline{a}] = \lambda[a,\overline{a}] \tag{5.30}$$
$$[a,\overline{a}_1+\overline{a}_2] = [a,\overline{a}_1] + [a,\overline{a}_2] \tag{5.31}$$
$$[a,[\overline{a}_1,\overline{a}_2]] = [[a,\overline{a}_1],\overline{a}_2] + [\overline{a}_1,[a,\overline{a}_2]] \tag{5.32}$$
$$[a+\overline{a}_1,\overline{a}_2] = [a,\overline{a}_2] + [\overline{a}_1,\overline{a}_2]. \tag{5.33}$$

We can now further extend the bracket operation to elements of $A$, as follows.

**Definition 5.8.** (i) *For $\overline{a}\in\overline{A}$ and $a\in A$, we put $[\overline{a},a] = -[a,\overline{a}]$.*

(ii) *For every two points $a_1,a_2\in A$ with $\overline{a}_{12} = a_2 - a_1$: $[a_1,a_2] = [a_1,\overline{a}_{12}]$.*

Observe that the extended bracket in (ii) is a map from $A\times A$ to $\overline{A}$. As we will show below, it also has Lie algebra type properties. Let 0 denote the zero vector in $\overline{A}$.

**Proposition 5.9.** *The bracket* $[\cdot,\cdot] : A \times A \to \overline{A}$*, has the following properties:*

$$[a_1, a_2 + \overline{a}] = [a_1, a_2] + [a_1, \overline{a}], \tag{5.34}$$
$$[a_1, a_2] = -[a_2, a_1], \tag{5.35}$$
$$[[a_1, a_2], a_3] + [[a_2, a_3], a_1] + [[a_3, a_1], a_2] = 0. \tag{5.36}$$

PROOF: The first property follows immediately from the definition and (5.31). Next, we have $[a_2, a_1] = [a_2, \overline{a}_{21}] = -[a_1 + \overline{a}_{12}, \overline{a}_{12}] = -[a_1, a_2]$. For the Jacobi identity, using a summation sign to indicate the cyclic sum over the three sections in each summand, we have

$$\sum [[a_1, a_2], a_3] = \sum [[a_1, a_2], a_2 + \overline{a}_{23}] = \sum [[a_1, a_2], \overline{a}_{23}],$$

in view of the linearity properties and the skew-symmetry of the bracket. Substituting subsequently $a_1 + \overline{a}_{12}$ for $a_2$, we obtain

$$\sum [[a_1, a_2], a_3] = \sum [[a_1, \overline{a}_{12}], \overline{a}_{23}],$$

which is zero in view of (5.27). □

**Proposition 5.10.** *A Lie algebra structure over an affine space $A$ is equivalent to a Lie algebra extension of the trivial Lie algebra $\mathbb{R}$ by $\overline{A}$. Explicitly, it is equivalent to the exact sequence of vector spaces $0 \longrightarrow \overline{A} \stackrel{\overline{\imath}}{\longrightarrow} \tilde{A} \stackrel{\jmath}{\longrightarrow} \mathbb{R} \longrightarrow 0$ being an exact sequence of Lie algebras.*

PROOF: If the exact sequence is one of Lie algebras, we of course have a Lie algebra structure on $\overline{A}$. The assumption that $\jmath$ is a Lie algebra homomorphism, implies that $\jmath([\tilde{a}_1, \tilde{a}_2]) = 0$, and thus $[\tilde{a}_1, \tilde{a}_2] \in \mathrm{Im}\overline{\imath}$. The map $D$ determined by $\overline{\imath}(D_a\overline{a}) = [\iota(a), \overline{\imath}(\overline{a})]$ satisfies all requirements to define a Lie algebra structure on $A$. For example,

$$\overline{\imath}(D_{a+\overline{a}_1}\overline{a}_2) = [\iota(a+\overline{a}_1), \overline{\imath}(\overline{a}_2)] = \overline{\imath}(D_a\overline{a}_2) + [\overline{\imath}(\overline{a}_1), \overline{\imath}(\overline{a}_2)] = \overline{\imath}\big(D_a\overline{a}_2 + [\overline{a}_1, \overline{a}_2]\big)$$

In the last step we have used the fact that $\overline{\imath}$ is a Lie algebra homomorphism. Since $\overline{\imath}$ is injective, property (5.28) follows.

Conversely, assume we have a Lie algebra structure on the affine space $A$. If we fix an element $a_0 \in A$, then every element $\tilde{a} \in \tilde{A}$ can be written in the form $\tilde{a} = \lambda\iota(a_0) + \overline{\iota}(\overline{a}_0)$. We can define a bracket of two elements $\tilde{a}_1 = \lambda_1\iota(a_0) + \overline{\iota}(\overline{a}_{01})$ and $\tilde{a}_2 = \lambda_2\iota(a_0) + \overline{\iota}(\overline{a}_{02})$ by

$$[\tilde{a}_1, \tilde{a}_2] = \overline{\iota}\big([\overline{a}_{01}, \overline{a}_{02}] + \lambda_1 D_a\overline{a}_{02} - \lambda_2 D_a\overline{a}_{01}\big)$$

This bracket is clearly bi-linear and skew-symmetric, and a straightforward calculation shows that it satisfies the Jacobi identity. Moreover, the definition does not depend on the choice of the point $a_0$; if $a_0'$ is another point in $A$, then $a_0 = a_0' + \overline{b}$ for some $\overline{b} \in \overline{A}$ and the compatibility condition implies that the result is independent of that choice. Finally, it is obvious that the maps $\iota$ and $\jmath$ then are Lie algebra homomorphisms. □

Notice that the only condition for a Lie algebra structure on $\tilde{A}$ to be an extension of $\mathbb{R}$ by $\overline{A}$ is that the bracket takes values in $\operatorname{Im}\overline{\iota}$, symbolically: $[\tilde{A}, \tilde{A}] \subset \overline{A}$. Indeed, this requirement means that $\jmath$ is a Lie algebra homomorphism, so that one can define a Lie algebra on $\overline{A}$ in the following way: if $\overline{a}_1, \overline{a}_2 \in \overline{A}$, $[\overline{a}_1, \overline{a}_2]$ is the unique $\overline{b}$ in $\overline{A}$, such that $\overline{\iota}(\overline{b}) = [\overline{\iota}(\overline{a}_1), \overline{\iota}(\overline{a}_2)]$. By construction, $\overline{\iota}$ is a Lie algebra homomorphism between the Lie algebras on $\overline{A}$ and $\tilde{A}$, and one can use the previous proposition to conclude that there will exist a Lie algebra on the affine space $A$.

Once we have chosen an affine frame on $A$, the bracket on $\tilde{A}$ is determined by the brackets of the associated basis elements. These must be of the form

$$[e_0, e_0] = 0 \qquad [e_0, e_\alpha] = C^\gamma_{0\alpha} e_\gamma \qquad [e_\alpha, e_\beta] = C^\gamma_{\alpha\beta} e_\gamma, \tag{5.37}$$

since all brackets must take values in the image of the map $\overline{\iota}$.

# Chapter 6

# *Lie algebroid structure on an affine bundle*

## 6.1 Affine bundles: generalities

The following definition is taken from [20].

**Definition 6.1.** *A fibre bundle $\pi : E \to M$ is an affine bundle modelled on the vector bundle $\overline{\pi} : \overline{E} \to M$ if*

1. *each fibre $\pi^{-1}(m) = E_m$ is an affine space modelled on the vector space $\overline{\pi}^{-1}(m) = \overline{E}_m$;*
2. *the standard fibre is $\mathbb{R}^N$, where $N$ is the fibre dimension of $\overline{\pi} : \overline{E} \to M$;*
3. *about each point of $M$ there is an affine local trivialisation, that is to say, a local trivialisation $\psi : \pi^{-1}(O) \to O \times \mathbb{R}^N$, with $O \subset M$ open, such that for each point $m \in O$ the map $\psi_m : E_m \to \mathbb{R}^N$ defined by $\psi_m = \Pi_2 \circ \psi|_{E_m}$ is an affine isomorphism from the affine space $E_m$ to the affine space $\mathbb{R}^N$, whose linear part is the linear isomorphism from $\overline{E}_m$ to $\mathbb{R}^N$ corresponding to a linear local trivialisation of the vector bundle $\overline{\pi} : \overline{E} \to M$.*

In the above definition, $\Pi_2$ stands for the projection onto the second factor of a product of two manifolds. For convenience, sections $\sigma$ of $\pi$ and sections $\overline{\sigma}$ of $\overline{\pi}$ will often be referred to as *affine sections* and *vector sections*, respectively.

Every fibre $E_m$ has a corresponding extended dual $E^\dagger_m$. The union of these spaces over all points $m \in M$ gives us a vector bundle $\pi^\dagger : E^\dagger \to M$. We also consider the bi-dual bundle $\tilde{\pi} \colon \tilde{E} \to M$, whose fibre at $m$ is $\tilde{E}_m = (E^\dagger_m)^*$. This bundle is sometimes called the *vector hull of* $\pi : E \to M$. Following Corollary 5.4, at every point $m$, we have the exact sequence of vector spaces

$$0 \longrightarrow \overline{E}_m \xrightarrow{\iota_m} \tilde{E}_m \xrightarrow{\jmath_m} \mathbb{R} \longrightarrow 0 \tag{6.1}$$

and therefore an exact sequence of vector bundles over $M$

$$0 \longrightarrow \overline{E} \overset{\bar{\iota}}{\longrightarrow} \tilde{E} \overset{\jmath}{\longrightarrow} M \times I\!R \longrightarrow 0. \tag{6.2}$$

On the other hand there is also the canonical immersion $\iota\colon E \to \tilde{E}$, so that $\iota(E_m) = \jmath^{-1}((m,1))$.

By taking sections, we have the exact sequence of real vector spaces (and $C^\infty(M)$-modules)

$$0 \longrightarrow \mathrm{Sec}(\overline{\pi}) \overset{\bar{\iota}}{\longrightarrow} \mathrm{Sec}(\tilde{\pi}) \overset{\jmath}{\longrightarrow} C^\infty(M) \longrightarrow 0.$$

Since $\overline{\pi}$ is the vector bundle associated to the affine bundle $\pi$, we have that $\mathrm{Sec}(\overline{\pi})$ is the (real) vector space associated to the affine space $\mathrm{Sec}(\pi)$. We therefore have an inclusion $\iota\colon \mathrm{Sec}(\pi) \to \mathrm{Sec}(\tilde{\pi})$, whereby we make no notational distinction between the bundle maps and the induced maps of sections (i.e. if $\sigma$ is a section and $r$ is a bundle map over the identity, we write $r(\sigma)$ instead of $r \circ \sigma$). It follows that if we fix a section $\sigma$ of $\pi$ then any section $\tilde{\zeta}$ of $\tilde{\pi}$ can be written as $\tilde{\zeta} = f\iota(\sigma) + \bar{\iota}(\overline{\eta})$, for some section $\overline{\eta}$ of $\overline{\pi}$ and where $f = \jmath(\tilde{\zeta})$.

Let us coordinatise $E$ as follows: $(x^I)_{1\leq I\leq m}$ denote coordinates on $M$; we further choose a local section $o$ of $\pi$ to play the role of zero section and a local basis $(\overline{e}_\alpha)_{1\leq\alpha\leq N}$ for $\mathrm{Sec}(\overline{\pi})$. Then, if $e$ is a point in the fibre $E_m$ over $m \in M$, it can be written in the form: $e = o(m) + y^\alpha \overline{e}_\alpha(m)$; $(x^I, y^\alpha)$ are (affine) coordinates of $e$ ($(x^I)$ being the coordinates of $m$). In an analogous way as it is done for an affine space, the affine frame $(o, \{e_\alpha\})$ gives rise to a basis $\{e^0, e^\alpha\}$ for $\mathrm{Sec}(\pi^\dagger)$ and its dual basis $\{e_0, e_\alpha\}$ for $\mathrm{Sec}(\tilde{\pi})$. Remark that, again, $e^0$ is global, since it corresponds on any fibre to the constant function 1. Any point $\tilde{e} \in \tilde{E}_m$ can thus be written as $\tilde{e} = y^0 e_0 + y^\alpha e_\alpha$ for some coordinate set $(y^0, y^\alpha)$. In particular, for the above elements $e$ and $\tilde{e}$, $\iota(e) = e_0 + y^\alpha e_\alpha$ and $\jmath(\tilde{e}) = y^0$.

We need to say a few words about the structure of the $C^\infty(E)$-module $\mathrm{Sec}(\pi^*\tilde{\pi})$. The injection of $E$ into $\tilde{E}$ provides a *canonical section* of $\pi^*\tilde{\pi}$, which will be denoted by $\mathcal{I}$. Furthermore, there exists a *canonical map* $\vartheta : \mathrm{Sec}(\pi^*\tilde{\pi}) \to \mathrm{Sec}(\pi^*\tilde{\pi})$, which can be discovered as follows. First, within a fixed fibre $\tilde{E}_m$, choosing an arbitrary $a \in E_m$, we get a map $\vartheta_a : \tilde{E}_m \to \tilde{E}_m, \tilde{e} \mapsto \tilde{e} - \jmath(\tilde{e})\iota(a)$, which actually takes values in $\overline{E}_m$, and therefore a map

$$\vartheta : \pi^*\tilde{E} \to \pi^*\overline{E} \subset \pi^*\tilde{E},\ (a, \tilde{e}) \mapsto (a, \tilde{e} - \jmath(\tilde{e})\iota(a)). \tag{6.3}$$

We will use the same notation for the extension of this map to $\mathrm{Sec}(\pi^*\tilde{\pi})$, i.e. for $\tilde{X} \in \mathrm{Sec}(\pi^*\tilde{\pi})$, $\vartheta(\tilde{X})(e) = \vartheta(\tilde{X}(e))$. It follows that every $\tilde{X} \in \mathrm{Sec}(\pi^*\tilde{\pi})$ can be written in the form

$$\tilde{X} = f_{\tilde{X}}\,\mathcal{I} + \vartheta(\tilde{X}), \qquad \text{with} \quad f_{\tilde{X}} \in C^\infty(E): \; f_{\tilde{X}}(e) = \jmath(\tilde{X}(e)). \tag{6.4}$$

Clearly, if $\tilde{X} = \vartheta(\tilde{X})$ in some open neighbourhood in $E$, it means that $\jmath(\tilde{X}(e)) = 0$, so that $\tilde{X}(e) \in \bar{\imath}(\overline{E})$ in that neighbourhood, and such a $\tilde{X}$ cannot at the same time be in the span of $\mathcal{I}$. We conclude that locally:

$$\mathrm{Sec}(\pi^*\tilde{\pi}) = \langle \mathcal{I} \rangle \oplus \mathrm{Sec}(\pi^*\overline{\pi}). \tag{6.5}$$

As a consequence, if $\{\overline{\sigma}_\alpha\}$ is a local basis for $\mathrm{Sec}(\overline{\pi})$, then $\{\mathcal{I}, \overline{\sigma}_\alpha\}$ is a local basis for $\mathrm{Sec}(\pi^*\tilde{\pi})$.

Since sections of $\tilde{\pi}$ can be regarded also as (basic) sections of $\pi^*\tilde{\pi}$, $\{e_0, e_\alpha\}$ can serve at the same time as local basis for $\mathrm{Sec}(\pi^*\tilde{\pi})$. Hence, every $\tilde{X} \in Sec(\pi^*\tilde{\pi})$ can be represented in the form $\tilde{X} = \tilde{X}^0(x,y)e_0 + \tilde{X}^\alpha(x,y)e_\alpha$. But more interestingly, with the use of the canonical section $\mathcal{I}$, we have

$$\mathcal{I} = e_0 + y^\alpha e_\alpha, \qquad \tilde{X} = \tilde{X}^0\mathcal{I} + (\tilde{X}^\alpha - \tilde{X}^0 y^\alpha)e_\alpha. \tag{6.6}$$

The coordinate expression of the canonical map is given by

$$\vartheta = (e^\alpha - y^\alpha e^0) \otimes e_\alpha.$$

On an affine bundle there exists a *vertical lift* $v : \pi^*\tilde{E} \to TE$. It is defined by the following sequence of natural constructions. Given a point $e \in E_m$ and a vector $\overline{e} \in \overline{E}_m$, we define the vector $v(e, \overline{e}) \in T_eE$ by its action on functions $f \in C^\infty(E)$:

$$v(e,\overline{e})f = \frac{d}{dt} f(e + t\overline{e})\big|_{t=0}. \tag{6.7}$$

Next, if $\tilde{e} \in \tilde{E}_m$ and $e \in E_m$, the vertical lift of $\tilde{e}$ to the point $e$ is defined as

$$v(e,\tilde{e}) = v\big(e, \vartheta_e(\tilde{e})\big). \tag{6.8}$$

For the special case that $\tilde{e} = \bar{\imath}(\overline{e})$, this is consistent with the preceding step: $v(e,\tilde{e}) = v(e,\overline{e})$. The final step of course is to extend this construction in the obvious way to an operation:

$$v : \mathrm{Sec}(\pi^*\tilde{\pi}) \to \mathcal{X}(E).$$

It follows in particular that

$$v(\mathcal{I}) = 0. \tag{6.9}$$

Given a vertical $Q$ in $TE$, we will use the notation $Q_v$ for the unique element in $\pi^*\overline{E}$ such that

$$v(Q_v) = Q. \tag{6.10}$$

## 6.2 Lie algebroid structure on an affine bundle

**Definition 6.2.** *A Lie algebroid structure on the affine bundle $\pi : E \to M$ consists of a Lie algebra structure on the (real) affine space of sections of $\pi$ together with an affine bundle map $\rho\colon E \to TM$ (the anchor), satisfying the following compatibility condition*

$$D_\sigma(f\overline{\zeta}) = \rho(\sigma)(f)\,\overline{\zeta} + fD_\sigma\overline{\zeta}, \tag{6.11}$$

*for every $\sigma \in \mathrm{Sec}(\pi)$, $\overline{\zeta} \in \mathrm{Sec}(\overline{\pi})$ and $f \in C^\infty(M)$, and where $D_\sigma$ is the action $\sigma \mapsto D_\sigma$ of $\mathrm{Sec}(\pi)$ on $\mathrm{Sec}(\overline{\pi})$.*

We will denote the underlying vector bundle map $\overline{E} \to TM$ of $\rho$ by $\overline{\rho}$. Note that we make no notational distinction between, on the one hand, the affine and linear anchor maps, regarded as maps between total spaces of bundles, and their action on sections of bundles on the other hand.

The compatibility condition ensures that the association $\sigma \mapsto D_\sigma$, which acts by derivations on the real Lie algebra $\mathrm{Sec}(\overline{\pi})$, also acts by derivations on the $C^\infty(M)$-module $\mathrm{Sec}(\overline{\pi})$. From now on we will also use the bracket notation $[\sigma, \overline{\zeta}] = D_\sigma\overline{\zeta}$.

Various bracket properties are gathered in the following table ($\lambda \in \mathbb{R}$, $f \in C^\infty(M)$):

| Lie bracket on $\mathrm{Sec}(\overline{\pi})$ | Action of $\mathrm{Sec}(\pi)$ on $\mathrm{Sec}(\overline{\pi})$ |
|---|---|
| (a1) $[\overline{\xi}, \overline{\sigma}] = -[\overline{\sigma}, \overline{\xi}]$ | |
| (a2) $[\overline{\xi}, \lambda\overline{\sigma}] = \lambda[\overline{\xi}, \overline{\sigma}]$ | (b2) $[\zeta, \lambda\overline{\sigma}] = \lambda[\zeta, \overline{\sigma}]$ |
| (a3) $[\overline{\xi}, \overline{\sigma}_1 + \overline{\sigma}_2] = [\overline{\xi}, \overline{\sigma}_1] + [\overline{\xi}, \overline{\sigma}_2]$ | (b3) $[\zeta, \overline{\sigma}_1 + \overline{\sigma}_2] = [\zeta, \overline{\sigma}_1] + [\zeta, \overline{\sigma}_2]$, |
| (a4) $[\overline{\xi} + \overline{\sigma}, \overline{\eta}] = [\overline{\xi}, \overline{\eta}] + [\overline{\sigma}, \overline{\eta}]$ | (b4) $[\zeta + \overline{\sigma}, \overline{\eta}] = [\zeta, \overline{\eta}] + [\overline{\sigma}, \overline{\eta}]$, |
| (a5) $[\overline{\xi}, [\overline{\sigma}, \overline{\eta}]] = [[\overline{\xi}, \overline{\sigma}], \overline{\eta}] + [\overline{\sigma}, [\overline{\xi}, \overline{\eta}]]$ | (b5) $[\zeta, [\overline{\sigma}, \overline{\eta}]] = [[\zeta, \overline{\sigma}], \overline{\eta}] + [\overline{\sigma}, [\zeta, \overline{\eta}]]$, |
| | (b6) $[\zeta, f\overline{\sigma}] = f\,[\zeta, \overline{\sigma}] + \rho(\zeta)(f)\overline{\sigma}$. |

There are a number of properties that can be derived from the defining relations in Definition 6.2. Substituting $\zeta + \overline{\sigma}$ for $\zeta$ in (b6), recalling that $\rho(\zeta + \overline{\xi}) = \rho(\zeta) + \overline{\rho}(\overline{\xi})$, it follows that

$$[\overline{\xi}, f\overline{\sigma}] = f\,[\overline{\xi}, \overline{\sigma}] + \overline{\rho}(\overline{\xi})(f)\overline{\sigma}. \tag{6.12}$$

Secondly, replacing $\overline{\eta}$ by $f\overline{\eta}$ in (b5) and making use of (b6) and (6.12), one obtains the additional compatibility property

$$[\rho(\zeta), \overline{\rho}(\overline{\sigma})] = \overline{\rho}([\zeta, \overline{\sigma}]), \tag{6.13}$$

from which it further follows that

$$[\overline{\rho}(\overline{\xi}), \overline{\rho}(\overline{\sigma})] = \overline{\rho}\left([\overline{\xi}, \overline{\sigma}]\right). \tag{6.14}$$

This means that the linear anchor map $\overline{\rho} : \overline{E} \to TM$ defines a Lie algebra homomorphism from $\mathrm{Sec}(\overline{\pi})$ into the real Lie algebra of vector fields on $M$, and that we have a classical Lie algebroid structure on the vector bundle $\overline{\pi} : \overline{E} \to M$.

**Remark**: For an alternative and equivalent definition of an affine Lie algebroid, we could impose first the Lie algebra structure of the bracket on $\mathrm{Sec}(\overline{\pi})$, together with the compatibility condition (6.12) for the anchor map $\overline{\rho}$, and subsequently require that the properties (b2-b6) in the right column of the tabular hold true for *at least one* $\zeta \in \mathrm{Sec}(\pi)$ and for an affine map $\rho : E \to TM$ whose linear part is $\overline{\rho}$. It then follows that such properties hold for all $\zeta$.

By means of Definition 5.8 we can further extend the bracket operation to $\mathrm{Sec}(\pi)$. Next to the properties of Proposition 5.9, we obtain now

**Proposition 6.3.** *The bracket* $[\cdot,\cdot] : \mathrm{Sec}(\pi) \times \mathrm{Sec}(\pi) \to \mathrm{Sec}(\overline{\pi})$*, has the property:*

$$\overline{\rho}([\zeta_1, \zeta_2]) = [\rho(\zeta_1), \rho(\zeta_2)]. \tag{6.15}$$

Proof: For every two sections $\zeta_1, \zeta_2 \in \mathrm{Sec}(\pi)$, putting $\overline{\zeta}_{12} = \zeta_2 - \zeta_1$, we have by definition $[\zeta_1, \zeta_2] = [\zeta_1, \overline{\zeta}_{12}]$. The compatibility property (6.15) now easily follows from (6.13). □

## 6.3 Local description of a Lie algebroid on an affine bundle

Let us now try to understand what an affine Lie algebroid structure means in coordinates. We have

$$[\overline{e}_\alpha, \overline{e}_\beta] = C^\gamma_{\alpha\beta}(x)\overline{e}_\gamma, \qquad [o, \overline{e}_\alpha] = C^\beta_{0\alpha}(x)\overline{e}_\beta, \tag{6.16}$$

for some *structure functions* $C^\gamma_{\alpha\beta} = -C^\gamma_{\beta\alpha}$ and $C^\beta_{0\alpha}$ on $M$. The affine map $\rho$ and its linear part $\overline{\rho}$ are fully determined by

$$\rho(o) = \rho^I_0(x)\frac{\partial}{\partial x^I}, \qquad \overline{\rho}(\overline{e}_\alpha) = \rho^I_\alpha(x)\frac{\partial}{\partial x^I}. \tag{6.17}$$

The further characterisation of the Lie algebroid structure now has the following coordinate translation. The derivation property (b5) and the resulting Jacobi identity for the bracket on $\mathrm{Sec}(\overline{\pi})$ mean that we have:

$$\rho^I_0\frac{\partial C^\mu_{\alpha\beta}}{\partial x^I} + C^\gamma_{\alpha\beta}C^\mu_{0\gamma} = C^\mu_{\alpha\gamma}C^\gamma_{0\beta} - C^\mu_{\beta\gamma}C^\gamma_{0\alpha} + \rho^I_\alpha\frac{\partial C^\mu_{0\beta}}{\partial x^I} - \rho^I_\beta\frac{\partial C^\mu_{0\alpha}}{\partial x^I}, \tag{6.18}$$

$$\sum_{\alpha,\beta,\gamma}\left(\rho^I_\alpha\frac{\partial C^\mu_{\beta\gamma}}{\partial x^I} + C^\mu_{\alpha\nu}C^\nu_{\beta\gamma}\right) = 0, \tag{6.19}$$

where the summation this time refers to a cyclic sum over $\alpha, \beta, \gamma$ and also the compatibility conditions (b6), (6.12) have been invoked. Finally, the properties (6.13) and (6.14), for which it is sufficient to express that $[\rho(o), \overline{\rho}(\overline{e}_\alpha)] = \overline{\rho}([o, \overline{e}_\alpha])$ and $[\overline{\rho}(\overline{e}_\alpha), \overline{\rho}(\overline{e}_\beta)] = \overline{\rho}([\overline{e}_\alpha, \overline{e}_\beta])$, require that

$$\rho^I_0\frac{\partial \rho^J_\beta}{\partial x^I} - \rho^I_\beta\frac{\partial \rho^J_0}{\partial x^I} = C^\alpha_{0\beta}\rho^J_\alpha, \tag{6.20}$$

$$\rho^I_\alpha\frac{\partial \rho^J_\beta}{\partial x^I} - \rho^I_\beta\frac{\partial \rho^J_\alpha}{\partial x^I} = C^\gamma_{\alpha\beta}\rho^J_\gamma. \tag{6.21}$$

It is now easy to verify that the model for a Lie algebroid structure on an affine bundle indeed satisfies all requirements that we have encountered in Chapter 4, in the context of Lagrangian equations of type (4.11). Remember that the set-up of the time-dependent case in Section 4.2 was somehow more restricted than the one we have used so far. First, we assumed the base manifold $M$ to be fibred over $I\!R$ and we use bundle coordinates $(t, x^i)$ on $M$. The affine bundle $J^1M \to M$ can then be regarded as a subbundle of $TM$. Its injection $\iota_{J^1M} : J^1M \to TM$ is an affine bundle map that is modelled on the injection $\iota_{VM} : VM \to TM$. Moreover, we supposed the anchor map to take values in $J^1M$ (technically $\rho = \iota_{J^1M} \circ \lambda$ for some affine map $\lambda : E \to J^1M$ and $\overline{\rho} = \iota_{VM} \circ \overline{\lambda}$ for the underlying linear map $\overline{\lambda} : \overline{E} \to VM$). In coordinates, we can think of $t$ as the zeroth coordinate and thus

$$\rho_0^0 = 1, \quad \rho_\alpha^0 = 0, \quad \rho_0^i = \lambda^i \quad \text{and} \quad \rho_\alpha^i = \lambda_\alpha^i$$

for some functions $\lambda^i, \lambda_\alpha^i \in C^\infty(M)$. Notice that such an extra fibration is not generally available, not even locally. For instance, if we take any affine bundle $\pi \colon E \to M$ with the trivial Lie algebroid structure (null bracket and anchor) then there is no fibration over $\mathbb{R}$ such that the image of $\rho$ is in $J^1M$, since the vectors in $\iota(J^1M)$ are non-zero. Coming back to the Lie algebroids, we find an identity for the relations (6.20) and (6.21) for $J = 0$ and the relations (4.12) and (4.13) we were looking for, for $J = j$.

It is of some interest to look at the way the various structure and anchor map functions transform under coordinate transformations. There are two distinct levels in making a change of coordinates on $E$, which we will describe separately. Firstly, we could choose a different (local) zero section $o'$ and a different local basis $\overline{e}'_\beta$ for $\mathrm{Sec}(\overline{\pi})$: say that $\overline{e}_\alpha = A_\alpha^\beta \overline{e}'_\beta$ and $o = o' + B^\alpha \overline{e}'_\alpha$. This amounts to making an affine change of coordinates in the fibres of the form: $y'^\alpha = A_\beta^\alpha(x) y^\beta + B^\alpha(x)$. Putting $[\overline{e}'_\alpha, \overline{e}'_\beta] = C'^\gamma_{\alpha\beta} \overline{e}'_\gamma$, $[o', \overline{e}'_\alpha] = C'^\beta_{0\alpha} \overline{e}'_\beta$, and also $\rho(o') = \rho'^I_0 \frac{\partial}{\partial x^I}$, $\overline{\rho}(\overline{e}'_\alpha) = \rho'^J_\alpha \frac{\partial}{\partial x^J}$, one can verify that the following transformation rules apply:

$$\rho_\alpha^I = A_\alpha^\beta \rho'^I_\beta, \qquad \rho'^I_0 = \rho_0^I - B^\alpha \rho'^I_\alpha,$$

and further

$$\begin{aligned} C_{\alpha\beta}^\gamma A_\gamma^\mu &= C'^\mu_{\gamma\nu} A_\alpha^\gamma A_\beta^\nu + \rho_\alpha^I \frac{\partial A_\beta^\mu}{\partial x^I} - \rho_\beta^I \frac{\partial A_\alpha^\mu}{\partial x^I}, \\ C_{0\beta}^\gamma A_\gamma^\alpha &= C'^\alpha_{0\mu} A_\beta^\mu + C'^\alpha_{\gamma\mu} B^\gamma A_\beta^\mu + \rho_0^I \frac{\partial A_\beta^\alpha}{\partial x^I} - \rho_\beta^I \frac{\partial B^\alpha}{\partial x^I}. \end{aligned}$$

At a different level, one can make a change of coordinates on $M$, of the form: $\hat{x}^I = \hat{x}^I(x)$. This has an effect on the anchor map functions of the form:

$$\hat{\rho}_\alpha^J = \rho_\alpha^I \frac{\partial \hat{x}^J}{\partial x^I}, \qquad \hat{\rho}_0^J = \rho_0^I \frac{\partial \hat{x}^J}{\partial x^I}.$$

A general change of adapted coordinates is of course a composition of the two steps described above.

## 6.4 Lie algebroids on the bi-dual

The following result shows that one can alternatively define an affine Lie algebroid structure on $E$ as a vector Lie algebroid structure $([\,,\,],\tilde{\rho})$ on $\tilde{E}$ such that the bracket of two sections in the image of $\iota$ belongs to the image of $\bar{\iota}$.

**Proposition 6.4.** *A Lie algebroid structure on the vector bundle $\tilde{\pi}\colon \tilde{E} \to M$ which is such that the bracket of sections in the image of $\iota$ lies in the image of $\bar{\iota}$ induces a Lie algebroid structure on the affine bundle $\pi\colon E \to M$, whereby the brackets and maps are determined by the following relations:*

$$\bar{\iota}([\overline{\eta}_1,\overline{\eta}_2]) = [\bar{\iota}(\overline{\eta}_1),\bar{\iota}(\overline{\eta}_2)] \tag{6.22}$$
$$\bar{\iota}([\sigma,\overline{\eta}]) = [\iota(\sigma),\bar{\iota}(\overline{\eta})] \tag{6.23}$$
$$\rho(\sigma) = \tilde{\rho}(\iota(\sigma)). \tag{6.24}$$

*Conversely, a Lie algebroid structure on the affine bundle $\pi\colon E \to M$ extends to a Lie algebroid structure on the vector bundle $\tilde{\pi}\colon \tilde{E} \to M$ such that the bracket of sections in the image of $\iota$ is in the image of $\bar{\iota}$. If we fix a section $\sigma$ of $\pi$ and write sections $\tilde{\zeta}$ of $\tilde{\pi}$ (locally) in the form $\tilde{\zeta} = f\iota(\sigma) + \bar{\iota}(\overline{\eta})$ then the anchor and the bracket are given by*

$$\tilde{\rho}(\tilde{\zeta}) = f\rho(\sigma) + \overline{\rho}(\overline{\eta}) \tag{6.25}$$
$$\begin{aligned}[\tilde{\zeta}_1,\tilde{\zeta}_2] = \Big(\tilde{\rho}(\tilde{\zeta}_1)(f_2) - \tilde{\rho}(\tilde{\zeta}_2)(f_1)\Big)\iota(\sigma)&\\ +\bar{\iota}\Big([\overline{\eta}_1,\overline{\eta_2}] + f_1[\sigma,\overline{\eta}_2] - f_2[\sigma,\overline{\eta}_1]\Big)&.\end{aligned} \tag{6.26}$$

Proof: The map $\bar{\iota}\colon \overline{E} \to \tilde{E}$ is a morphism of Lie algebroids, since we have

$$[\bar{\iota}(\overline{\eta}_1),\bar{\iota}(\overline{\eta}_2)] = \bar{\iota}([\overline{\eta}_1,\overline{\eta}_2]) \quad \text{and} \quad \tilde{\rho}\circ\bar{\iota} = \overline{\rho},$$

where $\overline{\rho}$ is the linear part of $\rho$. The verification of the other statements is straightforward but rather lengthy. We limit ourselves to checking that the compatibility conditions between brackets and anchors are satisfied. For the first part, we find

$$\begin{aligned}\bar{\iota}\Big([\sigma,f\overline{\eta}]\Big) &= [\iota(\sigma),\bar{\iota}(f\overline{\eta})] = [\iota(\sigma),f\bar{\iota}(\overline{\eta})]\\ &= \tilde{\rho}(\iota(\sigma))(f)\bar{\iota}(\overline{\eta}) + f[\iota(\sigma),\bar{\iota}(\overline{\eta})] = \bar{\iota}\Big(\rho(\sigma)(f)\overline{\eta} + f[\sigma,\overline{\eta}]\Big),\end{aligned}$$

form which it follows that $[\sigma, f\overline{\eta}] = \rho(\sigma)(f)\overline{\eta} + f[\sigma, \overline{\eta}]$.

For the converse, observe that

$$\begin{aligned}
[\tilde{\zeta}_1, f\tilde{\zeta}_2] - f[\tilde{\zeta}_1, \tilde{\zeta}_2] &= f_2\tilde{\rho}(\tilde{\zeta}_1)(f)\iota(\sigma) + \bar{\iota}\Big(f_1\rho(\sigma)(f)\overline{\eta_2} + \overline{\rho}(\overline{\eta_1})(f)\overline{\eta}_2\Big) \\
&= \tilde{\rho}(\tilde{\zeta}_1)(f)f_2\iota(\sigma) + (f_1\rho(\sigma) + \overline{\rho}(\overline{\eta_1}))(f)\bar{\iota}(\overline{\eta}_2) \\
&= \tilde{\rho}(\tilde{\zeta}_1)(f)\tilde{\zeta}_2,
\end{aligned}$$

which is the required compatibility condition. □

Remark that, unlike $\bar{\iota}$, the map $\jmath\colon \tilde{E} \to M \times I\!R$ is NOT a morphism of Lie algebroids, since we have that $\jmath([\tilde{\zeta}_1, \tilde{\zeta}_2]) = \tilde{\rho}(\tilde{\zeta}_1)f_2 - \tilde{\rho}(\tilde{\zeta}_2)f_1$, while $[\jmath(\tilde{\zeta}_1), \jmath(\tilde{\zeta}_2)] = 0$ since the fibres of $M \times I\!R$ are 1-dimensional.

Let us come to coordinate expressions. In coordinates, starting from an affine frame $(o, \{\overline{e}_\alpha\})$ for $\mathrm{Sec}(\pi)$ and a corresponding local basis $\{e_0, e_\alpha\}$ of sections of $\tilde{\pi}$ (see Section 6.1), we have

$$\tilde{\rho}(y^0 e_0 + y^\alpha e_\alpha) = (\rho^I_0 y^0 + \rho^I_\alpha y^\alpha)\frac{\partial}{\partial x^I}\,, \tag{6.27}$$

and the bracket is determined by

$$[e_0, e_0] = 0 \qquad [e_0, e_\beta] = C^\gamma_{0\beta} e_\gamma \qquad [e_\alpha, e_\beta] = C^\gamma_{\alpha\beta} e_\gamma. \tag{6.28}$$

## 6.5 Exterior differential

We first recall some features of the by now standard theory of Lie algebroids on a vector bundle (see [46]). It is well-known that the definition of an exterior derivative on forms involves the Lie algebroid bracket and the anchor map. It then turns out that the Jacobi identity of the Lie algebroid bracket and the compatibility with the bracket of vector fields via the anchor map are exactly the conditions for this exterior derivative to have the co-boundary property $d^2 = 0$ (see also [33, 34, 48]). In our opinion, such a feature in itself gives a strong indication that the generalisation from Lie algebra to Lie algebroid is indeed a meaningful step. We can feel confident that such a supportive property will equally hold for our extension to Lie algebroids on affine bundles, because of the link with a vector Lie algebroid established by Proposition 6.4. Again, however, it is of interest to see how one can introduce an exterior derivative for forms which act on sections of an affine bundle $\pi$, without recourse to the vector bundle encompassing $\pi$. So, we propose to approach the subject first as follows:

1. Forms on $\pi$ are introduced in the spirit of the algebraic Definition 5.5.
2. Then we show how to construct an exterior derivation $d$ when a bracket is available which satisfies (a1-a5), (b2-b4) and (b6), but not necessarily (b5).
3. We finally show that the requirement that the bracket also satisfies (b5) is equivalent with the condition that the exterior derivative satisfies the co-boundary condition $d^2 = 0$.

The manifold $\Omega^k(\pi^\dagger) = \bigcup_{m\in M} \Omega^k(E^\dagger_m)$ is the total manifold of a bundle over $M$. Its sections $\omega : M \to \Omega^k(\pi^\dagger), m \mapsto \omega_m \in \Omega^k(E^\dagger_m)$ are called *$k$-forms on the affine bundle $\pi$*. We can also give a definition of these forms in almost exactly the same way as we defined forms on an affine space:

**Definition 6.5.** *A $k$-form on an affine bundle $\pi : E \to M$ is a map $\omega\colon \mathrm{Sec}(\pi)\times\cdots\times\mathrm{Sec}(\pi) \to C^\infty(M)$ for which there exists a $k$-form $\overline{\omega}$ on the associated vector bundle $\overline{\pi}$, and a map $\omega_0\colon \mathrm{Sec}(\pi)\times\mathrm{Sec}(\overline{\pi})\times\cdots\times\mathrm{Sec}(\overline{\pi}) \to C^\infty(M)$ with the following properties*

1. *$\omega_0$ is skew-symmetric and $C^\infty(M)$-linear in its $k-1$ vector arguments.*
2. *For every $\zeta \in \mathrm{Sec}(\pi)$ and for every $\overline{\sigma}_1, \overline{\sigma}_2, \dots, \overline{\sigma}_k \in \mathrm{Sec}(\overline{\pi})$, we have*

$$\omega_0(\zeta+\overline{\sigma}_1, \overline{\sigma}_2, \dots, \overline{\sigma}_k) = \omega_0(\zeta, \overline{\sigma}_2, \dots, \overline{\sigma}_k) + \overline{\omega}(\overline{\sigma}_1, \overline{\sigma}_2, \dots, \overline{\sigma}_k). \quad (6.29)$$

3. *For every $\zeta_1, \dots, \zeta_k \in \mathrm{Sec}(\pi)$, if we choose an arbitrary $\zeta_0 \in \mathrm{Sec}(\pi)$ and put $\zeta_i = \zeta_0 + \overline{\zeta}_{0i}$, we have*

$$\omega(\zeta_1, \dots, \zeta_k) = \sum_{i=1}^{k} (-1)^{i-1} \omega_0(\zeta_0, \overline{\zeta}_{01}, \dots, \widehat{\overline{\zeta}_{0i}}, \dots, \overline{\zeta}_{0k}) + \overline{\omega}(\overline{\zeta}_{01}, \overline{\zeta}_{02}, \dots, \overline{\zeta}_{0k}). \quad (6.30)$$

The set of forms on $\pi$, which we will denote by $\bigwedge(\pi^\dagger)$, is a module over the ring $C^\infty(M)$, which also constitutes the set of 0-forms.

Before arriving at the development of an exterior calculus on forms, we will recall a few generalities about derivations. Derivations on $\bigwedge(\pi^\dagger)$ are defined in the usual way. Following the standard work of Frölicher and Nijenhuis [32], one easily shows that derivations are local operators and that they are completely determined by their action on functions and 1-forms. The commutator of two derivations $D_i$, of degree $r_i$ say, is again a derivation, of degree $r_1 + r_2$, defined by

$$[D_1, D_2] = D_1 \circ D_2 - (-1)^{r_1 r_2} D_2 \circ D_1. \tag{6.31}$$

Perhaps the simplest type of derivation is contraction with a section.

**Definition 6.6.** *For $\omega \in \bigwedge^k(\pi^\dagger)$ and $\zeta \in \mathrm{Sec}(\pi)$, $i_\zeta\omega \in \bigwedge^{k-1}(\pi^\dagger)$ is defined by*

$$i_\zeta\omega(\zeta_1, \dots, \zeta_{k-1}) = \omega(\zeta, \zeta_1, \dots \zeta_{k-1}). \tag{6.32}$$

The proof that this is a derivation of degree $-1$ is standard and does not depend on the peculiarities of our present theory. But perhaps we have to convince ourselves in the first place that $i_\zeta\omega$ is indeed a form in the sense of Definition 6.5.

**Proposition 6.7.** *$i_\zeta\omega$ is a $(k-1)$-form which, in the sense of the general defining relation (6.30), is determined by an operator $(i_\zeta\omega)_0$ and a $k$-form $\overline{i_\zeta\omega}$ on $\mathrm{Sec}(\overline{\pi})$, defined as follows: for all $\overline{\sigma}_i \in \mathrm{Sec}(\overline{\pi})$, $\zeta_0 \in \mathrm{Sec}(\pi)$,*

$$(i_\zeta\omega)_0(\zeta_0, \overline{\sigma}_2, \dots \overline{\sigma}_{k-1}) = -\omega_0(\zeta_0, \overline{\zeta}, \overline{\sigma}_2, \dots \overline{\sigma}_{k-1}), \tag{6.33}$$

*where $\overline{\zeta} = \zeta - \zeta_0$ and*

$$\overline{i_\zeta\omega}(\overline{\sigma}_2, \dots, \overline{\sigma}_k) = \omega_0(\zeta, \overline{\sigma}_2, \dots, \overline{\sigma}_k). \tag{6.34}$$

*We further have the property (with $\overline{\zeta}_{01} = \overline{\zeta}$ and $\overline{\zeta}_{0i} = \zeta_i - \zeta_0$ for $i \in \{2 \dots k\}$):*

$$i_\zeta\omega(\zeta_2, \dots, \zeta_k) = \sum_{i=1}^{k} (-1)^{i-1} (i_\zeta\omega_0) (\overline{\zeta}_{01}, \dots, \widehat{\overline{\zeta}_{0i}}, \dots, \overline{\zeta}_{0k}). \tag{6.35}$$

Proof: A direct computation, using (6.32) and (6.30), gives

$$\begin{aligned}
i_\zeta\omega(\zeta_2, \dots, \zeta_k) &= \omega(\zeta, \zeta_2, \dots, \zeta_k) \\
&= \omega_0(\zeta_0, \overline{\zeta}_{02}, \dots, \overline{\zeta}_{0k}) + \sum_{j=1}^{k-1} (-1)^j \omega_0(\zeta_0, \overline{\zeta}, \overline{\zeta}_{02}, \dots, \widehat{\overline{\zeta}_{0j}}, \dots, \overline{\zeta}_{0k}) \\
&\quad + \overline{\omega}(\overline{\zeta}, \overline{\zeta}_{02}, \dots, \overline{\zeta}_{0k}) \\
&= \sum_{j=1}^{k-1} (-1)^j \omega_0(\zeta_0, \overline{\zeta}, \overline{\zeta}_{02}, \dots, \widehat{\overline{\zeta}_{0j}}, \dots, \overline{\zeta}_{0k}) + \omega_0(\zeta, \overline{\zeta}_{02}, \dots, \overline{\zeta}_{0k}),
\end{aligned}$$

from which we are led to introduce $(i_\zeta\omega)_0$ and $\overline{i_\zeta\omega}$ as in (6.33) and (6.34). It is then straightforward to verify that these two operators are linked by a property of type (6.29), so the first statement follows. Observe that, with an obvious meaning for contraction of the operator $\omega_0$ with $\zeta$, we can write: $\overline{i_\zeta\omega} = i_\zeta\omega_0$. The somewhat peculiar feature of the additional property is that $i_\zeta\omega$ can be completely computed from $i_\zeta\omega_0$. To prove this we again start from (6.30) to write (with $\zeta_1 = \zeta$)

$$i_\zeta\omega(\zeta_2,\ldots,\zeta_k) = \sum_{i=1}^{k}(-1)^{i-1}\omega_0(\zeta_0,\overline{\zeta}_{01},\ldots,\widehat{\overline{\zeta}_{0i}},\ldots,\overline{\zeta}_{0k}) + \overline{\omega}(\overline{\zeta}_{01},\ldots,\overline{\zeta}_{0k}).$$

This time, we substitute $\zeta_1 - \overline{\zeta}_{01}$ for $\zeta_0$ and observe that, in view of (6.29), the second part of the sum involving $\omega_0$ then precisely cancels the last term. □

Before we can arrive now at the definition of an exterior derivative operator, we need to give a meaning also to the value of a $k$-form $\omega$ when, say, its first argument is taken to be a vector section.

**Definition 6.8.** *If $\omega$ is a $k$-form on $\pi$, then for $\overline{\sigma} \in \mathrm{Sec}(\overline{\pi})$ and $\zeta_i \in \mathrm{Sec}(\pi)$, we put*

$$\omega(\overline{\sigma},\zeta_2,\ldots,\zeta_k) = \omega(\zeta_1+\overline{\sigma},\zeta_2,\ldots,\zeta_k) - \omega(\zeta_1,\zeta_2,\ldots,\zeta_k), \tag{6.36}$$

*where $\zeta_1$ is chosen arbitrarily.*

For this to make sense of course, we need to be sure that the result does not depend on the choice of $\zeta_1$. Now, if we evaluate the right-hand side of the defining relation by using (6.30), we obtain

$$\omega(\overline{\sigma},\zeta_2,\ldots,\zeta_k) = \overline{\omega}(\overline{\sigma},\overline{\zeta}_{02},\ldots,\overline{\zeta}_{0k}) + \sum_{i=2}^{k}(-1)^{i-1}\omega_0(\zeta_0,\overline{\sigma},\overline{\zeta}_{02},\ldots,\widehat{\overline{\zeta}_{0i}},\ldots,\overline{\zeta}_{0k}). \tag{6.37}$$

The right-hand side of this explicit expression makes no mentioning of $\zeta_1$ anymore. It might seem at first sight that we have shifted the problem, because it does depend on the reference section $\zeta_0$. However, we have argued before that (6.30) does not depend on the choice of such a reference section, whence our newly defined concept makes sense.

The explicit formula (6.37) further shows that $i_{\overline{\sigma}}\omega$ is well defined as a $(k-1)$-form, in the sense of Definition 6.5. The first term on the right identifies its associated form on $\mathrm{Sec}(\overline{\pi})$, whereas the second term, upon swapping the first two arguments, reveals that $(i_{\overline{\sigma}}\omega)_0 = i_{\overline{\sigma}}\omega_0$. As for local computations, it follows from the definition (6.36) that $e^0(\overline{\sigma}) = 0$, whereas the $\overline{e}^\alpha$ act on $\overline{\sigma}$ simply as duals of $\mathrm{Sec}(\overline{\pi})$.

**Definition 6.9.** *The exterior derivative of $\omega$, denoted by $d\omega$ is defined by*

$$\begin{aligned} d\omega(\zeta_1,\dots,\zeta_{k+1}) &= \sum_{i=1}^{k+1}(-1)^{i-1}\rho(\zeta_i)\Big(\omega(\zeta_1,\dots,\widehat{\zeta_i},\dots,\zeta_{k+1})\Big) \\ &\quad + \sum_{1\le i<j\le k+1}(-1)^{i+j}\omega([\zeta_i,\zeta_j],\zeta_1,\dots,\widehat{\zeta_i},\dots,\widehat{\zeta_j},\dots,\zeta_{k+1}). \end{aligned} \tag{6.38}$$

Note first that we are making use of Definition 6.8 in the second term on the right, because the bracket of two affine sections is a vector section. It is fairly obvious that $d\omega$ is skew-symmetric in all its arguments. To justify the definition, however, we should be able to identify an operator $(d\omega)_0$ and a related $(k+1)$-form $\overline{d\omega}$ on $\mathrm{Sec}(\overline{\pi})$, such that $d\omega$ satisfies an expression of type (6.30). We of course have an exterior derivative at our disposal for the $k$-form $\overline{\omega}$ on $\mathrm{Sec}(\overline{\pi})$ which we also denote by $d$. We know that $d\overline{\omega}$ has a property of type (6.38) (or can be defined that way), with vector sections replacing affine sections and $\overline{\rho}$ as anchor map instead of $\rho$.

**Definition 6.10.** *For $\omega_0 : \mathrm{Sec}(\pi)\times\mathrm{Sec}(\overline{\pi})\times\cdots\times\mathrm{Sec}(\overline{\pi}) \to C^\infty(M)$, we define $d\omega_0$, an operator of the same type, but depending on one more vector section, by*

$$\begin{aligned} d\omega_0(\zeta,\overline{\sigma}_2,\dots,\overline{\sigma}_{k+1}) &= \rho(\zeta)\Big(\overline{\omega}(\overline{\sigma}_2,\dots,\overline{\sigma}_{k+1})\Big) \\ &\quad + \sum_{i=2}^{k+1}(-1)^{i-1}\overline{\rho}(\overline{\sigma}_i)\Big(\omega_0(\zeta,\overline{\sigma}_2,\dots,\widehat{\overline{\sigma}_i},\dots,\overline{\sigma}_{k+1})\Big) \\ &\quad + \sum_{j=2}^{k+1}(-1)^{j+1}\overline{\omega}([\zeta,\overline{\sigma}_j],\overline{\sigma}_2,\dots,\widehat{\overline{\sigma}_j},\dots,\overline{\sigma}_{k+1}) \\ &\quad - \sum_{2\le i<j\le k+1}(-1)^{i+j}\omega_0(\zeta,[\overline{\sigma}_i,\overline{\sigma}_j],\overline{\sigma}_2,\dots,\widehat{\overline{\sigma}_i},\dots,\widehat{\overline{\sigma}_j},\dots,\overline{\sigma}_{k+1}). \end{aligned} \tag{6.39}$$

This expression may look rather exotic at first, but it is obtained by formally copying the definition (6.38) and writing in that process either $\rho$ or $\overline{\rho}$, and either $\omega_0$ or $\overline{\omega}$, in such a way that every term on the right-hand side has a proper meaning. There are two important observations to be made here. First of all, the required linearity of $d\omega_0$ in its vector arguments relies on the properties (b6) and (6.12) of our Lie algebroid bracket. Secondly, replacing the affine section $\zeta$ in the definition by $\zeta + \overline{\sigma}$, we find:

$$d\omega_0(\zeta+\overline{\sigma}, \overline{\sigma}_2, \ldots, \overline{\sigma}_{k+1}) = d\omega_0(\zeta, \overline{\sigma}_2, \ldots, \overline{\sigma}_{k+1}) + d\overline{\omega}(\overline{\sigma}, \overline{\sigma}_2, \ldots, \overline{\sigma}_{k+1}). \quad (6.40)$$

We thus know what to expect for the decomposition (6.30) of $d\omega$ and this is confirmed by the following result.

**Proposition 6.11.** *We have* $(d\omega)_0 = d\omega_0$ *and* $\overline{d\omega} = d\overline{\omega}$.

PROOF: We start with the defining relation (6.38) of the exterior derivative, in which we make use of the decomposition (6.30) in the first term and (6.37) in the second. We first obtain,

$$\begin{aligned}
& d\omega(\zeta_1, \ldots, \zeta_{k+1}) \\
&= \sum_{i=1}^{k+1} (-1)^{i-1} (\rho(\zeta_0) + \overline{\rho}(\overline{\zeta}_{0i})) \Big( \sum_{j=1}^{i-1} (-1)^{j-1} \omega_0(\zeta_0, \overline{\zeta}_{01}, \ldots, \widehat{\overline{\zeta}_{0j}}, \ldots, \widehat{\overline{\zeta}_{0i}}, \ldots, \overline{\zeta}_{0k+1}) \\
&+ \sum_{j=i+1}^{k+1} (-1)^{j} \omega_0(\zeta_0, \overline{\zeta}_{01}, \ldots, \widehat{\overline{\zeta}_{0i}}, \ldots, \widehat{\overline{\zeta}_{0j}}, \ldots, \overline{\zeta}_{0k+1}) + \overline{\omega}(\overline{\zeta}_{01}, \ldots, \widehat{\overline{\zeta}_{0i}}, \ldots, \overline{\zeta}_{0k+1}) \Big) \\
&+ \sum_{1 \le i < j \le k+1} (-1)^{i+j} \Big( \overline{\omega}([\zeta_i, \zeta_j], \overline{\zeta}_{01}, \ldots, \widehat{\overline{\zeta}_{0i}}, \ldots, \widehat{\overline{\zeta}_{0j}}, \ldots, \overline{\zeta}_{0k+1}) \\
&+ \sum_{l=1}^{i-1} (-1)^{l} \omega_0(\zeta_0, [\zeta_i, \zeta_j], \overline{\zeta}_{01}, \ldots, \widehat{\overline{\zeta}_{0l}}, \ldots, \widehat{\overline{\zeta}}_i, \ldots, \widehat{\overline{\zeta}_{0j}}, \ldots, \overline{\zeta}_{0k+1}) \\
&+ \sum_{l=i+1}^{j-1} (-1)^{l-1} \omega_0(\zeta_0, [\zeta_i, \zeta_j], \overline{\zeta}_1, \ldots, \widehat{\overline{\zeta}_{0i}}, \ldots, \widehat{\overline{\zeta}_{0l}}, \ldots, \widehat{\overline{\zeta}_{0j}}, \ldots, \overline{\zeta}_{0k+1}) \\
&+ \sum_{l=j+1}^{k+1} (-1)^{l} \omega_0(\zeta_0, [\zeta_i, \zeta_j], \overline{\zeta}_{01}, \ldots, \widehat{\overline{\zeta}_{0i}}, \ldots, \widehat{\overline{\zeta}_{0j}}, \ldots, \widehat{\overline{\zeta}_{0l}}, \ldots, \overline{\zeta}_{0k+1}) \Big),
\end{aligned}$$

and now perform a number of manipulations on multiple sums. Interchanging the order of summation in the first line of the right-hand side, we have $\sum_{i=1}^{k+1} \sum_{j=1}^{i-1} = \sum_{j=1}^{k} \sum_{i=j+1}^{k+1}$. Interchanging subsequently the names of the indices $i$ and $j$ in the first line, the term involving $\rho(\zeta_0)$ cancels the corresponding one appearing in the second line. The last three lines involve triple sums, which can be rearranged as follows. The first triple sum, with suitable interchanges of the order of summation, becomes:

$$\sum_{i=1}^{k}\sum_{j=i+1}^{k+1}\sum_{l=1}^{i-1}=\sum_{i=1}^{k}\sum_{l=1}^{i-1}\sum_{j=i+1}^{k+1}=\sum_{l=1}^{k-1}\sum_{i=l+1}^{k}\sum_{j=i+1}^{k+1}=\sum_{1\le l<i<j\le k+1}.$$

For the second one, we have

$$\sum_{i=1}^{k}\sum_{j=i+1}^{k+1}\sum_{l=i+1}^{j-1}=\sum_{i=1}^{k}\sum_{l=i+1}^{k}\sum_{j=l+1}^{k+1}=\sum_{1\le i<l<j\le k+1}.$$

The last one can directly be written as $\sum_{1\le i<j<l\le k+1}$. Changing names of indices to make all triple sums look alike, we thus far arrive at the result:

$$\begin{aligned}
&d\omega(\zeta_1,\dots,\zeta_{k+1})\\
&\quad=\sum_{1\le i<j\le k+1}(-1)^{i+j}\overline{\rho}(\overline{\zeta}_{0j}-\overline{\zeta}_{0i})\Big(\omega_0(\zeta_0,\overline{\zeta}_{01},\dots,\widehat{\overline{\zeta}_{0i}},\dots,\widehat{\overline{\zeta}_{0j}},\dots,\overline{\zeta}_{0k+1})\Big)\\
&\quad+\sum_{i=1}^{k+1}(-1)^{i-1}(\rho(\zeta_0)+\overline{\rho}(\overline{\zeta}_{0i}))\Big(\overline{\omega}(\overline{\zeta}_{01},\dots,\widehat{\overline{\zeta}_{0i}},\dots,\overline{\zeta}_{0k+1})\Big)\\
&\quad+\sum_{1\le i<j\le k+1}(-1)^{i+j}\,\overline{\omega}([\zeta_0,\overline{\zeta}_{0j}]-[\zeta_0,\overline{\zeta}_{0i}]+[\overline{\zeta}_{0i},\overline{\zeta}_{0j}],\overline{\zeta}_{01},\dots\\
&\qquad\qquad\qquad\dots,\widehat{\overline{\zeta}_{0i}},\dots,\widehat{\overline{\zeta}_{0j}},\dots,\overline{\zeta}_{0k+1})\\
&\quad+\sum_{1\le i<j<l\le k+1}(-1)^{i+j+l}\omega_0(\zeta_0,[\zeta_i,\zeta_j]+[\zeta_j,\zeta_l]+[\zeta_l,\zeta_i],\overline{\zeta}_{01},\dots\\
&\qquad\qquad\qquad\dots\widehat{\overline{\zeta}_{0i}},\dots,\widehat{\overline{\zeta}_{0j}},\dots,\widehat{\overline{\zeta}_{0l}},\dots,\overline{\zeta}_{0k+1}).
\end{aligned}$$

It is clear now that the terms which do not involve $\zeta_0$ combine exactly to $d\overline{\omega}(\overline{\zeta}_1, \ldots, \overline{\zeta}_{k+1})$. What remains is

$$\begin{aligned}
&\sum_{i=1}^{k+1}(-1)^{i-1}\rho(\zeta_0)\Big(\overline{\omega}(\overline{\zeta}_{01}, \ldots, \widehat{\overline{\zeta}_{0i}}, \ldots, \overline{\zeta}_{0k+1})\Big)\\
&+ \sum_{1\le i<j\le k+1}(-1)^{i+j}\overline{\rho}(\overline{\zeta}_{0j} - \overline{\zeta}_{0i})\Big(\omega_0(\zeta_0, \overline{\zeta}_{01}, \ldots, \widehat{\overline{\zeta}_{0i}}, \ldots, \widehat{\overline{\zeta}_{0j}}, \ldots, \overline{\zeta}_{0k+1})\Big)\\
&+ \sum_{1\le i<j\le k+1}(-1)^{i+j}\, \overline{\omega}([\zeta_0, \overline{\zeta}_{0j}] - [\zeta_0, \overline{\zeta}_{0i}], \overline{\zeta}_{01}, \ldots, \widehat{\overline{\zeta}_{0i}}, \ldots, \widehat{\overline{\zeta}_{0j}}, \ldots, \overline{\zeta}_{0k+1})\\
&+ \sum_{1\le i<j<l\le k+1}(-1)^{i+j+l}\omega_0(\zeta_0, [\zeta_i, \zeta_j] + [\zeta_j, \zeta_l] + [\zeta_l, \zeta_i], \overline{\zeta}_{01}, \ldots\\
&\qquad\qquad\qquad \ldots, \widehat{\overline{\zeta}_{0i}}, \ldots, \widehat{\overline{\zeta}_{0j}}, \ldots, \widehat{\overline{\zeta}_{0l}}, \ldots, \overline{\zeta}_{0k+1}),
\end{aligned}$$

and should be compared to $\sum_{i=1}^{k+1}(-1)^{i-1}d\omega_0(\zeta_0, \overline{\zeta}_{01}, \ldots, \widehat{\overline{\zeta}_{0i}}, \ldots, \overline{\zeta}_{0k+1})$, with $d\omega_0$ as defined by (6.39). It is obvious that the first three lines in the computation of $(d\omega)_0$ are exactly the ones we have in the above expression. The last term in (6.39) gives rise to triple sums of the form

$$\sum_{l=1}^{k+1}(-1)^{l-1}\sum_{i<j<l}(-1)^{i+j-1}\omega_0(\zeta_0, [\overline{\zeta}_{0i}, \overline{\zeta}_{0j}], \overline{\zeta}_{01}, \cdot\cdot, \overline{\zeta}_{0i}, \cdot\cdot, \overline{\zeta}_{0j}, \cdot\cdot, \overline{\zeta}_{0l}, \cdot\cdot, \overline{\zeta}_{0k+1})$$

(there is a similar term with $\sum_{i<l<j}$ and one with $\sum_{l<i<j}$). With suitable interchanges of summations, similar to what was explicitly explained before, these three terms combine to:

$$\sum_{1\le i<j<l\le k+1}(-1)^{i+j+l}\omega_0(\zeta_0, [\overline{\zeta}_{0i}, \overline{\zeta}_{0j}]+[\overline{\zeta}_{0j}, \overline{\zeta}_{0l}]+[\overline{\zeta}_{0l}, \overline{\zeta}_{0i}], \overline{\zeta}_{01}, \cdot\cdot, \widehat{\overline{\zeta}_{0i}}, \cdot\cdot, \widehat{\overline{\zeta}_{0j}}, \cdot\cdot, \widehat{\overline{\zeta}_{0l}}, \cdot\cdot, \overline{\zeta}_{0k+1}).$$

The proof now becomes complete if we observe that:

$$[\zeta_i, \zeta_j] + [\zeta_j, \zeta_l] + [\zeta_l, \zeta_i] = [\overline{\zeta}_{0i}, \overline{\zeta}_{0j}] + [\overline{\zeta}_{0j}, \overline{\zeta}_{0l}] + [\overline{\zeta}_{0l}, \overline{\zeta}_{0i}].$$

□

It is of some interest to work out some simple cases in detail. For a function $f \in C^\infty(M)$, $df$ is defined by

$$df(\zeta) = \rho(\zeta)(f) = \rho(\zeta_0)(f) + \overline{\rho}(\overline{\zeta})(f), \tag{6.41}$$

from which we learn that $(df)_0 = df$ (as expected) and $\overline{df}(\overline{\sigma}) = \overline{\rho}(\overline{\sigma})(f)$. If $\varphi$ is a 1-form, the defining relation (6.38) for its exterior derivative reads

$$d\varphi(\zeta_1, \zeta_2) = \rho(\zeta_1)(\varphi(\zeta_2)) - \rho(\zeta_2)(\varphi(\zeta_1)) - \overline{\varphi}([\zeta_1, \zeta_2]). \tag{6.42}$$

Introducing an arbitrary reference section $\zeta_0$, it is easy to verify that this can be rewritten as

$$d\varphi(\zeta_1, \zeta_2) = (d\varphi)_0(\zeta_0, \overline{\zeta}_{02}) - (d\varphi)_0(\zeta_0, \overline{\zeta}_{01}) + \overline{d\varphi}(\overline{\zeta}_{01}, \overline{\zeta}_{02}), \tag{6.43}$$

where $\overline{d\varphi} = d\overline{\varphi}$ and

$$(d\varphi)_0(\zeta, \overline{\sigma}) = \rho(\zeta)(\overline{\varphi}(\overline{\sigma})) - \overline{\rho}(\overline{\sigma})(\varphi(\zeta)) - \overline{\varphi}([\zeta, \overline{\sigma}]). \tag{6.44}$$

This is in perfect agreement with the results of Proposition 6.11 and Definition 6.10.

Concerning derivation properties, it is trivial to verify that for the product of functions: $d(fg) = fdg + gdf$. Also, from (6.42) applied to $f\varphi$ we get:

$$\begin{aligned} d(f\varphi)(\zeta_1, \zeta_2) &= \rho(\zeta_1)(f\varphi(\zeta_2)) - \rho(\zeta_2)(f\varphi(\zeta_1)) - f\overline{\varphi}([\zeta_1, \zeta_2]) \\ &= df(\zeta_1)\varphi(\zeta_2) - df(\zeta_2)\varphi(\zeta_1) \\ &\quad + f\Big(\rho(\zeta_1)(\varphi(\zeta_2)) - \rho(\zeta_2)(\varphi(\zeta_1)) - \overline{\varphi}([\zeta_1, \zeta_2])\Big), \end{aligned}$$

from which we conclude that

$$d(f\varphi) = f\, d\varphi + df \wedge \varphi. \tag{6.45}$$

Recalling now the general statements about derivations we made before, we can conclude that there exists a unique derivation $\hat{d}$ on $\bigwedge(\pi^\dagger)$, of degree 1, which coincides with our $d$ on functions and 1-forms. If we can show that $\hat{d}\omega = d\omega$ for an arbitrary $\omega \in \bigwedge(\pi^\dagger)$, we will know that the operator $d$ defined by (6.38) is a derivation. To this end, let us introduce

$$\hat{d}_\zeta = [i_\zeta, \hat{d}],$$

which, as commutator of two derivations, is itself a derivation of degree 0 on $\bigwedge(\pi^\dagger)$. We extend the action of $\hat{d}_\zeta$ to $\mathrm{Sec}(\pi)$ 'by duality'. That is to say, for $\eta \in \mathrm{Sec}(\pi)$, $\hat{d}_\zeta \eta$ is defined by requiring that for all $\varphi \in \bigwedge^1(\pi^\dagger)$:

$$\langle \hat{d}_\zeta \eta, \varphi \rangle = \hat{d}_\zeta(\varphi(\eta)) - \hat{d}_\zeta \varphi(\eta). \tag{6.46}$$

It is easy to see that $\hat{d}_\zeta \eta$ is skew-symmetric in $\zeta$ and $\eta$, so that it makes sense to introduce a bracket notation for it: $[\zeta, \eta]^\wedge = \hat{d}_\zeta \eta$. We now recall a result proved in [54] which, although stated there in an entirely different context, has a quite universal validity.

**Lemma 6.12.** *Given a derivation $\hat{d}$ of degree 1, and introducing $\hat{d}_\zeta$ and $[\zeta,\eta]^\wedge$ as above, we have for all $\omega \in \bigwedge^k(\pi^\dagger)$:*

$$\begin{aligned}\hat{d}\omega(\zeta_1,\dots,\zeta_{k+1}) &= \sum_{i=1}^{k+1}(-1)^{i-1}\hat{d}_{\zeta_i}\Big(\omega(\zeta_1,\dots,\widehat{\zeta_i},\dots,\zeta_{k+1})\Big) \\ &\quad + \sum_{1\le i<j\le k+1}(-1)^{i+j}\omega([\zeta_i,\zeta_j]^\wedge,\zeta_1,\dots,\widehat{\zeta_i},\dots,\widehat{\zeta_j},\dots,\zeta_{k+1}).\end{aligned}$$

To show that $\hat{d}\omega = d\omega$ now, it suffices to verify that $\hat{d}_\zeta f = \rho(\zeta)(f)$ on functions, and that the bracket $[\,,\,]^\wedge$ coincides with the Lie algebroid bracket. We have $\hat{d}_\zeta f = i_\zeta \hat{d} f = i_\zeta df = \rho(\zeta)(f)$, and for all $\varphi \in \bigwedge^1(\pi^\dagger)$

$$\begin{aligned}\langle [\zeta,\eta]^\wedge,\varphi\rangle &= \langle \hat{d}_\zeta\eta,\varphi\rangle = \rho(\zeta)(\varphi(\eta)) - \hat{d}\varphi(\zeta,\eta) - \hat{d}i_\zeta\varphi(\eta) \\ &= \rho(\zeta)(\varphi(\eta)) - \rho(\eta)(\varphi(\zeta)) - d\varphi(\zeta,\eta) = \overline{\varphi}([\zeta,\eta]),\end{aligned}$$

from which the desired result follows.

We now reach the main question which is about the relationship between $d^2$ and the compatibility requirements in the definition of an affine Lie algebroid. To appreciate the meaning of the following lemma, we take a step back and assume now that the bracket $[\zeta_i,\zeta_j]$ figuring in the definition (6.38) of $d$ satisfies the 'Leibniz-type property' (b6) with respect to the module structure of $\mathrm{Sec}(\pi)$ (and the resulting property (6.12)), but no further compatibility or Lie algebra conditions a priori. Remember that the property (b6) of the bracket was necessary to make sure that $d\omega$ is a form in the first place.

**Lemma 6.13.** *For all $\omega \in \bigwedge^k(\pi^\dagger)$ and $\zeta_i \in \mathrm{Sec}(\pi)$, we have*

$$\begin{aligned}&d^2\omega(\zeta_1,\dots,\zeta_{k+2}) = \\ &\quad\sum_{1\le i<j\le k+2}(-1)^{i+j}\Big(\overline{\rho}([\zeta_i,\zeta_j]) - [\rho(\zeta_i),\rho(\zeta_j)]\Big)\Big(\omega(\zeta_1,\dots,\widehat{\zeta_i},\dots,\widehat{\zeta_j},\dots,\zeta_{k+2})\Big) \\ &\quad + \sum_{1\le i<j<l\le k+2}(-1)^{i+j+l}\omega\big(\textstyle\sum_{i,j,l}[\zeta_i,[\zeta_j,\zeta_l]],\zeta_1,\,.\,,\widehat{\zeta_i},\,.\,,\widehat{\zeta_j},\,.\,,\widehat{\zeta_l},\,.\,,\zeta_{k+2}\big), \qquad (6.47)\end{aligned}$$

*(where the smaller summation sign of course refers again to a cyclic sum over the three indices involved).*

Proof: Remark that this lemma, with suitable adaptations, also has a rather universal validity. If $\omega$ is a $k$-form, then $d^2\omega$ is a $(k+2)$-form with

$$\begin{aligned} d^2\omega(\zeta_1,\ldots,\zeta_{k+2}) &= \sum_{i=1}^{k+2}(-1)^{i-1}\rho(\zeta_i)\Big(d\omega(\zeta_1,\ldots,\widehat{\zeta_i},\ldots,\zeta_{k+2})\Big) \\ &+\sum_{1\le i<j\le k+2}(-1)^{i+j}d\omega([\zeta_i,\zeta_j],\zeta_1,\ldots,\widehat{\zeta_i},\ldots,\widehat{\zeta_j},\ldots,\zeta_{k+2}). \end{aligned} \tag{6.48}$$

If we plug in the definition of $d\omega$, the first term on the right will further decompose into two parts, one involving double and the other involving triple sums. Based on our experience with such combinatorics in one of the preceding proofs, we can right away conclude that the first line of the right-hand side of (6.48) equals:

$$\begin{aligned} &\sum_{1\le i<j\le k+2}(-1)^{i+j}\Big(\rho(\zeta_j)\rho(\zeta_i)-\rho(\zeta_i)\rho(\zeta_j)\Big)\Big(\omega(\zeta_1,\ldots,\widehat{\zeta_i},\ldots,\widehat{\zeta_j},\ldots,\zeta_{k+2})\Big) \\ &+\sum_{1\le i<j<l\le k+2}(-1)^{i+j+l-1}\textstyle\sum_{i,j,l}\Big\{\rho(\zeta_i)\Big(\omega([\zeta_j,\zeta_l]\Big\},\zeta_1,\,.\,,\widehat{\zeta_i},\,.\,,\widehat{\zeta_j},\,.\,,\widehat{\zeta_l},\,.\,,\zeta_{k+2})\Big), \end{aligned} \tag{6.49}$$

where the smaller summation sign, as before, refers to a cyclic sum, the range of which is delimited by the curly brackets. For the second term on the right in (6.48), we have to remember that the first argument is a vector section. Using the defining relation (6.36), applied to $d\omega$, we obtain:

$$\begin{aligned} d\omega(\overline{\sigma},\zeta_1,\ldots,\zeta_k) &= \overline{\rho}(\overline{\sigma})\Big(\omega(\zeta_1,\ldots,\zeta_k)\Big) \\ &+\sum_{i=1}^{k}(-1)^i\rho(\zeta_i)\Big(\omega(\overline{\sigma},\zeta_1,\ldots,\widehat{\zeta_i},\ldots,\zeta_k)\Big) \\ &+\sum_{j=1}^{k}(-1)^j\omega([\overline{\sigma},\zeta_j],\zeta_1,\ldots,\widehat{\zeta_j},\ldots,\zeta_k) \\ &+\sum_{1\le i<j\le k}(-1)^{i+j}\omega([\zeta_i,\zeta_j],\overline{\sigma},\zeta_1,\ldots,\widehat{\zeta_i},\ldots,\widehat{\zeta_j},\ldots,\zeta_k). \end{aligned} \tag{6.50}$$

The last line here has two vector arguments, but this is consistent with the application of Definition 6.8 to a form of type $i_{\overline{\zeta}}\omega$. We look at the effect of each of these four terms, when inserted in the second sum of (6.48). The first one simply gives:

$$\sum_{1\leq<i<j\leq k+2}(-1)^{i+j}\overline{\rho}([\zeta_i,\zeta_j])\Big(\omega(\zeta_1,\ldots,\widehat{\zeta_i},\ldots,\widehat{\zeta_j},\ldots,\zeta_{k+2})\Big). \qquad (6.51)$$

The second one is easily seen to give rise to terms which cancel exactly the second sum in (6.49). The third term of (6.50) gives rise to expressions involving double brackets, which combine to:

$$\sum_{1\leq i<j<l\leq k+2}(-1)^{i+j+l}\omega(\textstyle\sum_{i,j,l}[[\zeta_i,\zeta_j],\zeta_l],\zeta_1,\,.\,,\widehat{\zeta_i},\,.\,,\widehat{\zeta_j},\,.\,,\widehat{\zeta_l},\,.\,,\zeta_{k+2}). \quad (6.52)$$

The fourth term of (6.50) finally creates terms which involve two double sums, and in each of the summands the first two arguments of $\omega$ are brackets. One has to look at all possible orderings, six in total, of the four different indices involved, but when the same procedure is applied to shuffle the order of summations suitably around and rename indices where appropriate, one easily finds that the six terms cancel each other two by two in view of the skew-symmetry of $\omega$. What we are left with in the end is the first term of (6.49), (6.51) and (6.52): they precisely combine to the statement of the lemma. □

**Proposition 6.14.** *The exterior derivative has the property $d^2 = 0$ if and only if the bracket further satisfies the Jacobi identity (5.36) (or, equivalently, (b5)).*

PROOF: If (b6) and (5.36) hold true, we also have (b5) and we know from previous considerations that (6.13) then holds as well. The above lemma this way trivially implies $d^2 = 0$. For the converse, we observe that $d^2 f = 0$, for all $f \in C^\infty(M)$, implies (6.15), from which it subsequently follows that $d^2\varphi = 0$, for all $\varphi \in \bigwedge^1(\pi^\dagger)$, implies (5.36). □

To complete the picture of basic derivations on $\bigwedge(\pi^\dagger)$, we have a closer look at the analogue of the classical Lie derivative.

**Definition 6.15.** *For every $\zeta \in \mathrm{Sec}(\pi)$, the derivation $d_\zeta$ of degree zero is defined as*

$$d_\zeta = [i_\zeta, d] = i_\zeta \circ d + d \circ i_\zeta. \qquad (6.53)$$

So, since $d_\zeta$ is defined as a commutator of derivations of degree -1 and 1, we know that it will itself be a derivation of degree zero: $d_\zeta(\Lambda^k(\pi^\dagger)) \subset \Lambda^k(\pi^\dagger)$ and

$$d_\zeta(\omega \wedge \mu) = d_\zeta\omega \wedge \mu + \omega \wedge d_\zeta\mu. \tag{6.54}$$

Likewise, we can rely on proofs similar to those in the standard theory to conclude that the following commutator properties will hold true:

$$[d_\zeta, i_\eta] = i_{[\zeta,\eta]}, \qquad [d_\zeta, d] = 0, \qquad [d_\zeta, d_\eta] = d_{[\zeta,\eta]}. \tag{6.55}$$

Note, however, that a Lie-type derivation with respect to a vector section turns up in the last property, and this is indeed well defined as follows: $d_{\overline{\sigma}} = [i_{\overline{\sigma}}, d]$. It is further natural to extend the action of $d_\zeta$ to $\mathrm{Sec}(\pi)$ by duality, i.e. to require that a property of type (6.46) holds true. It then follows, as expected, that for $\eta, \zeta \in \mathrm{Sec}(\pi)$,

$$d_\zeta\eta = [\zeta, \eta]. \tag{6.56}$$

As a result of such an extension, $d_\zeta$ has Leibniz-type properties also with respect to the evaluation of forms on the appropriate number of affine (or vector) sections; the following property, which could be verified by a direct computation from the definition of $d_\zeta$, thus becomes self-evident:

$$d_\eta\omega(\zeta_1,\ldots,\zeta_k) = \rho(\eta)\Big(\omega(\zeta_1,\ldots,\zeta_k)\Big) + \sum_{j=1}^{k}(-1)^j\omega([\eta,\zeta_j],\zeta_1,\ldots,\widehat{\zeta_j},\ldots,\zeta_k). \tag{6.57}$$

Finally, we list what the two composing parts of $d_\zeta\omega$ are, in the sense of the defining relation (6.30) of forms.

**Proposition 6.16.** *For $\omega \in \bigwedge^k(\pi^\dagger)$, we have*

$$\begin{aligned}(d_\zeta\omega)_0(\zeta_0,\overline{\zeta}_1,\ldots,\overline{\zeta}_{k-1}) &= \rho(\zeta)\Big(\omega_0(\zeta_0,\overline{\zeta}_1,\ldots,\overline{\zeta}_{k-1})\Big) - \overline{\omega}([\zeta,\zeta_0],\overline{\zeta}_1,\ldots,\overline{\zeta}_{k-1}) \\ &\quad + \sum_{j=1}^{k-1}(-1)^j\omega_0(\zeta_0,[\zeta,\zeta_j],\overline{\zeta}_1,\ldots,\widehat{\overline{\zeta}_j},\ldots,\overline{\zeta}_{k-1}), \qquad (6.58)\\ \overline{d_\zeta\omega}(\overline{\zeta}_1,\ldots,\overline{\zeta}_k) &= \rho(\zeta)\Big(\overline{\omega}(\overline{\zeta}_1,\ldots,\overline{\zeta}_k)\Big) \\ &\quad + \sum_{j=1}^{k-1}(-1)^j\overline{\omega}([\zeta,\zeta_j],\overline{\zeta}_1,\ldots,\widehat{\overline{\zeta}_j}\ldots,\overline{\zeta}_k). \qquad (6.59)\end{aligned}$$

These are exactly the sort of expressions one expects. The proof is a matter of a direct computation, starting from the formula (6.57) and using the decompositions (6.30) and (6.37). It further requires manipulations of double sums of the same nature as those in some previous proofs.

## 6.6 Coordinate expressions

We list now coordinate expressions for the basic exterior derivatives. Let $(x^I)$ as before be coordinates on $M$. For their exterior derivatives we obtain the following: for all $\zeta \in \mathrm{Sec}(\pi)$,

$$dx^I(\zeta) = \rho(\zeta)(x^I). \tag{6.60}$$

In terms of Definition 6.5, it follows that $\overline{dx^I}(\overline{\sigma}) = \overline{\rho}(\overline{\sigma})(x^I)$ (and of course $(dx^I)_0 = dx^I$). For the general representation (5.18) of $dx^I$ as a 1-form, we find:

$$dx^I = \rho^I_0 e^0 + \rho^I_\alpha e^\alpha \tag{6.61}$$

and for an arbitrary function $f \in C^\infty(M)$

$$df = \rho^I_0 \frac{\partial f}{\partial x^I} e^0 + \rho^I_\alpha \frac{\partial f}{\partial x^I} e^\alpha.$$

The defining properties of an affine Lie algebroid immediately imply that $de^0 = 0$. We further calculate, making use for example of (5.23,5.24), the general formula (6.42) and the coordinate expressions (6.16), that

$$de^\alpha = -C^\alpha_{0\beta} e^0 \wedge e^\beta - \tfrac{1}{2} C^\alpha_{\beta\gamma} e^\beta \wedge e^\gamma. \tag{6.62}$$

It is instructive to verify that expressing the properties $d^2 e^\alpha = 0$ and $d^2 x^I = 0$ is indeed equivalent to the requirements (6.18,6.19) and (6.20,6.21), respectively.

In the interest of doing computations, we also list the Lie-type derivatives of functions $f \in C^\infty(M)$ and the local basis of 1-forms. For $\zeta = e_0 + \zeta^\alpha \overline{e}_\alpha$,

$$d_\zeta f = \rho(\zeta)(f), \qquad d_\zeta e^0 = 0, \qquad d_\zeta e^\alpha = C^\alpha_{0\beta}\zeta^\beta e^0 - C^\alpha_{0\beta} e^\beta + C^\alpha_{\beta\gamma}\zeta^\gamma e^\beta + d\zeta^\alpha.$$

## 6.7 Relation with the exterior derivative on the bi-dual

There exists a far less technical way to prove that $d^2 = 0$. Results such as Proposition 5.6 about exteriors forms on affine spaces obviously carry over to differential forms on affine bundle. Hence, if $\omega$ is a $k$-form on $\pi$ there exist a unique $\tilde{\omega}$ on $\tilde{\pi}$ such that $\omega = \iota^*(\tilde{\omega})$. One can then simply define

$$d\omega = \iota^*(d\tilde{\omega}). \tag{6.63}$$

One can verify that this gives the same $d$ as the one defined by the direct formula (6.38). Of course, it has the advantage that $d^2 = 0$ now becomes an obvious property:

$$d^2\omega = d(d(\iota^*\tilde{\omega})) = d(\iota^* d\tilde{\omega}) = \iota^*(d^2\tilde{\omega}) = 0.$$

The purpose of the more lengthy considerations in the previous section, however, was to understand in detail how the various ingredients of the direct definition of an affine Lie algebroid are related to the construction of an exterior calculus.

Recall that, for any Lie algebroid on $\pi$, there is a corresponding Lie algebroid structure on $\overline{\pi}$. For any form $\overline{\omega}$ on $\overline{\pi}$, one can likewise compute its exterior derivative by taking any $\tilde{\omega}$ on $\tilde{\pi}$ for which $\overline{\iota}^*\tilde{\omega} = \overline{\omega}$ and putting $d\overline{\omega} = \overline{\iota}^* d\tilde{\omega}$. It follows that the set of forms on $\tilde{\pi}$ for which $\overline{\iota}^*\tilde{\omega} = 0$ is a differential ideal. This ideal is generated by the 1-form $e^0$ which, as we have seen, is actually closed. The point we wish to make here is that $de^0 = 0$ fully characterises the fact that the Lie algebroid structure on $\tilde{\pi}$ comes from an affine Lie algebroid on $\pi$, in the following sense. Suppose we are given a (vector) Lie algebroid on $\tilde{E} \to M$ and want to figure out whether it restricts to a Lie algebroid on the affine subbundle $E \to M$.

**Proposition 6.17.** *A Lie algebroid structure on $\tilde{\pi}$ restricts to a Lie algebroid structure on the affine bundle $\pi$ if and only if the exterior differential satisfies $de^0 = 0$.*

Proof: Indeed, taking two sections $\zeta_1$ and $\zeta_2$ of $\pi$ we have

$$\begin{aligned} de^0(\iota(\zeta_1), \iota(\zeta_2)) &= \tilde{\rho}(\iota(\zeta_1))\langle e^0, \iota(\zeta_2)\rangle - \tilde{\rho}(\iota(\zeta_2))\langle e^0, \iota(\zeta_1)\rangle \\ &\quad - \langle e^0, [\iota(\zeta_1), \iota(\zeta_2)]\rangle \\ &= -\langle e^0, [\iota(\zeta_1), \iota(\zeta_2)]\rangle \end{aligned}$$

It follows that $[\iota(\zeta_1), \iota(\zeta_2)]$ is in $\mathrm{Im}(\bar{\iota}) = \mathrm{Ker}(e^0)$ if and only if $de^0$ vanishes on the image of $\iota$, which spans $\tilde{E}$. Similar observations hold for $[\bar{\iota}(\bar{\zeta}_1), \bar{\iota}(\bar{\zeta}_2)]$ and $[\iota(\zeta_1), \bar{\iota}(\bar{\zeta}_2)]$. □

To end this section, we come back to the case where the base manifold $M$ is fibred over the real line $\pi_R \colon M \to I\!R$ and the anchor map takes values in $J^1M$. In this case we have that $dt = e^0$ (see expression 6.61) so that $e^0$ is not only closed but also exact. In fact, this is the condition for a Lie algebroid structure on an affine bundle to have a 1-jet-valued anchor. Indeed, if there exists an $f \in C^\infty(M)$ such that $df = e^0$, then the partial derivatives of $f$ cannot simultaneously vanish, hence $f$ defines a local fibration and then for any section $\zeta$ of $\pi$ we have that $\rho(\zeta)f = \langle df, \zeta\rangle = \langle e^0, \zeta\rangle = 1$, which is the condition for the anchor being 1-jet-valued.

## 6.8 Poisson structure

When we have a Lie algebroid structure on $\tilde{E}$, there is a Poisson bracket on the dual bundle $\tilde{E}^* = E^\dagger$. Any section $\zeta$ of $\tilde{\pi}$, and in particular any section of $\pi$, determines a fibre linear function $\widehat{\zeta}$ on $E^\dagger$ by

$$\widehat{\zeta}(\varphi) = \langle \zeta_m, \varphi\rangle \qquad \text{for every } \varphi \in E^\dagger_m.$$

Then the Poisson bracket is determined by the condition

$$\{\widehat{\zeta}_1, \widehat{\zeta}_2\} = \widehat{[\zeta_1, \zeta_2]},$$

which for consistency (using linearity and the Leibnitz rule) requires that we put

$$\{\widehat{\zeta}, g\} = \tilde{\rho}(\zeta)(g), \qquad \text{and} \qquad \{f, g\} = 0,$$

for $f$ and $g$ functions on $M$.

In coordinates, we have

$$\begin{aligned} &\{x^I, x^J\} = 0 && \{\mu_0, x^I\} = \rho^I_0 && \{\mu_\alpha, x^I\} = \rho^I_\alpha \\ &\{\mu_0, \mu_\beta\} = C^\gamma_{0\beta}\mu_\gamma && \{\mu_\alpha, \mu_\beta\} = C^\gamma_{\alpha\beta}\mu_\gamma \end{aligned}$$

and therefore the Poisson tensor is

$$\begin{aligned} \Lambda_{E^\dagger} = \rho^I_\alpha \frac{\partial}{\partial\mu_\alpha} \wedge \frac{\partial}{\partial x^I} + \frac{1}{2}\mu_\gamma\, C^\gamma_{\alpha\beta} \frac{\partial}{\partial\mu_\alpha} \wedge \frac{\partial}{\partial\mu_\beta} + \\ + \frac{\partial}{\partial\mu_0} \wedge \left(\rho^I_0 \frac{\partial}{\partial x^I} + \mu_\gamma\, C^\gamma_{0\beta} \frac{\partial}{\partial\mu_\beta}\right). \end{aligned}$$

It is of some interest to mention yet another characterisation of the result described in Proposition 6.4. The above Poisson bracket in fact is determined by the bracket of linear functions coming from sections of $\pi$, since these span the set of all linear functions on $E^\dagger$. But the bracket of sections of $\pi$ is a section of $\overline{\pi}$; it follows that the corresponding Poisson brackets are independent of the coordinate $\mu_0$, and therefore, $\frac{\partial}{\partial\mu_0}$ is a symmetry of the Poisson tensor. Conversely, it is obvious that the latter symmetry property will imply that the bracket of sections in the image of $\iota$ belongs to the image of $\overline{\iota}$.

## 6.9 Examples

**The canonical affine Lie algebroid** The canonical example of a Lie algebroid over an affine bundle is the first jet bundle $J^1M \to M$ of a manifold $M$ fibered over the real line $\pi_R\colon M \to I\!R$. The elements of the manifold $J^1M$ are equivalence classes $j^1_t\gamma$ of sections $\gamma$ of the bundle $\pi_R\colon M \to I\!R$, where two sections are equivalent if they have first order contact at the point $t$. It is an affine bundle whose associated vector bundle is $\overline{\pi}_M : VM \to M$ the set of vectors tangent to $M$ which are vertical over $I\!R$. In this case it is well-known that $J^1M^\dagger = T^*M$, and therefore $\widetilde{J^1M} = TM$. The canonical immersion is given by

$$\iota_{J^1M}(j^1_t\gamma) = \dot{\gamma}(t),$$

i.e. it maps the 1-jet of the section $\gamma$ at the point $t$ to the vector tangent to $\gamma$ at the point $t$. In coordinates, if $j^1_t\gamma$ has coordinates $(t, x, v)$ then

$$\iota_{J^1M}(t, x, v) = \frac{\partial}{\partial t}\Big|_{(t,x)} + v^i \frac{\partial}{\partial x^i}\Big|_{(t,x)}.$$

An element $w$ of the associated vector bundle $\overline{\pi}_M : VM \to M$ is of the form

$$w = w^i \frac{\partial}{\partial x^i}\Big|_{(t,x)}.$$

The bracket of sections of $J^1M$ is defined precisely by means of the above identification of 1-jets with vectors $v \in TM$ which project onto the vector $\partial/\partial t$. In coordinates a section $X$ of $J^1M$ is identified with the vector field

$$X = \frac{\partial}{\partial t} + X^i(t, x)\frac{\partial}{\partial x^i},$$

and the bracket is

$$[X, Y] = \{X(Y^i) - Y(X^i)\}\frac{\partial}{\partial x^i}$$

which is obviously a section of the vector bundle.

**Affine distributions** An affine sub-bundle $E$ of $J^1M$ is called involutive if the bracket of sections of the sub-bundle is a section of the associated vector bundle. Therefore, taking as anchor the natural inclusion into $TM$ and as bracket the restriction of the bracket in $J^1M$ to $E$, we have an affine Lie algebroid structure on $E$.

**Lie algebra structures on affine spaces** We consider the case in which the manifold $M$ reduces to one point $M = \{m\}$. Thus our affine bundle is $E = \{m\} \times A$ and the associated vector bundle is $\overline{E} \equiv \{m\} \times \overline{A}$ for some affine space $A$ modelled on the vector space $\overline{A}$. Then, a Lie algebroid structure on the affine bundle $E$ is just an affine Lie algebra structure on $A$. Indeed, every section of $E$ and of $\overline{E}$ is determined by a point in $A$ and $\overline{A}$, respectively. The anchor must vanishes since $TM = \{0_m\}$, so it does not carry any additional information.

**Trivial affine Lie algebroids** By a trivial affine space we mean just a point $A = \{o\}$, and the associated vector space is the trivial one $\overline{A} = \{0\}$. The extended affine dual of $A$ is $A^\dagger = I\!R$ since the only affine maps defined on a space of just a point are the constant maps. It follows that the extended bi-dual is $\tilde{A} = I\!R$

Given a manifold $M$, we consider the affine bundle $E = M \times \{o\}$ with associated vector bundle $\overline{E} = M \times \{0\}$. On $\overline{E}$ we consider the trivial bracket $[\,,\,] = 0$ and the anchor $\overline{\rho} = 0$, and as derivation $D_o$ we also take $D_o = 0$. Now, to construct a Lie algebroid structure on $E$, we take an arbitrary vector field $X_0$ on $M$ as given and define the map $\rho\colon E \to TM$ by $\rho(m, o) = X_0(m)$. Then it follows that $\rho$ is compatible with $D_o$.

The extended dual of $E$ is $E^\dagger = M \times I\!R$ and the extended bi-dual is $\tilde{E} = M \times I\!R$. We therefore have one section $e^0$ spanning the set of sections of $E^\dagger$, and the dual element $e_0$ (which is just the image under the canonical immersion of the constant section of value 0).

We want to study the associated exterior differential operator and Poisson bracket.

For the exterior differential operator, since $E^\dagger$ has 1-dimensional fibres it follows that $de^0 = 0$. On functions $f \in C^\infty(M)$ we have $df = \rho(e_0)(f)e^0 = X_0(f)e^0$.

For the Poisson structure, since the fibre of $E^\dagger$ is 1-dimensional, it is determined by the equation $\{\hat{e_0}, f\} = \rho(e_0)(f)$ and $\hat{e_0} = \mu_0$. We have that the only non-trivial brackets are $\{\mu_0, f\} = X_0(f)$. Therefore, the Poisson tensor is

$$\Lambda = \frac{\partial}{\partial \mu_0} \wedge X_0.$$

**Quotient by a group** If $p\colon Q \to M$ is a principal $G$-bundle, $M$ is fibred over $\mathbb{R}$ and $J^1Q$ the first jet manifold of $Q \to \mathbb{R}$, then $E = J^1Q/G \to M$ is an affine Lie algebroid. The anchor is $\rho([j^1_t\gamma]) = j^1_t(p \circ \gamma)$. The bracket is obtained by projecting the bracket on $J^1Q$. We have that $\tilde{E}$ is the Atiyah algebroid $TQ/G$, see for instance [10].

**Affine actions of Lie algebras** Let $A$ be an affine space endowed with a Lie algebra structure. By an action of $A$ on a manifold $M$ we mean an affine map $\phi\colon A \to \mathcal{X}(M)$, such that $[\phi(a), \phi(b)] = \overline{\phi}([a,b])$. Then $E = M \times A \to M$ has an affine Lie algebroid structure. The anchor is $\rho(m,\xi) = \phi(\xi)(m)$ and the bracket can be defined in terms of constant sections: the bracket of two constant sections $\sigma_i(m) = (m, \xi_i)$, is the constant section corresponding to the bracket of the values

$$[\sigma_1, \sigma_2](m) = (m, [\xi_1, \xi_2]_A).$$

If we consider the Lie algebra $\tilde{A}$ then $\tilde{A}$ acts also on the manifold $M$. The extension $\tilde{E}$ is the Lie algebroid associated to the action of $\tilde{A}$.

**Poisson manifolds with symmetry** Consider a Poisson manifold $(M, \Lambda)$ and an infinitesimal symmetry $Y \in \mathcal{X}(M)$ of $\Lambda$, that is $\mathcal{L}_Y\Lambda = 0$ . Take $E$ to be $T^*M$ with its natural affine structure, where the associated vector bundle $\overline{E}$ is $T^*M$ itself. On $\overline{E}$ we consider the Lie algebroid structure defined by the canonical Poisson structure. For a section $\alpha$ of $\pi$ (i.e. a 1-form on $M$) we define the map $D_\alpha\colon \mathrm{Sec}(\overline{\pi}) \to \mathrm{Sec}(\overline{\pi})$ by

$$D_\alpha\beta = \mathcal{L}_Y\beta + [\alpha, \beta].$$

Since $Y$ is a symmetry of $\Lambda$, $D_\alpha$ is a derivation and clearly satisfies the required compatibility condition. If we further consider the affine anchor

$\rho\colon E \to TM$, determined by $\rho(\alpha_m) = \Lambda(\alpha_m) + Y_m$, then we have a Lie algebroid structure on the affine bundle $\pi$.

In this case, since there is a distinguished section of $E$ (the zero section), we have that $E^\dagger = TM \times \mathbb{R}$ and $\tilde{E} = T^*M \times \mathbb{R}$.

**Jets of sections in a groupoid** Let $G$ be a Lie groupoid over a manifold $M$ with source $\alpha$ and target $\beta$ (the notation is as in [10]). Let $T^\alpha_{G^{(0)}}G = \ker T\alpha|_{G^{(0)}}$ be the associated Lie algebroid, that is, the set of $\alpha$-vertical vectors at points in $G^{(0)}$ (the set of identities). The anchor is the map $\rho = T\beta$. Assume that $M$ is further fibred over the real line, $\pi_R\colon M \to \mathbb{R}$ and consider the bundle $E = J^\alpha_{G^{(0)}}G$ of 1-jets of sections of $\pi_R \circ \beta$ which are $\alpha$-vertical, at points in $G^{(0)}$. This is an affine bundle whose associated vector bundle is $(T^\alpha_{G^{(0)}}G)^{\mathrm{ver}}$ the set of $(\pi_R \circ \beta)$-vertical vectors on $T^\alpha_{G^{(0)}}G$. If $\bar{\imath}$ is the natural inclusion of $(T^\alpha_{G^{(0)}}G)^{\mathrm{ver}}$ into $T^\alpha_{G^{(0)}}G$ and we define the map $j\colon T^\alpha_{G^{(0)}}G \to M \times \mathbb{R}$ by $j(v) = (\alpha(\tau_G(v)), t(v))$ (where $t = \pi_R \circ \beta \circ \tau_G$), then we have the exact sequence of vector bundles over $M$

$$0 \longrightarrow (T^\alpha_{G^{(0)}}G)^{\mathrm{ver}} \xrightarrow{\bar{\imath}} T^\alpha_{G^{(0)}}G \xrightarrow{j} M \times \mathbb{R} \longrightarrow 0$$

and $j^{-1}(M \times \{1\}) = J^\alpha_{G^{(0)}}G$. Moreover, the bracket of two sections of $J^\alpha_{G^{(0)}}G$ is vertical over $\mathbb{R}$ from where it follows that the Lie algebroid structure of $T^\alpha_{G^{(0)}}G$ restricts to a Lie algebroid structure on the affine bundle $J^\alpha_{G^{(0)}}G$.

# Chapter 7

# *Dynamics on affine Lie algebroids*

## 7.1 $\lambda$-admissible curves and dynamics

As we expressed in chapter 4, the model of affine Lie algebroids we are developing should in the first place offer an environment in which one can accommodate the time-dependent Lagrange-type equations (4.11). At present, we wish to look in more detail at the geometric nature of the more general dynamical systems, which we call pseudo-second-order equations, and are those described by differential equations of the form (4.10). For this purpose in fact, we do not need the full machinery of algebroids: it suffices to assume that $M$ is fibred over $\mathbb{R}$ (with projection $\pi_{\mathbb{R}}$), $E$ is an affine bundle over $M$ with projection $\pi$ and $\lambda : E \to J^1M$ an affine bundle map over the identity. As shown in the previous chapter, we can then easily find an anchor map $\rho$ that takes values in $TM$ by composing $\lambda$ with the natural injection of $J^1M$ into $TM$: $\rho = \iota_{J^1M} \circ \lambda$. Using the fibration $\pi_{\mathbb{R}}$ of $M$, we can also consider $E$ as being fibred over $\mathbb{R}$ with projection $\pi_{\mathbb{R}} \circ \pi$. Therefore, it is possible to look at the the first jet bundle of this fibration. We will use the notation $\pi_E$ for the bundle $J^1E \to E$.

**Definition 7.1.** *A curve $\psi$ in $E$ which, in addition, is a section of $\pi_{\mathbb{R}} \circ \pi$, is said to be $\lambda$-admissible, if $\lambda \circ \psi = j^1(\pi \circ \psi)$.*

One could say that $\psi$ is the $\lambda$-prolongation of a curve in $M$. In coordinates, we have

$$\psi : t \mapsto (t, x^i(t), y^\alpha(t)), \qquad \text{with} \quad \dot{x}^i(t) = \lambda^i(t, x(t)) + \lambda^i_\alpha(t, x(t))\, y^\alpha(t).$$

Note in passing that, not unexpectedly, one can characterise $\lambda$-admissibility via a concept of contact forms: putting $\Theta^i = \lambda^*\theta^i$, where the $\theta^i = dx^i - v^i dt$ are the contact forms on $J^1M$, we have that $\psi$ is a $\lambda$-admissible curve in $E$ if and only if $\psi^*\Theta^i = 0$.

Pseudo-second-order equation fields on $E$ are vector fields whose integral curves all are $\lambda$-admissible curves. As in the standard theory of SODEs on a tangent bundle or first jet bundle, however, there is a simple direct characterisation of such vector fields.

**Definition 7.2.** $\Gamma \in \mathcal{X}(E)$ *is a pseudo-second-order equation field if*

$$T\pi \circ \Gamma = \iota_{J^1 M} \circ \lambda,$$

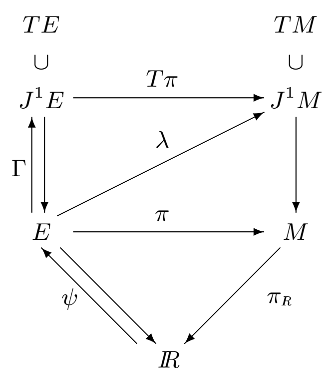
*where* $\iota_{J^1 M}$ *is the injection of* $J^1 M$ *into* $TM$.

Clearly, in coordinates, a pseudo-SODE is of the form

$$\Gamma = \frac{\partial}{\partial t} + (\lambda^i(t,x) + \lambda^i_\alpha(t,x) y^\alpha) \frac{\partial}{\partial x^i} + f^\alpha(t,x,y) \frac{\partial}{\partial y^\alpha}, \tag{7.1}$$

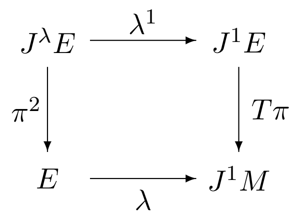
for some functions $f^\alpha$, and it is obvious that all its integral curves will be $\lambda$-admissible.

The following diagram visualises the notions of $\lambda$-admissible curves and pseudo-SODEs.




An important point now, however, is that there is a natural way of interpreting the vector field $\Gamma$ as section of a different bundle.

From the above definition, it is clear that a pseudo-SODE is actually a section of $\pi_E : J^1 E \to E$, with the additional property that for all $p \in E$, $T\pi|_{J^1 E}(\Gamma(p)) = \lambda(p)$. An equivalent way of saying the same thing, by definition of the concept of a pullback bundle, is that $(p, \Gamma(p))$ is a point of $\lambda^* J^1 E$, with $J^1 E$ regarded as fibred over $J^1 M$ via $T\pi|_{J^1 E}$. From now on, we will write $J^\lambda E$ for $\lambda^* J^1 E$, and denote its two projections as indicated in the following diagram:




If we finally put $\pi^1 = \pi_E \circ \lambda^1$, there is yet another way of expressing the characterisation of a pseudo-SODE. Indeed, from the trivial observation that $\pi_E(\Gamma(p)) = p = \pi^2((p, \Gamma(p)))$, it follows that a pseudo-SODE $\Gamma$ can also be regarded as a section of the bundle $\pi^1 : J^\lambda E \to E$, with the property that $\pi^2 \circ \Gamma = \pi^1 \circ \Gamma$.

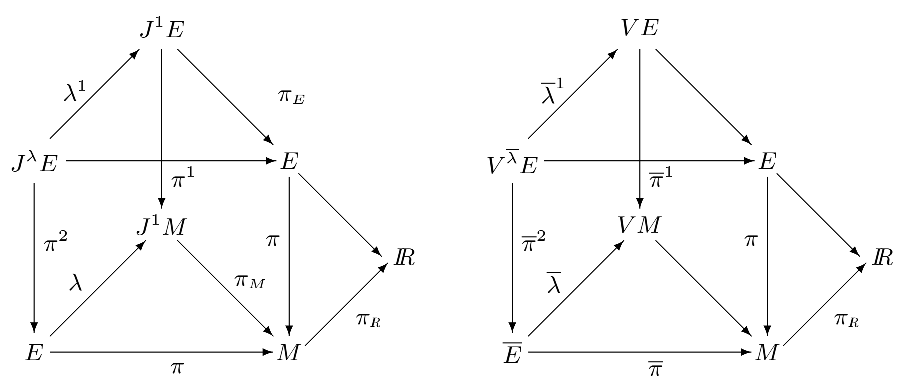
The various spaces and projections, described in this discussion, are depicted in the next diagrams. Remark that we are now looking at the bundle $\pi^1 : J^\lambda E \to E$, whose total space is the manifold

$$J^\lambda E = \lambda^* J^1 E = \{(e, Z) \in E \times J^1 E \mid \lambda(e) = T\pi|_{J^1E}(Z)\}, \tag{7.2}$$

but the fibration we focus on is not one of the projections which define $J^\lambda E$, but rather the map $\pi^1 = \pi_E \circ \lambda^1$. As such, we are looking at an affine bundle, modelled on the vector bundle $\overline{\pi}^1 : V^{\overline{\lambda}} E \to E$, with total space

$$V^{\overline{\lambda}} E = \{(\overline{e}, \overline{V}) \in \overline{E} \times VE \mid \overline{\lambda}(\overline{e}) = T\pi|_{VE}(\overline{V})\}. \tag{7.3}$$

Here, $\overline{\lambda}$ is the linear bundle map $\overline{E} \to VM$ that is associated to $\lambda : E \to J^1M$.




The above diagrams immediately suggests the following question: if we put more structure into the scheme by assuming that $\pi : E \to M$ carries an affine Lie algebroid structure (and thus $\overline{\pi} : \overline{E} \to M$ a vector Lie algebroid structure), is it possible to prolong this structure to the bundle $\pi^1 : J^\lambda E \to E$, in such a way of course that $\lambda^1$ becomes the anchor map of the induced affine Lie algebroid (related to a Lie algebroid on $\overline{\pi}^1$)? In fact, some aspects related to this question can be formulated also in a more general set-up. Therefore, we will first explore the prolongation idea in sufficient detail in this more general framework and come back to the above special situation later. Moreover, as above, the fibration of $M$ over $I\!R$ will only be important if we want the dynamical systems to depend explicitly on time. Usually, we will therefore omit this extra requirement.

## 7.2 $\varrho$-prolongation of a fibre bundle

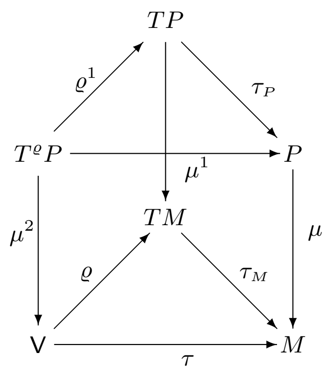
In this section we define the prolongation of a fibre bundle with respect to an anchored vector bundle, following ideas first introduced in [37]. We are primarily interested in the prolongation of the affine bundle $E \to M$, but we will describe explicitly a more general construction first, since this does not introduce extra complications. Let $\mu : P \to M$ be an arbitrary fibre bundle and $\tau : \mathsf{V} \to M$ a vector bundle. Assume we have an anchor map on $\tau$ at our disposal, i.e. a vector bundle morphism $\varrho : \mathsf{V} \to TM$ over the identity.

**Definition 7.3.** *The $\varrho$-prolongation of $\mu : P \to M$ is the bundle $\mu^1 : T^\varrho P \to P$, constructed as follows: (i) the total space $T^\varrho P$ is the total space of the pullback bundle $\varrho^* TP$*

$$T^\varrho P = \{(\mathsf{v}, X_p) \in \mathsf{V} \times TP \mid \varrho(\mathsf{v}) = T\mu(X_p)\}; \tag{7.4}$$

*(ii) if $\varrho^1$ denotes the projection of $\varrho^* TP$ onto $TP$ and $\tau_P$ is the tangent bundle projection, then $\mu^1 = \tau_P \circ \varrho^1$.*

The situation is summarised in the following diagram, whereby the projection on the first element of a pair $(\mathsf{v}, X_p) \in T^\varrho P$ is denoted by $\mu^2$.




The vector bundle structure of $\tau$ and $\tau_P$ carries over to the bundle $\mu^1$. In a lot of situations, the prolonged bundle $\mu^1$ shows a 'tangent-bundle-like behaviour'. For example, it is possible to get a notion of 'vertical elements'. An element $Z$ of $T^\varrho P$ is said to be *vertical* if $\mu^2(Z) = 0$. The set of all vertical elements in $T^\varrho P$ is a vector subbundle of $\mu^1$ and will be denoted by $\mathcal{V}^\varrho P$. If $(0, V_p) \in \mathcal{V}^\varrho P$, then $V_p = \varrho^1(0, V_p)$ will also be vertical in $TP$, since $T\mu(V_p) = \varrho(0) = 0$. One should realise, however, that (because $\varrho$ need not be injective) there can exist elements $Z$ of $T^\varrho P$ for which $\varrho^1(Z)$ is vertical on $P$, but $Z$ itself is not vertical. Such elements are of the form $(\mathsf{v}, V_p)$, with $\mathsf{v}$ in the kernel of $\varrho$.

We will also need the fibre linear map $j : T^\varrho P \to \mu^*\mathsf{V}$, defined by

$$j(\mathsf{v}, X_p) = (p, \mathsf{v}) \tag{7.5}$$

(we will use the same notation for its extension $j : \mathrm{Sec}(\mu^1) \to \mathrm{Sec}(\mu^*\tau)$ to sections). $j$ is surjective and its kernel is $\mathcal{V}^\varrho P$. Therefore, we have the following short exact sequence of vector bundles:

$$0 \to \mathcal{V}^\varrho P \to T^\varrho P \xrightarrow{j} \mu^*\mathsf{V} \to 0. \tag{7.6}$$

Notice also that, if we have two bundles $\mu_i : P_i \to M$ $(i = 1, 2)$, and a bundle map $F$ (over the identity on $M$) between them, the tangent map $TF : TP_1 \to TP_2$ extends to a map $T^\varrho P_1 \to T^\varrho P_2 : (\mathsf{v}, X_p) \mapsto (\mathsf{v}, TF(X_p))$. Indeed, we have $T\mu_2(TF(X_p)) = T\mu_1(X_p) = \varrho(\mathsf{v})$. There exists even a generalised notion of tangent map between prolongations with different anchor maps. Suppose that two vector bundles $\mathsf{V}_1 \to M_1$ and $\mathsf{V}_2 \to M_2$ with anchors $\varrho_1$ and $\varrho_2$ (respectively) are given, together with the above two arbitrary fibre bundles $\mu_1$ and $\mu_2$. Suppose further that $F : P_1 \to P_2$ is a bundle map over some $f : M_1 \to M_2$ and that $\mathsf{f} : \mathsf{V}_1 \to \mathsf{V}_2$ is a vector bundle morphism over the same $f$, satisfying $Tf \circ \varrho_1 = \varrho_2 \circ \mathsf{f}$. Then, we can define a map

$$T^{\varrho_1,\varrho_2}F : T^{\varrho_1}P_1 \to T^{\varrho_2}P_2, (\mathsf{v}_1, X_1) \mapsto (\mathsf{f}(\mathsf{v}_1), TF(X_1)). \tag{7.7}$$

There is more to say about the tangent-bundle-like behaviour of $T^\varrho P$, but we will not elaborate on that here.

Let us have a look at the local description of sections of $\mu^1$. Let $\{\mathsf{e}_a\}$ be a basis for $\mathrm{Sec}(\tau)$. A section $\mathcal{Z}$ of $\mu^1$ is completely determined once we know the maps $\mu^2 \circ \mathcal{Z} : P \to \mathsf{V}$ and $\varrho^1 \circ \mathcal{Z} : P \to TP$. For example, let $p \in P$ be a point with coordinates $(x^I, u^A)$, so that $(x^I)$ are the coordinates of $\mu(p) \in M$. If then $\mathcal{Z}$ is a section of $\mu^1$, we will have:

$$\begin{aligned} \pi_2 \circ \mathcal{Z} : &\qquad (x,u) \longmapsto (x, \mathsf{Y}^a(x,u)), \\ \varrho^1 \circ \mathcal{Z} : &\qquad (x,u) \longmapsto \left( \rho^I_a \mathsf{Y}^a(x,u) \frac{\partial}{\partial x^I} + X^A(x,u) \frac{\partial}{\partial u^A} \right)\Big|_p , \end{aligned}$$

and determining $\mathcal{Z}$ in coordinates of course amounts to specifying the functions $(\mathsf{Y}^a, X^A)$ on $P$.

It is worthwhile looking at the representation of such a $\mathcal{Z}$ with respect to suitably selected local sections of $\mu^1$, which exhibit the vector bundle structure of $\mu^1$ and are adapted to the basis $\{\mathsf{e}_a\}$ of $\mathrm{Sec}(\tau)$. To this end, we introduce two sets of local sections $\mathcal{X}_a$ and $\mathcal{V}_A$ of $\mu^1$ which will span $\mathrm{Sec}(\mu^1)$. The $\mathcal{V}_A$ span 'vertical sections' (whose set will be denoted by $\mathrm{Ver}(\mu^1)$) and are determined by: $\mu^2 \circ \mathcal{V}_A = 0$, while we let $\varrho^1 \circ \mathcal{V}_A$ be $\frac{\partial}{\partial u^A}$. Verticality is an intrinsic property whereas, as usual, there is no intrinsic notion of horizontality. The determination of the sections $\mathcal{X}_a$ will therefore rely on pure coordinate arguments. For the projection onto $\mathsf{V}$ we put $\mu^2 \circ \mathcal{X}_a = \mathsf{e}_a \circ \mu$ and then, fixing $\varrho^1 \circ \mathcal{X}_a$ (as a vector field on $P$) further requires making a prescription for the vertical components, which we simply take to be zero. Thus we have:

$$\mathcal{X}_a(p) = \left( \mathsf{e}_a(\mu(p)),\, \varrho^I_a(x) \frac{\partial}{\partial x^I}\Big|_p \right) \qquad \mathcal{V}_A(p) = \left( \mathsf{o}(\mu(p)),\, \frac{\partial}{\partial u^A}\Big|_p \right), \tag{7.8}$$

$\mathsf{o}$ being the zero section of $\tau$.

The above general section $\mathcal{Z}$ of $\mu^1$ then has the local representation:

$$\mathcal{Z} = \mathsf{Y}^a(x,u)\mathcal{X}_a + X^A(x,u)\mathcal{V}_A. \tag{7.9}$$

## 7.3 Lie algebroid prolongation of a fibre bundle

Suppose now that the vector bundle $\tau : \mathsf{V} \to M$ is a Lie algebroid with anchor $\varrho$. We wish to establish that there is an induced Lie algebroid structure on the vector bundle $\mu^1$ with anchor map $\varrho^1$. The idea is to define the new bracket by requiring roughly that its two projections are determined by the known brackets of the projected sections. But there are some technical complications which we will address now.

For $\mathcal{Z}_1, \mathcal{Z}_2 \in \mathrm{Sec}(\mu^1)$, a preliminary observation is that the Lie bracket of their image under $\varrho^1$ (which gives rise to vector fields on $P$), belongs to the image of $\varrho^1$. A coordinate calculation can confirm this. Putting

$$\varrho^1(\mathcal{Z}_i) = \mathsf{Y}_i^a \varrho_a^J \frac{\partial}{\partial x^J} + X_i^A \frac{\partial}{\partial u^A},$$

we have

$$\begin{aligned}[][\varrho^1(\mathcal{Z}_1), \varrho^1(\mathcal{Z}_2)] &= \Big(\varrho^1(\mathcal{Z}_1)(\mathsf{Y}_2^a) - \varrho^1(\mathcal{Z}_2)(\mathsf{Y}_1^a)\Big)\varrho_a^J \frac{\partial}{\partial x^J} \\ &\quad + \Big(\mathsf{Y}_2^a \varrho^1(\mathcal{Z}_1)(\varrho_a^J) - \mathsf{Y}_1^a \varrho^1(\mathcal{Z}_2)(\varrho_a^J)\Big)\frac{\partial}{\partial x^J} + \cdots \frac{\partial}{\partial u^A}.\end{aligned}$$

The first term on the right manifestly belongs to the image of $\varrho^1$, whereas the last term is irrelevant for that purpose. The middle term can be rewritten as

$$\mathsf{Y}_2^a \mathsf{Y}_1^b \left(\varrho_b^J \frac{\partial \varrho_a^I}{\partial x^J} - \varrho_a^J \frac{\partial \varrho_b^I}{\partial x^J}\right)\frac{\partial}{\partial x^I},$$

which is seen to belong to the image of $\varrho^1$ in view of the fact that $\varrho$ is a Lie algebra homomorphism (see e.g. expression (4.4)). It is therefore natural to impose right away that the bracket $[\cdot,\cdot]^1$ under construction, which of course is required to be skew-symmetric and $\mathbb{R}$-bilinear, should satisfy

$$\varrho^1\left([\mathcal{Z}_1, \mathcal{Z}_2]^1\right) = [\varrho^1(\mathcal{Z}_1), \varrho^1(\mathcal{Z}_2)]. \tag{7.10}$$

This will have for consequence that for $F_i \in C^\infty(P)$,

$$\begin{aligned}\varrho^1\left([F_1\,\mathcal{Z}_1, F_2\,\mathcal{Z}_2]^1\right) &= \\ &F_1 F_2\,[\varrho^1(\mathcal{Z}_1), \varrho^1(\mathcal{Z}_2)] + F_1\,\varrho^1(\mathcal{Z}_1)(F_2)\,\varrho^1(\mathcal{Z}_2) - F_2\,\varrho^1(\mathcal{Z}_2)(F_1)\,\varrho^1(\mathcal{Z}_1).\end{aligned}$$

It remains then to make sure that the projection under $\mu^2$ can be specified in a compatible way. The above coordinate calculation to some extent illustrates how one should proceed. If we apply $T\mu$ to the preceding equality, we get (pointwise)

$$\begin{aligned} T\mu\left(\varrho^1\left([F_1\,\mathcal{Z}_1, F_2\,\mathcal{Z}_2]^1\right)\right) &= (F_1F_2)\,T\mu\left([\varrho^1(\mathcal{Z}_1), \varrho^1(\mathcal{Z}_2)]\right) \\ &+ F_1\,\varrho^1(\mathcal{Z}_1)(F_2)\,\varrho(\mu^2(\mathcal{Z}_2)) - F_2\,\varrho^1(\mathcal{Z}_2)(F_1)\,\rho(\mu^2(\mathcal{Z}_1)). \end{aligned}$$

In general, the $\varrho(\mu^2(\mathcal{Z}_i))$ are vector fields along $\mu$ for which there is no standard Lie bracket available. If the $\mathcal{Z}_i$ are projectable, however, meaning that there exist sections $\mathsf{s}_i$ of $\tau$ such that $\mu^2 \circ \mathcal{Z}_i = \mathsf{s}_i \circ \mu$, then the vector fields $\varrho^1(\mathcal{Z}_i)$ on $P$ are $\mu$-related to the vector fields $\varrho(\mathsf{s}_i)$ on $M$. Hence, the corresponding brackets are also $\mu$-related, meaning that for projectable $\mathcal{Z}_i$, we can put

$$\mu^2\left([\mathcal{Z}_1, \mathcal{Z}_2]^1\right) = [\mu^2(\mathcal{Z}_1), \mu^2(\mathcal{Z}_2)], \tag{7.11}$$

and then the property that $\varrho$ is a Lie algebra homomorphism (4.2) (which in coordinates gives (4.4)) ensures that

$$T\mu\left(\varrho^1\left([\mathcal{Z}_1, \mathcal{Z}_2]^1\right)\right) = \varrho \circ \mu_2\left([\mathcal{Z}_1, \mathcal{Z}_2]^1\right)$$

as it should. The expression for $T\mu\left(\varrho^1\left([F_1\,\mathcal{Z}_1, F_2\,\mathcal{Z}_2]^1\right)\right)$ further shows that the $\mu^2$ and $\varrho^1$ projections of the bracket under construction will still match up if for projectable $\mathcal{Z}_i$ and for any $F_i \in C^\infty(P)$, we define

$$\begin{aligned} \mu^2\left([F_1\,\mathcal{Z}_1, F_2\,\mathcal{Z}_2]^1\right) &= F_1F_2\,[\mu^2(\mathcal{Z}_1), \mu^2(\mathcal{Z}_2)] \\ &+ F_1\,\varrho^1(\mathcal{Z}_1)(F_2)\,\mu^2(\mathcal{Z}_2) - F_2\,\varrho^1(\mathcal{Z}_2)(F_1)\,\mu^2(\mathcal{Z}_1). \end{aligned} \tag{7.12}$$

It then follows that

$$[F_1\,\mathcal{Z}_1, F_2\,\mathcal{Z}_2]^1 = F_1F_2\,[\mathcal{Z}_1, \mathcal{Z}_2]^1 + F_1\,\varrho^1(\mathcal{Z}_1)(F_2)\,\mathcal{Z}_2 - F_2\,\varrho^1(\mathcal{Z}_2)(F_1)\,\mathcal{Z}_1, \tag{7.13}$$

since both sides have the same $\mu^2$ and $\varrho^1$ projections.

The final point to observe now is that sections of $\mu^1$ (locally) are finitely generated, over the ring $C^\infty(P)$, by projectable sections. Hence, the defining relations (7.10) and (7.12) are sufficient to define the bracket $[\cdot,\cdot]^1$ on vector sections. The property (7.13) will hold by extension for all sections and the bracket will satisfy the Jacobi identity as a result of the Jacobi identity of the Lie algebroid bracket we started from and the same identity for vector fields on $P$. This concludes the construction of the prolonged Lie algebroid. From now on, we simply write $[\cdot,\cdot]$ for $[\cdot,\cdot]^1$. For completeness we summarise our result in the following proposition.

**Proposition 7.4.** *The prolongation of $P$ inherits a Lie algebroid structure from the one on $\mathsf{V}$ and the standard one on $TP$. The anchor map is $\varrho^1 \colon T^\varrho P \to TP, \varrho^1(\mathsf{v}, X_p) = X_p$, and the bracket can be defined in terms of projectable sections as follows. If $\mathcal{Z}_1$, $\mathcal{Z}_2$ are two projectable sections of $\mu^1$ given by $\mathcal{Z}_k(p) = (\mathsf{s}_k(m), X_k(p))$, $k = 1,2$ for some sections $\mathsf{s}_k$ of $\tau$ and $X_k$ of $\tau_P$, then the bracket $[\mathcal{Z}_1, \mathcal{Z}_2]$ is the section given by*

$$[\mathcal{Z}_1, \mathcal{Z}_2](p) = ([\mathsf{s}_1, \mathsf{s}_2](m), [X_1, X_2](p)).$$

For computational purposes, it remains to list the brackets of the local sections which are used in the general representation of a section of $\mu^1$ as in (7.9). Suppose that the brackets of the Lie algebroid on $\tau$ are given by (4.3). We thus have

$$[\mathcal{X}_a, \mathcal{X}_b] = \mathsf{C}^{\,c}_{ab}\mathcal{X}_c, \quad [\mathcal{X}_a, \mathcal{V}_A] = 0, \quad [\mathcal{V}_A, \mathcal{V}_B] = 0.$$

It is perhaps worthwhile to repeat hereby that both projections have to be looked at to verify these statements, although of course they are bound to match up if our new bracket has been defined consistently. Thus we have, for example:

$$\begin{aligned} \mu^2([\mathcal{X}_a, \mathcal{X}_b]) &= [\mu^2(\mathcal{X}_a), \mu^2(\mathcal{X}_b)] = [\mathsf{e}_a, \mathsf{e}_b] = \mathsf{C}^{\,c}_{ab}\mathsf{e}_c, \\ \varrho^1([\mathcal{X}_a, \mathcal{X}_b]) &= [\varrho^1(\mathcal{X}_a), \varrho^1(\mathcal{X}_b)] = \left[\varrho^I_a \frac{\partial}{\partial x^I}, \varrho^J_b \frac{\partial}{\partial x^J}\right] = \mathsf{C}^{\,c}_{ab}\,\varrho^1(\mathcal{X}_c), \end{aligned}$$

where (4.4) has been used again in the last line.

Let us have a look now at the exterior differential which is associated to the Lie algebroid structure on $\mu^1$. It is determined by

$$dx^I = \varrho^I_a \mathcal{X}^a, \qquad du^A = \mathcal{V}^A, \tag{7.14}$$

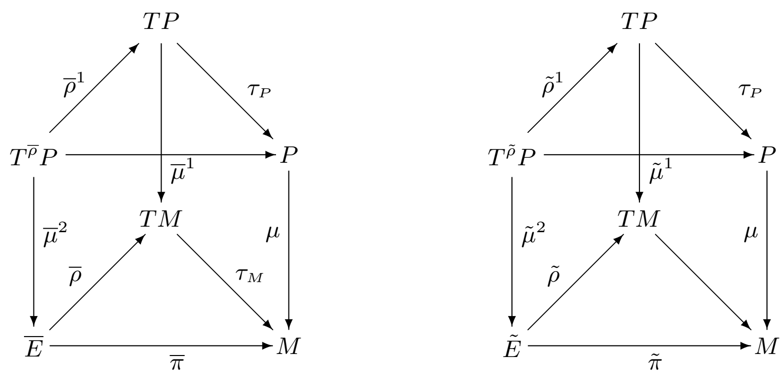
and

$$d\mathcal{X}^c = -\frac{1}{2}\mathsf{C}^{\,c}_{ab}\mathcal{X}^a \wedge \mathcal{X}^b, \qquad d\mathcal{V}^A = 0, \tag{7.15}$$

where $\{\mathcal{X}^a, \mathcal{V}^A\}$ denotes the local basis of $(\mu^1)^*$ that is dual to $\{\mathcal{X}_a, \mathcal{V}_A\}$.

## 7.4 Affine Lie algebroid prolongation of a fibre bundle

As a first step towards the situation we are most interested in, we consider the case where the vector bundle $\mathsf{V}$ is either $\overline{E}$ or $\tilde{E}$ and the anchor map is respectively $\overline{\rho}$ and $\tilde{\rho}$ (associated to an affine bundle map $\rho : E \to TM$), but $\mu$ remains an arbitrary fibre bundle. The notations will be as in the following diagrams.




Suppose that there is a Lie algebroid structure on the affine bundle $\pi$ and consider the associated Lie algebroid structure on $\tilde{\pi}$, described in Proposition 6.4. We will show now that also the Lie algebroid structure on $\tilde{\mu}^1 : T^{\tilde{\rho}}P \to P$ then is associated to an affine Lie algebroid. Let the maps $\bar{\imath}$ and $\jmath$ be as in (6.1) and (6.2).

**Proposition 7.5.** *Let $\overline{I}\colon T^{\overline{\rho}}P \to T^{\tilde{\rho}}P$ and $J\colon T^{\tilde{\rho}}P \to P \times I\!R$ be the maps*

$$\overline{I}(\overline{e}, X_p) = (\bar{\imath}(\overline{e}), X_p) \qquad \textit{and} \qquad J(\tilde{e}, X_p) = (p, \jmath_m(\tilde{e})),$$

*with $m = \mu(p)$. Then the following sequence of vector bundles is exact*

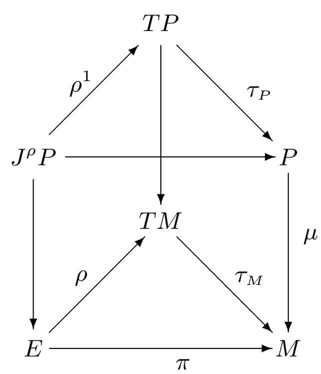
$$0 \longrightarrow T^{\overline{\rho}}P \xrightarrow{\overline{I}} T^{\tilde{\rho}}P \xrightarrow{J} P \times I\!R \longrightarrow 0.$$

PROOF: Since $\jmath_m \circ \bar{\imath} = 0$, we clearly have that $J \circ \overline{I} = 0$, so that $\mathrm{Im}(\overline{I}) \subset \mathrm{Ker}(J)$. Moreover, if $(\tilde{e}, X_p) \in \mathrm{Ker}(J)$ then $\jmath(\tilde{e}) = 0$; hence there exists a $\overline{e}$ such that $\tilde{e} = \bar{\imath}(\overline{e})$, so that $(\tilde{e}, X_p) = (\bar{\imath}(\overline{e}), X_p) = \overline{I}(\overline{e}, X_p)$ is in $\mathrm{Im}(\overline{I})$. □

Therefore, following the reasoning at the end of Corollary 5.4, the set of points $J^{-1}(P \times \{1\})$ is an affine bundle whose associated vector bundle is

$\overline{\mu}^1 : T^{\overline{\rho}}P \to P$. We will use the notation $J^\rho P \to P$ for this affine bundle, where

$$J^\rho P = \{(\iota(e), X_p) \in T^{\tilde{\rho}}P\} \simeq \{\, (e, X_p) \in E \times TP \mid \rho(e) = T\mu(X_p) \,\}. \quad (7.16)$$




**Proposition 7.6.** *The Lie algebroid structure of $T^{\tilde{\rho}}P$ restricts to $J^\rho P$, defining therefore a Lie algebroid structure on that affine bundle.*

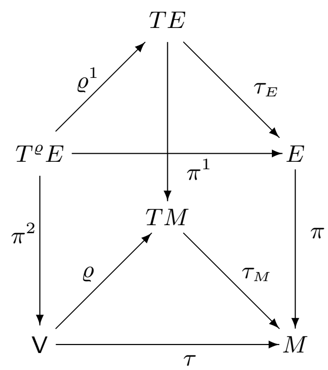
PROOF: To see this, we have to prove that the bracket of sections of $J^\rho P$ is a section of $T^{\overline{\rho}}P$. For that it is enough to consider projectable sections, since they form a generating set. If $\mathcal{Z}_1$, $\mathcal{Z}_2$ are sections of $J^\rho P$, projecting onto sections $\sigma_1$ and $\sigma_2$ of $E$, then for every $p \in P$ we have

$$[\mathcal{Z}_1, \mathcal{Z}_2](p) = ([\sigma_1, \sigma_2](m), [X_1, X_2](p)), \qquad \text{with } m = \mu(p),$$

which is an element of $T^{\overline{\rho}}P$, since $[\sigma_1, \sigma_2](m) \in \overline{E}_m$. □

## 7.5 $\varrho$-prolongation of an affine bundle

Let us come back again to the original set-up of the prolongation idea: $\tau : \mathsf{V} \to M$ is a vector bundle, not necessarily related to the affine bundle $\pi : E \to M$. We will now specialise to the case where $\mu$ is precisely the affine bundle $\pi$ and show that it is possible to extend the notion of vertical lift.




Recall that in Section 6.1 we have defined the vertical lift of $(e, \overline{e}) \in \pi^*\overline{E}$ for a point $e \in E_m$ and a vector $\overline{e} \in \overline{E}_m$ as the vector $v(e, \overline{e}) \in T_eE$ (see expression (6.7)) and further also the vertical lift of an element $\tilde{e} \in \tilde{E}_m$ to the point $e \in E_m$ as the tangent vector $v(e, \tilde{e})$ given by expression (6.8). We can now extend the map $v$ to a map ${}^V : \pi^*\tilde{\pi} \to T^{\varrho}E$, by means of

$$(e, \tilde{e})^V = (0, v(e, \tilde{e})). \tag{7.17}$$

The final step in the construction now is obvious: if $\tilde{\zeta}$ is a section of $\tilde{E}$, we define the section $\tilde{\zeta}^V$ of $T^{\varrho}E$, called the vertical lift of $\tilde{\zeta}$, by putting

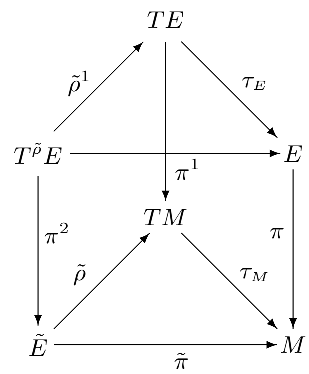
$$\tilde{\zeta}^V(e) = (e, \tilde{\zeta}(m))^V, \qquad \text{with } e \in E \text{ and } m = \pi(e).$$

In coordinates, if $\tilde{\zeta} = \zeta^0 e_0 + \zeta^\alpha e_\alpha$ then

$$\tilde{\zeta}^V = (\zeta^\alpha - y^\alpha \zeta^0)\mathcal{V}_\alpha.$$

## 7.6 Affine Lie algebroid prolongation of an affine bundle

The final step is the case where the vector bundle $\tau$ is just the bundle $\tilde{\pi}$. This is the space where Lagrangian-type systems can be defined, as will be shown below. So the final diagram under consideration is the following.




We will suppose in this section that there exists an affine Lie algebroid structure on $\pi$, and thus (in view of Propositions 6.4 and Proposition 7.4) a Lie algebroid structure on the vector bundles $\tilde{\pi} : \tilde{E} \to M$ and $\pi^1 : T^{\tilde{\rho}}E \to E$. Furthermore, Proposition 7.6 shows that the Lie algebroid on $\pi^1$ is related to a Lie algebroid structure on the affine bundle $J^\rho E \to E$. If, in particular, $M$ is fibred over $I\!R$ and $\rho$ is of the form $\iota_{J^1M} \circ \lambda$, the question raised at the end of Section 7.1 is seen to have a positive answer, with $J^\lambda E$ from (7.2) coinciding with the space $J^\rho E$ constructed in (7.16). In Section 7.1 we have identified this affine bundle as the environment needed to study the dynamical systems of the pseudo-SODE-form (4.10). There is however a subtle point which needs to be addressed. For the purpose of defining (geometrically) time-dependent systems of Lagrangian type, one can construct all the necessary tools on the affine prolongation $J^\lambda E \to E$. However, we wish to discuss in the forthcoming chapters aspects such as the non-linear and linear connections which are naturally associated to a pseudo-SODE on an affine bundle $\pi$, and for that purpose, even specifically for the time-dependent framework, it turns out to be more appropriate to use the prolongation structure $T^{\tilde{\rho}}E \to E$, which is a vector bundle, rather than an affine bundle.

So, for that reason, we will reformulate the concepts of 'admissibility' and 'pseudo-SODE', on the appropriate bundle and for arbitrary $\rho : E \to TM$ in Section 7.8. Before arriving there, however, it is necessary to develop some more technical tools.

## 7.7 Admissible elements, vertical endomorphism and complete lift

An important concept for the study of dynamical systems on affine Lie algebroids is that of admissible elements. An element $Z \in T^{\tilde{\rho}}E$ is said to be *admissible* if, roughly speaking, it has the same projection under $\pi^1$ and $\pi^2$. More precisely, we set

$$\mathrm{Adm}(E) = \{ Z \in T^{\tilde{\rho}}E \,|\, \pi^2(Z) = \iota(\pi^1(Z)) \}.$$

An equivalent characterisation uses the map $j : T^{\tilde{\rho}}E \to \pi^*\tilde{E}$ (defined in (7.5)): $Z$ is admissible if and only if it belongs to $J^\rho E$ and the canonical map $\vartheta$ (defined in Section 6.1 in Chapter 6) vanishes on its projection onto $\pi^*\tilde{E}$ under $j$, hence

$$\mathrm{Adm}(E) = \{ Z \in J^\rho E \,|\, \vartheta(j(Z)) = 0 \}.$$

Indeed, if $Z$ is of the form $Z = (\tilde{e}, X_e)$, then the first condition means that $\tilde{e}$ is in the image of $\iota$ and the second then further specifies that $\tilde{e} = \iota(e)$.

By a *contact 1-form* we mean a 1-form $\theta$ on $T^{\tilde{\rho}}E$ (i.e. a $C^\infty(E)$-linear map from sections of $\pi^1$ to $C^\infty(E)$), which vanishes on sections whose image lies in $\mathrm{Adm}(E)$. It follows from the characterisation of admissible elements that contact forms are locally spanned by

$$\theta^\alpha = \mathcal{X}^\alpha - y^\alpha \mathcal{X}^0.$$

Any 1-form $\theta$ on $\tilde{E}$ determines a contact 1-form $\overrightarrow{\theta}$ by means of the canonical map: If $\mathcal{Z} \in \mathrm{Sec}(\pi^1)$, then

$$\langle \overrightarrow{\theta}, \mathcal{Z} \rangle = \langle \theta, \vartheta(j(\mathcal{Z})) \rangle.$$

In coordinates, if $\theta$ is of the form $\theta = \theta_0 e^0 + \theta_\alpha e^\alpha$ then

$$\overrightarrow{\theta} = \theta_\alpha(\mathcal{X}^\alpha - y^\alpha \mathcal{X}^0).$$

Notice that the elements of the basis $\{\theta^\alpha\}$ of contact 1-forms are of this type: $\theta^\alpha = \overrightarrow{e^\alpha}$. We further will need the affine function $\overleftarrow{\theta} \in C^\infty(E)$ associated to $\iota^*(\theta)$. To be precise, there is of course a linear function on $\tilde{E}$ associated to $\theta$, but we will reserve the notation $\overleftarrow{\theta}$ for its restriction to $E$, meaning that in coordinates:

$$\overleftarrow{\theta} = \theta_0 + \theta_\alpha y^\alpha.$$

With these definitions, we can split (the pullback of) a 1-form $\theta$ on $\tilde{E}$ as follows

$$(\pi^2)^*\theta = \overleftarrow{\theta}\,\mathcal{X}^0 + \overrightarrow{\theta}.$$

This decomposition is important for various calculations in one of the next sections.

Given a section $\tilde{\zeta} \in \mathrm{Sec}(\tilde{\pi})$ we can now define a section $\tilde{\zeta}^C$ of $\pi^1$ which is called the complete lift of $\tilde{\zeta}$. It is defined by the following two conditions which completely characterise it:

- $\tilde{\zeta}^C$ projects to $\tilde{\zeta}$, i.e. $\pi^2 \circ \tilde{\zeta}^C = \tilde{\zeta} \circ \pi$, and
- $\tilde{\zeta}^C$ preserves the set of contact forms, that is, if $\theta$ on $T^{\tilde{\rho}}E$ is a contact form then $d_{\tilde{\zeta}^C}\theta = [i_{\tilde{\zeta}^C}, d]\theta$ is contact.

In the case of the pullback of a 1-form on $\tilde{E}$, making use of the decomposition as sum of a contact plus a non-contact form $(\pi^2)^*\theta = \overleftarrow{\theta}\,\mathcal{X}^0 + \overrightarrow{\theta}$, and taking into account that $\tilde{\zeta}^C$ projects onto $\tilde{\zeta}$ and that $\pi^2$ is a morphism of Lie algebroids, one can verify that

$$\begin{aligned} d_{\tilde{\zeta}^C}\overrightarrow{\theta} &= \overrightarrow{d_{\tilde{\zeta}}\theta} + \overleftarrow{\theta}\,\overrightarrow{d_{\tilde{\zeta}}e^0}, \\ d_{\tilde{\zeta}^C}\overleftarrow{\theta} &= \overleftarrow{d_{\tilde{\zeta}}\theta} - \overleftarrow{\theta}\,\overleftarrow{d_{\tilde{\zeta}}e^0}. \end{aligned}$$

In fact, any of these two conditions is equivalent to the second condition in our definition of complete lift.

The coordinate expression of the complete lift of the section $\tilde{\zeta} = \zeta^0 e_0 + \zeta^\alpha e_\alpha$ is

$$\tilde{\zeta}^C = \zeta^0\mathcal{X}_0 + \zeta^\alpha\mathcal{X}_\alpha + [(\dot{\zeta}^\alpha - y^\alpha\dot{\zeta}^0) + C^\alpha_\beta(\zeta^\beta - y^\beta\zeta^0)]\mathcal{V}_\alpha,$$

where $C^\alpha_\beta = C^\alpha_{0\beta} + C^\alpha_{\gamma\beta}y^\gamma$ and, for a function $f \in C^\infty(M)$, the complete lift $\dot{f} \in C^\infty(E)$ is defined by $\dot{f} = \overleftarrow{df}$. The first two terms of $\tilde{\zeta}^C$ are determined by the projectability condition, whereas the third term can be obtained by applying the preceding formula to $\theta = e^\alpha$.

The vertical and complete lift satisfy the following properties

$$\begin{aligned} d_{\tilde{\zeta}^V} f &= 0 & \qquad d_{\tilde{\zeta}^V}\overleftarrow{\theta} &= i_{\tilde{\zeta}^C}\overrightarrow{\theta} \\ d_{\tilde{\zeta}^C} f &= d_{\tilde{\zeta}} f & \qquad d_{\tilde{\zeta}^C}\overleftarrow{\theta} &= \overleftarrow{d_{\tilde{\zeta}}\theta} - \overleftarrow{\theta}\,\overleftarrow{d_{\tilde{\zeta}}e^0} \end{aligned}$$

for $f \in C^\infty(M)$ and $\theta$ a 1-form on $\tilde{E}$. We prove only the third; if $\overline{a} = \vartheta_a(\tilde{\zeta}(a))$ then

$$\begin{aligned} d_{\tilde{\zeta}^V} \overleftarrow{\theta}(a) &= v(a, \tilde{\zeta}(m))\overleftarrow{\theta} = \frac{d}{dt}\overleftarrow{\theta}(a + t\overline{a})\big|_{t=0} = \frac{d}{dt}(\overleftarrow{\theta}(a) + t\langle \overrightarrow{\theta}_m, \overline{a}\rangle)\big|_{t=0} \\ &= \langle \overrightarrow{\theta}_m, \overline{a}\rangle = \langle \overrightarrow{\theta}_m, \vartheta_a(\tilde{\zeta}(a))\rangle = (i_{\tilde{\zeta}^C}\overrightarrow{\theta})(a). \end{aligned}$$

Using the above equations it is a matter of a routine calculation to prove the following commutation relations.

$$\begin{aligned} &[\tilde{\zeta}_1^C, \tilde{\zeta}_2^C] = [\tilde{\zeta}_1, \tilde{\zeta}_2]^C \\ &[\tilde{\zeta}_1^C, \tilde{\zeta}_2^V] = [\tilde{\zeta}_1, \tilde{\zeta}_2]^V + \langle \tilde{\zeta}_1, e^0\rangle \tilde{\zeta}_2^V \\ &[\tilde{\zeta}_1^V, \tilde{\zeta}_2^V] = \langle \tilde{\zeta}_1, e^0\rangle \tilde{\zeta}_2^V - \langle \tilde{\zeta}_2, e^0\rangle \tilde{\zeta}_1^V. \end{aligned}$$

The above definitions and relations are greatly simplified if we restrict to sections of the associated vector bundle $\overline{E}$. Indeed, if $\overline{\sigma} = \sigma^\alpha \overline{e}_\alpha$ is a section of $\overline{E}$, then the complete and vertical lifts of $\overline{\iota}(\overline{\sigma})$ have the coordinate expressions

$$\overline{\sigma}^C = \sigma^\alpha \mathcal{X}_\alpha + (\dot{\sigma}^\alpha + C^\alpha_\beta \sigma^\beta)\mathcal{V}_\alpha \qquad \overline{\sigma}^V = \sigma^\alpha \mathcal{V}_\alpha,$$

and the action of the complete lift over linear functions and contact forms is given by

$$d_{\overline{\sigma}^C}\overrightarrow{\theta} = \overrightarrow{d_{\overline{\sigma}}\theta} \qquad d_{\overline{\sigma}^C}\overleftarrow{\theta} = \overleftarrow{d_{\overline{\sigma}}\theta},$$

since $d_{\overline{\sigma}} e^0 = 0$. Furthermore, the commutation relations are as in the usual vector Lie algebroid case:

$$[\overline{\sigma}^C, \overline{\eta}^C] = [\overline{\sigma}, \overline{\eta}]^C \qquad [\overline{\sigma}^C, \overline{\eta}^V] = [\overline{\sigma}, \overline{\eta}]^V \qquad [\overline{\sigma}^V, \overline{\eta}^V] = 0.$$

Finally, in the case that $\tau = \tilde{\pi}$, we can combine the vertical lift ${}^V : \mathrm{Sec}(\pi^*\tilde{\pi}) \to \mathrm{Sec}(\pi^1)$ with the projection $j : \mathrm{Sec}(\pi^1) \to \mathrm{Sec}(\pi^*\tilde{\pi})$.

**Definition 7.7.** *The map $S = {}^V \circ j$ is called the vertical endomorphism on* $\mathrm{Sec}(\pi^1)$.

In coordinates, the type (1,1) tensor field $S$ reads

$$S = (\mathcal{X}^\alpha - y^\alpha \mathcal{X}^0) \otimes \mathcal{V}_\alpha.$$

It follows that

$$S(\tilde{\zeta}^C) = \tilde{\zeta}^V \qquad \text{and} \qquad S(\tilde{\zeta}^V) = 0.$$

## 7.8 $\rho$-admissible curves and dynamics

In this section, we show how the dynamical systems on $E$ (4.14) of the pseudo-SODE type, can be defined in the way standard second-order differential equations on a tangent bundle or first-jet bundle are conceived, although they do not necessarily correspond, locally, to second-order equations.

A curve $\gamma\colon \mathbb{R} \to E$ is said to be *admissible* if

$$\rho \circ \gamma = \dot{\gamma}_M,$$

where $\gamma_M = \pi \circ \gamma$ is the projected curve to the base. In coordinates, if $\gamma(u) = (x^I(u), y^\alpha(u))$ then $\gamma$ is admissible if

$$\dot{x}^I(u) = \rho^I_0(x(u)) + \rho^I_\alpha(x(u))y^\alpha(u).$$

A curve is admissible if and only if its prolongation takes values in the set of admissible elements. Indeed, the prolongation of the curve $\gamma$ is the curve $\gamma^c(u) = (\iota(\gamma(u)), \dot{\gamma}(u))$, and this is an admissible element if and only if it is in $T^{\tilde{\rho}}E$, that is $\rho(\gamma(u)) = T\pi(\dot{\gamma})(u) = \dot{\gamma}_M(u)$.

**Definition 7.8.** *A pseudo-*SODE *on $E$ is a section $\Gamma$ of* Adm$(E)$*, that is, a section of $T^{\tilde{\rho}}E$ which takes values in the set of admissible elements.*

From this definition, it readily follows that $\langle \Gamma, \mathcal{X}^0 \rangle = 1$ and that the integral curves of $\tilde{\rho}^1(\Gamma)$ are admissible curves. Conversely, any section $\mathcal{Z}$ of $T^{\tilde{\rho}}E$ such that the integral curves of $\tilde{\rho}^1(\mathcal{Z})$ are admissible is a pseudo-SODE. From the alternative characterisation of Adm$(E)$ as the set of elements of $J^\rho E$ which vanish under $\vartheta$, it follows that:

**Corollary 7.9.** *A section $\Gamma$ of $T^{\tilde{\rho}}E$ is a pseudo-*SODE *if and only if $S(\Gamma) = 0$ and $\langle \Gamma, \mathcal{X}^0 \rangle = 1$.*

Locally, a SODE $\Gamma$ is of the form

$$\Gamma = \mathcal{X}_0 + y^\alpha \mathcal{X}_\alpha + f^\alpha \mathcal{V}_\alpha. \tag{7.18}$$

and the vector field $\tilde{\rho}^1(\Gamma)$ is of the form

$$\tilde{\rho}^1(\Gamma) = (\rho^I_0 + \rho^I_\alpha y^\alpha)\frac{\partial}{\partial x^I} + f^\alpha \frac{\partial}{\partial y^\alpha}.$$

Evidently, its integral curves are solutions of the dynamical system

$$\begin{aligned} \dot{x}^I &= \rho^I_\alpha(x)y^\alpha + \rho^I_0(x), \\ \dot{y}^\alpha &= f^\alpha(x,y). \end{aligned} \tag{7.19}$$

## 7.9 Lagrange-type systems on an affine Lie algebroid

Remark that nowhere in the definition of a pseudo-SODE we required that the affine bundle possessed a Lie algebroid structure. We shall now subsequently discuss a class of pseudo-SODEs, which come from a (constrained) variational problem and therefore are said to be of Lagrangian type. To define this Lagrangian-type equations in a coordinate free way, we can simply mimic the usual construction on a first-jet bundle. To do so, we need to require now that $\pi$ is an affine Lie algebroid.

For a given function $L$ on $E$, we define the *Cartan 1-form* $\Theta_L$ on $T^{\tilde{\rho}}E$ by

$$\Theta_L = dL \circ S + L\mathcal{X}^0$$

and the *Cartan 2-form* $\Omega_L$ by $\Omega_L = -d\Theta_L$. We say that a Lagrangian $L$ is *regular* if the matrix $\left(\frac{\partial^2 L}{\partial y^\alpha \partial y^\beta}\right)$ is non-singular at every point.

**Proposition 7.10.** *If $L$ is a regular Lagrangian, then there is a unique section $\Gamma$ of $\pi^1$ such that*

$$i_\Gamma \Omega_L = 0 \qquad \text{and} \qquad \langle \Gamma, \mathcal{X}^0 \rangle = 1.$$

*Moreover, the section $\Gamma$ is a pseudo-*SODE *and is said to be of Lagrangian type.*

PROOF: We first calculate coordinate expressions for the Cartan forms. For the first Cartan form we get

$$\Theta_L = \frac{\partial L}{\partial y^\alpha}\theta^\alpha + L\mathcal{X}^0,$$

while the expression of $\Omega_L$ is given by

$$\begin{aligned}\Omega_L &= d\Big(\frac{\partial L}{\partial y^\alpha}\Big) \wedge \theta^\alpha + \frac{\partial L}{\partial y^\alpha} d\theta^\alpha + dL \wedge \mathcal{X}^0 + L d\mathcal{X}^0 \\ &= \theta^\alpha \wedge \Big( (\rho^I_\alpha \frac{\partial L}{\partial x^I} + C^\beta_{0\alpha}\frac{\partial L}{\partial y^\beta})\mathcal{X}^0 - d\Big(\frac{\partial L}{\partial y^\alpha}\Big)\Big) - \frac{1}{2} C^\alpha_{\gamma\beta}\frac{\partial L}{\partial y^\alpha}\mathcal{X}^\gamma \wedge \mathcal{X}^\beta.\end{aligned}$$

(in the transition to the second line we have used (7.14) and (7.15)). Let now $\Gamma = \mathcal{X}_0 + g^\alpha \mathcal{X}_\alpha + f^\alpha \mathcal{V}_\alpha$ be a section of $\pi^1$ (taking already into account that $\langle \Gamma, \mathcal{X}^0 \rangle = 1$). It follows that

$$\begin{aligned} i_\Gamma \Omega_L &= (y^\alpha - g^\alpha)\Big( (\rho^I_\alpha \frac{\partial L}{\partial x^I} + C^\beta_{0\alpha}\frac{\partial L}{\partial y^\beta})\mathcal{X}^0 - d(\frac{\partial L}{\partial y^\alpha})\Big) \\ &\quad + \Big( \langle \Gamma, d(\frac{\partial L}{\partial y^\alpha})\rangle - (\rho^I_\alpha \frac{\partial L}{\partial x^I} + C^\beta_{0\alpha}\frac{\partial L}{\partial y^\beta})\Big)\theta^\alpha - \frac{\partial L}{\partial y^\alpha} C^\alpha_{\gamma\beta} g^\gamma \mathcal{X}^\beta \end{aligned}$$

Now, putting $i_\Gamma \Omega_L(\mathcal{Z}) = 0$ for a vertical section $\mathcal{Z}$, it follows that

$$\frac{\partial^2 L}{\partial y^\alpha \partial y^\beta}(y^\alpha - g^\alpha) = 0,$$

or, since $\left(\frac{\partial^2 L}{\partial y^\alpha \partial y^\beta}\right)$ is supposed to be non-singular, $g^\alpha = y^\alpha$. This observation proves that $\Gamma$ is a pseudo-SODE and it simplifies the 1-form $i_\Gamma \Omega_L$. Due to the skew-symmetry of the structure functions it is clear that $C^\alpha_{\gamma\beta} y^\gamma y^\beta = 0$ and we find

$$0 = i_\Gamma \Omega_L = \Big( f^\beta \frac{\partial^2 L}{\partial y^\alpha \partial y^\beta} + (\rho^I_\beta y^\beta + \rho^I_0) \frac{\partial L^2}{\partial x^I \partial y^\alpha} - \rho^I_\alpha \frac{\partial L}{\partial x^I} - (C^\beta_{0\alpha} + C^\beta_{\gamma\alpha} y^\gamma) \frac{\partial L}{\partial y^\beta} \Big) \theta^\alpha$$

The functions $f^\mu$ which now uniquely determine the pseudo-SODE are given by

$$f^\mu = g^{\mu\alpha} \left( \rho^I_\alpha \frac{\partial L}{\partial x^I} + (C^\beta_{0\alpha} + C^\beta_{\gamma\alpha} y^\gamma) \frac{\partial L}{\partial y^\beta} - (\rho^I_\beta y^\beta + \rho^I_0) \frac{\partial^2 L}{\partial x^I \partial y^\alpha} \right), \quad (7.20)$$

where $(g^{\alpha\beta})$ stands for the inverse matrix of $(g_{\alpha\beta}) = \left(\frac{\partial^2 L}{\partial y^\alpha \partial y^\beta}\right)$. □

The integral curves of $\varrho^1(\Gamma)$ are solutions of the *Lagrangian-type equations on an affine Lie algebroid*,

$$\begin{aligned} \dot{x}^I &= \rho^I_0 + \rho^I_\alpha y^\alpha, \\ \frac{d}{dt}\Big(\frac{\partial L}{\partial y^\alpha}\Big) &= \rho^I_\alpha \frac{\partial L}{\partial x^I} + (C^\beta_{0\alpha} + C^\beta_{\gamma\alpha} y^\gamma) \frac{\partial L}{\partial y^\beta}. \end{aligned} \quad (7.21)$$

In the time-dependent case where $M \to \mathbb{R}$ and $\rho = \iota_{J^1 M} \circ \lambda$ they reduce exactly to (4.17).

It is interesting to verify that the equations (7.21) also can be obtained from a geometric calculus of variations approach. We explain how this works without working out all the technical details. Given a function $L \in C^\infty(E)$ and two points $m_0$ and $m_1$ on $M$, consider the problem of determining the critical curves of the functional

$$J(\gamma) = \int_\gamma L = \int_{t_0}^{t_1} L(\gamma(t))\, dt$$

defined on the set of admissible curves $\gamma\colon [t_0, t_1] \to E$, for which $\gamma_M$ in the base manifold has fixed endpoints $m_0$ and $m_1$. This is a constrained problem, since the curves we consider are restricted to be admissible, i.e.

they have to satisfy the constraints $\dot{x}^I = \rho_0^I + \rho_\alpha^I y^\alpha$. We should therefore be more specific about the class of admissible variations we will allow; they will be generated by complete lifts of sections of $\overline{E}$, as follows.

Let $\overline{\sigma}$ be a section of $\overline{E}$ such that $\overline{\sigma}(m_0) = \overline{\sigma}(m_1) = 0$. We consider the vector fields $X = \overline{\rho}(\overline{\sigma})$ and $Y = \overline{\rho}^1(\overline{\sigma}^C)$, and we denote their flows by $\psi_s$ and $\Psi_s$, respectively. It follows that $\psi_s(m_0) = m_0$ and $\psi_s(m_1) = m_1$. The family of curves $\chi(s,t) = \Psi_s(\gamma(t))$ is a 1-parameter family of admissible variations of $\gamma$: that $\chi(s,t)$ projects onto $\chi_M(s,t) = \psi_s(\gamma_M(t))$ is obvious; the fact that $\chi(s,t)$ is an admissible curve for every fixed $s$ requires more work and is left to the reader. At the endpoints $t_0$ and $t_1$, we have

$$\chi_M(s,t_i) = \psi_s(\gamma_M(t_i)) = \psi_s(m_i) = m_i.$$

The infinitesimal variation fields we consider are of the form $Z = Y \circ \gamma$; their projection to $M$ is $W = X \circ \gamma_M$. Therefore, the variation of $L$ along $\chi(s,t)$ at $s = 0$ is given by

$$\frac{\partial(L \circ \chi)}{\partial s}(0,t) = Z(t)(L) = Y(L)(\gamma(t)) = \rho^1(\overline{\sigma}^C)(L)(\gamma(t)) = d_{\overline{\sigma}^C} L(\gamma(t)),$$

from which it follows that

$$\frac{d}{ds} J(\chi_s)\Big|_{s=0} = \int_\gamma d_{\overline{\sigma}^C} L.$$

If $\overline{\sigma}$ is a section satisfying the conditions given above, then so is $f\overline{\sigma}$ for every function $f$ on $M$. Taking into account that $(f\overline{\sigma})^C = f\overline{\sigma}^C + \dot{f}\overline{\sigma}^V$, we have that

$$\begin{aligned} 0 &= \int_\gamma d_{(f\overline{\sigma})^C} L \\ &= \int_\gamma f d_{\overline{\sigma}^C} L + \dot{f} d_{\overline{\sigma}^V} L \\ &= f\langle dL, \overline{\sigma}^V\rangle\Big|_{\gamma(t_0)}^{\gamma(t_1)} + \int_\gamma f\{d_{\overline{\sigma}^C} L - d_\Gamma\langle dL \circ S, \overline{\sigma}^C\rangle\} \\ &= \int_\gamma f i_{\overline{\sigma}^C}\{dL - d_\Gamma(dL \circ S)\}, \end{aligned}$$

whereby $\Gamma$ is the pseudo-SODE of which the extremals we are looking for will be solutions, and we have made use of the property that $[\sigma^C, \Gamma]$ is vertical, as one can easily verify in coordinates.

Since $f$ is arbitrary, the fundamental lemma of the calculus of variations implies that its coefficient must vanish along extremals $\gamma(t)$ and therefore also in an open neighbourhood of $\gamma$ in $E$. So the vanishing of the variation of $J$ is equivalent to

$$i_{\overline{\sigma}^C}(dL - d_\Gamma(dL \circ S)) = 0.$$

One easily verifies that $d_\Gamma(dL \circ S) - dL$ is 'semi-basic', and since it vanishes on the complete lift of arbitrary sections $\overline{\sigma}$ of $\overline{E}$, it follows that $d_\Gamma(dL \circ S) - dL = \lambda \mathcal{X}^0$. The value of $\lambda$ can be found by contraction with $\Gamma$:

$$\lambda = \langle d_\Gamma(dL \circ S) - dL, \Gamma \rangle = i_\Gamma d_\Gamma(dL \circ S) - d_\Gamma L = d_\Gamma i_\Gamma(dL \circ S) - d_\Gamma L = -d_\Gamma L.$$

Thus the Euler-Lagrange equations can be written as $d_\Gamma \Theta_L = dL$, from which it follows, since $d i_\Gamma \Theta_L = dL$, that $i_\Gamma \Omega_L = 0$.

# Chapter 8

# *Generalised connections and affine bundles*

## 8.1 Connections for Lagrange-type systems on affine Lie algebroids

In the previous chapters, we were able to establish a geometrical model for the Lagrange-type equations (7.21) on an affine Lie algebroid. More generally, the class of dynamical systems of interest were of the kind (7.19) and were given the name 'pseudo-SODE'.

In order to start a geometrical study of a number of features of such systems, we will have to bring 'connections' back into the picture. We have seen that, in the geometrical study of SODEs, connections appear both in linear and non-linear form. First, every SODE gives rise to an associated non-linear connection (Proposition 2.4). Secondly, every non-linear connection has associated 'linearised versions' which constitute the class of Berwald-type connections. Berwald-type connections obtained in these two steps from SODEs are applied to investigate certain qualitative aspects of the theory (see e.g. the applications in Chapter 3). In the next chapters we will show that, if we bring a generalised type of non-linear connection into the framework of an affine Lie algebroid prolongation of an affine bundle, the model to study pseudo-SODEs highly resembles the classical one for time-dependent SODEs.

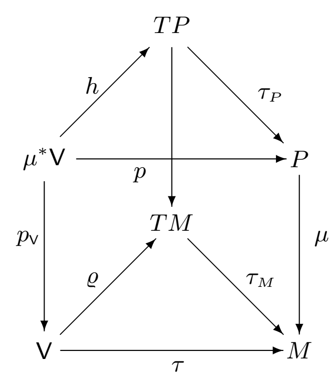
## 8.2 Connections over a vector bundle map and the horizontal subbundle of a prolonged bundle

Let us come back to the most general prolongation idea, as defined in Section 7.2 and visualised in the diagram on page 109. In that situation, $\mu : P \to M$ is a fibre bundle, $\tau : \mathsf{V} \to M$ a vector bundle and $\varrho : \mathsf{V} \to TM$ a linear bundle map.

We have already mentioned that there exists a canonical way to define 'vertical elements' for the prolongation $T^\varrho P$. If $Z \in \mathcal{V}^\varrho P$, then also $\varrho^1(Z) \in VP$ ($VP$ being the vertical submanifold of $TP$). The idea of arriving at a notion of 'horizontality' on $TP$, adapted to the presence of the anchor map in the picture, lies at the basis of the following concept, introduced in [11].

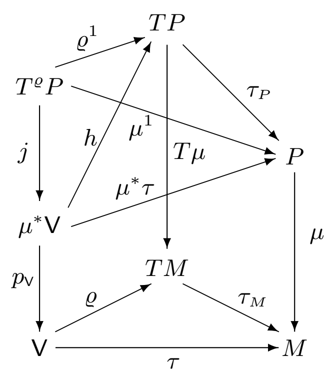
**Definition 8.1.** *A $\varrho$-connection on $\mu$ is a linear bundle map $h : \mu^*\mathsf{V} \to TP$ (over the identity on $P$), such that $\varrho \circ p_{\mathsf{V}} = T\mu \circ h$, where $p_{\mathsf{V}}$ is the projection of $\mu^*\mathsf{V}$ onto $\mathsf{V}$.*

There is a quite striking similarity between the diagram on page 109 and the one we can draw here for the illustration of all spaces involved in the definition of a $\varrho$-connection:




We have already mentioned that a vector in the image $\varrho^1(T^\varrho P)$ can be vertical in $TP$ even though the corresponding vector in the domain is not vertical in $T^\varrho P$. This is related to the observation that $\operatorname{Im} h$ can have a non-empty intersection with the vertical vectors on $P$. As discussed in detail in [11], $\operatorname{Im} h$ will in general also fail to determine a full complement to the vertical vectors on $P$. That is why one refers to a $\varrho$-connection on $\mu$ also as a 'generalised connection'. For completeness we should mention that the idea of a 'generalised connection' in [11] was inspired by earlier work in e.g. [28, 29, 88].

The point we would like to emphasise, however, is that it is perhaps not such a good idea to concentrate on horizontality on $TP$. Instead, as one may conjecture from an inspection of the prolongation-diagram and the diagram above, the better fibration to look for horizontality in this framework is the prolonged bundle $\mu^1 : T^\varrho P \to P$. In other words, we think it is important to bring the pictures of *ϱ-prolongation* and *ϱ-connection* together into the following scheme.




What we propose to discuss in detail now is that, given a $\varrho$-connection on $\mu$, there is an associated, genuine decomposition of the bundle $\mu^1$, i.e. a 'horizontal subspace', at each point $p \in P$, of the fibre of $T^\varrho P$, which is complementary to the vertical subspace at $p$. In other words, instead of considering a horizontal lift operation from sections of $\tau$ to sections of $\tau_P$, as is done in [11], it is more appropriate to focus on a horizontal lift from sections of $\tau$, and by extension sections of the pullback bundle $\mu^*\tau : \mu^*\mathsf{V} \to P$, to sections of the bundle $\mu^1$ (see also [51]).

Recall that the fibre linear map $j : T^\varrho P \to \mu^*\mathsf{V}$ (defined by expression (7.5)) is surjective and its kernel is $\mathcal{V}^\varrho P$. Therefore, we have the following short exact sequence of vector bundles:

$$0 \to \mathcal{V}^\varrho P \to T^\varrho P \xrightarrow{j} \mu^*\mathsf{V} \to 0, \tag{8.1}$$

where the second arrow is the natural injection. Recall that we have denoted the projection $(\mathsf{v}, X_p) \in T^\varrho P \mapsto \mathsf{v} \in \mathsf{V}$ by $\mu^2$.

**Proposition 8.2.** *The existence of a ϱ-connection on μ is equivalent to the existence of a splitting* $^H$ *of the short exact sequence (8.1); we have* $\varrho^1 \circ {}^H = h$.

Proof: Let $h : \mu^*\mathsf{V} \to TP$ be given and satisfy the requirements of a $\varrho$-connection on $\mu$. To define the 'horizontal lift' of a point $(p, \mathsf{v}) \in \mu^*\mathsf{V}$, as a point in $T^{\varrho}P$, it suffices to fix the projections of $(p,\mathsf{v})^H$ under $\varrho^1$ and $\mu^2$ in a consistent way. We put:

$$\varrho^1\big((p,\mathsf{v})^H\big) := h(p,\mathsf{v}) \qquad \text{and} \qquad \mu^2\big((p,\mathsf{v})^H\big) := \mathsf{v}. \tag{8.2}$$

This determines effectively an element of $T^{\varrho}P$ since $\varrho\circ\mu^2((p,\mathsf{v})^H) = \varrho(\mathsf{v}) = \varrho\circ p_{\mathsf{V}}((p,\mathsf{v})) = T\mu\circ h((p,\mathsf{v})) = T\mu\circ\varrho^1((p,\mathsf{v})^H)$. The horizontal lift is obviously a splitting of (8.1), since by construction $j\left((p,\mathsf{v})^H\right) = (\tau_P(h(p,\mathsf{v})),\mathsf{v}) = (p,\mathsf{v})$.

Conversely, if a splitting ${}^H$ of (8.1) is given, we define $h:\mu^*\mathsf{V}\to TP$ by $h(p,\mathsf{v}) = \varrho^1((p,\mathsf{v})^H)$. It satisfies the required properties, i.e. $h$ is a linear bundle map and we have

$$T\mu\circ h = T\mu\circ\varrho^1\circ{}^H = \varrho\circ\mu^2\circ{}^H = \varrho\circ p_{\mathsf{V}}\circ j\circ{}^H = \varrho\circ p_{\mathsf{V}},$$

which concludes the proof. □

Denoting the subbundle of $T^{\varrho}P$, determined by a splitting ${}^H$, which is complementary to $\mathcal{V}^{\varrho}P$ by $\mathcal{H}^{\varrho}P$, it follows that

$$T^{\varrho}P = \mathcal{H}^{\varrho}P\oplus\mathcal{V}^{\varrho}P. \tag{8.3}$$

An equivalent way of expressing this decomposition (analogous to what is familiar for the case of a classical Ehresmann connection) is the following: there exist two complementary projection operators $P_H$ and $P_V$ on $T^{\varrho}P$, i.e. we have $P_H + P_V = id$, and

$${P_H}^2 = P_H, \qquad {P_V}^2 = P_V, \qquad P_H\circ P_V = P_V\circ P_H = 0.$$

As usual, (8.1) leads to an associated short exact sequence for the set of sections of these spaces, regarded as bundles over $P$:

$$0\to \mathrm{Ver}(\mu^1)\to \mathrm{Sec}(\mu^1)\xrightarrow{j}\mathrm{Sec}(\mu^*\tau)\to 0, \tag{8.4}$$

where $\mathrm{Ver}(\mu^1)$ denotes the set of vertical sections of $\mu^1$. The same symbol $j$ is used for this second interpretation, so that for $\mathcal{Z}\in\mathrm{Sec}(\mu^1)$ and $p\in P$: $j(\mathcal{Z})(p) = j(\mathcal{Z}(p))$. Via the composition with $\mu$, sections of $\tau$ can be regarded as maps from $P$ to $\mathsf{V}$ and, as such, are (*basic*) sections of $\mu^*\tau: \mu^*\mathsf{V}\to P$. We will use the notations $P_V$ and $P_H$ also when we regard these projectors as acting on sections of $\mu^1$, rather than on points of $T^{\varrho}P$.

Let us take coordinates $(x^I,u^A)$ for a point $p$ in $P$. Following [11], we know that the map $h:\mu^*\mathsf{V}\to TP$ locally is of the form:

$$h(x^I,u^A,\mathsf{v}^a) = (x^I,u^A,\varrho^I_a(x)\mathsf{v}^a,-\Gamma^A_a(x,u)\mathsf{v}^a), \tag{8.5}$$

whereby we have adopted a different sign convention concerning the *connection coefficients* $\Gamma_a^A$. As shown in Proposition 8.2, a $\varrho$-connection on $\mu$ is equivalent to a decomposition of the bundle $T^\varrho P$, originating from a horizontal lift operation from $\mu^*\mathsf{V}$ to $T^\varrho P$ (or sections thereof). In the representation where points of $T^\varrho P$ are couples of an element of $\mathsf{V}$ and a suitable tangent vector of $P$, the horizontal lift is given by

$$(x^I, u^A, \mathsf{v}^a)^H = \left((x^I, \mathsf{v}^a), \mathsf{v}^a\left(\rho_a^I \frac{\partial}{\partial x^I} - \Gamma_a^A \frac{\partial}{\partial u^A}\right)\right).$$

Next to the basis (7.8) for $\mathrm{Sec}(\mu^1)$, it is convenient to introduce a local basis for the horizontal sections of $\mu^1$, which is given by

$$\mathcal{H}_a = P_H(\mathcal{X}_a) = \mathcal{X}_a - \Gamma_a^A(x, y)\mathcal{V}_A.$$

A representation of a section $\mathcal{Z} = \mathsf{Y}^a(x, u)\mathcal{X}_a + X^A\mathcal{V}_A$, adapted to the given connection, is then

$$\mathcal{Z} = \mathsf{Y}^a\mathcal{H}_a + (X^A + \mathsf{Y}^a\Gamma_a^A)\mathcal{V}_A.$$

Its projection onto $\mathrm{Sec}(\mu^*\tau)$, by means of $j$, is $\mathsf{Y} = \mathsf{Y}^a\mathsf{e}_a$.

### The case of a Lie algebroid

So far, we have only talked about anchored bundles. In order to introduce curvature for generalised connections, we have to require, next to the anchor map, also the presence of a Lie bracket structure. In Proposition 7.4 we have shown that, if $\tau : \mathsf{V} \to M$ is a Lie algebroid with anchor $\varrho$, there will be an associated Lie algebroid on $\mu^1$ (this time with anchor map $\varrho^1$). Once we have a Lie algebroid structure on a vector bundle, it becomes possible to define a bracket operation also on 'vector-valued forms' on sections of that bundle, in exactly the same way as it is done in the standard Frölicher and Nijenhuis theory (see e.g. [69]). Coming back now to our present situation, we have already come across a type (1,1) tensor field on the Lie algebroid $T^\varrho P$, namely the horizontal projector $P_H$. Using the Lie algebroid bracket on $\mu^1$, we thus can define curvature of the given (non-linear) $\varrho$-connection on $\mu$ in the way this is usually done for connections on a tangent bundle.

**Definition 8.3.** *The curvature $R$ of a $\varrho$-connection on $\mu$ is a skew-symmetric, $C^\infty(P)$-bilinear map:* $\mathrm{Sec}(\mu^1) \times \mathrm{Sec}(\mu^1) \to \mathrm{Sec}(\mu^1)$, *determined by* $R := \frac{1}{2}[P_H, P_H]$.

The coordinate expression for the curvature here is:

$$R = \frac{1}{2}\Big(\varrho^1(\mathcal{H}_a)(\Gamma^A_a) - \varrho^1(\mathcal{H}_a)(\Gamma^A_b) + \mathsf{C}\,^c_{ab}\Gamma^A_c\Big)\mathcal{X}^a \wedge \mathcal{X}^b \otimes \mathcal{V}_A.$$

Apart from applications to Lie algebroids [29, 72], it has recently been shown that $\varrho$-connections can be an important tool in, for example, nonholonomic mechanics [42], sub-Riemannian geometry [43], Poisson geometry [28] and in control theory [44]. Due to our new view on generalised connections (as established by Proposition 8.2), we can show next that generalised connections on an affine bundle $\pi$ are exactly what we need to study qualitative aspects of the theory of pseudo-SODEs.

## 8.3 Pseudo-SODE-connections

The appropriate arena to study pseudo-SODEs can be found in Section 7.6 (see the diagram on page 117). It is not immediately clear whether a pseudo-SODE comes with a canonically associated (non-linear) $\tilde{\rho}$-connection in this general setting. However, the construction of a connection becomes quite obvious when we have the additional structure of a Lie algebroid. So, let $\mu$ be an affine Lie algebroid $\pi : E \to M$ and $\tau$ its bi-dual $\tilde{\pi}$. In this case, also the vector bundle $\pi^1$ is a Lie algebroid.

Suppose that we have a $\tilde{\rho}$-connection on $\pi$. For later use, we list the following brackets of horizontal and vertical sections:

$$\begin{aligned}
[\mathcal{H}_\alpha, \mathcal{H}_\beta] &= C^\delta_{\alpha\beta}\mathcal{H}_\delta + (C^\delta_{\alpha\beta}\Gamma^\gamma_\delta + \tilde{\rho}^1(\mathcal{H}_\beta)(\Gamma^\gamma_\alpha) - \tilde{\rho}^1(\mathcal{H}_\alpha)(\Gamma^\gamma_\beta))\mathcal{V}_\gamma,\\
[\mathcal{H}_0, \mathcal{H}_\beta] &= C^\delta_{0\beta}\mathcal{H}_\delta + (C^\delta_{0\beta}\Gamma^\gamma_\delta + \tilde{\rho}^1(\mathcal{H}_\beta)(\Gamma^\gamma_0) - \tilde{\rho}^1(\mathcal{H}_0)(\Gamma^\gamma_\beta))\mathcal{V}_\gamma, \quad (8.6)\\
[\mathcal{H}_\alpha, \mathcal{V}_\beta] &= \frac{\partial \Gamma^\delta_\alpha}{\partial y^\beta}\mathcal{V}_\delta \qquad \text{and} \qquad [\mathcal{H}_0, \mathcal{V}_\beta] = \frac{\partial \Gamma^\delta_0}{\partial y^\beta}\mathcal{V}_\delta.
\end{aligned}$$

In the previous section, we have seen how curvature can be defined. In the current set-up, next to $P_H$, we also have a second type (1,1) tensor field on $T^{\tilde{\rho}}E$ at our disposal, namely the vertical endomorphism $S$ (see Definition 7.7).

**Definition 8.4.** *The torsion $T$ of a $\tilde{\rho}$-connection on $\pi$ is the skew-symmetric, $C^\infty(E)$-bilinear map:* $\mathrm{Sec}(\tilde{\pi}^1) \times \mathrm{Sec}(\tilde{\pi}^1) \to \mathrm{Sec}(\tilde{\pi}^1)$*, determined by $T = [P_H, S]$.*

The coordinate expression for the torsion is:

$$\begin{aligned} T &= \frac{1}{2}\left(\frac{\partial \Gamma^\gamma_\alpha}{\partial y^\beta} - \frac{\partial \Gamma^\gamma_\beta}{\partial y^\alpha} - C^\gamma_{\alpha\beta}\right)\mathcal{X}^\alpha \wedge \mathcal{X}^\beta \otimes \mathcal{V}_\gamma \\ &\quad + \left(\frac{\partial \Gamma^\gamma_0}{\partial y^\alpha} - \Gamma^\gamma_\alpha + y^\beta \frac{\partial \Gamma^\gamma_\alpha}{\partial y^\beta} - C^\gamma_{0\alpha}\right)\mathcal{X}^0 \wedge \mathcal{X}^\alpha \otimes \mathcal{V}_\gamma. \end{aligned}$$

In [11] the notions of torsion and curvature of a generalised connection were only introduced under rather special circumstances, namely that the connection is linear (which requires, in the current set-up, that $\pi$ is in fact a vector bundle, see the next section). That there exist such notions as torsion and curvature for the non-linear case also, becomes clear only if the interest is shifted, as we do, from horizontality on $TE \to E$ to horizontality on $\tilde{\pi}^1 : T^{\tilde{\rho}}E \to E$.

It is time to bring the dynamical systems back into the picture. Already in Proposition 2.4 we have seen that a time-dependent SODE $\Gamma$ gives rise to a (genuine) non-linear connection on $\pi_M : J^1M \to M$. Coming back to the general situation, the Lie algebroid structure on $\pi^1$ provides us with an exterior derivative; we use the standard notation $d_\Gamma$ for the commutator $[i_\Gamma, d]$, which plays the role of Lie derivative and extends, as a degree zero derivation, to tensor fields of any type.

**Proposition 8.5.** *If $\Gamma$ is a pseudo-SODE on an affine Lie algebroid $\pi$, then the operator*

$$P_H = \frac{1}{2}\Big(I - d_\Gamma S + \mathcal{X}^0 \otimes \Gamma\Big) \tag{8.7}$$

*defines a horizontal projector on $\mathrm{Sec}(\pi^1)$ and hence a $\tilde{\rho}$-connection on $\pi$.*

PROOF: The proof follows the lines of the classical one for time-dependent mechanics (see [26] or [20]). We will give a brief sketch of one possibility to proceed. For $\tilde{\sigma} \in \mathrm{Sec}(\tilde{\pi})$, define the horizontal lift $\tilde{\sigma}^H \in \mathrm{Sec}(\pi^1)$ by

$$\tilde{\sigma}^H = \frac{1}{2}\Big(\tilde{\sigma}^C + \langle \tilde{\sigma}, e^0\rangle \Gamma - [\Gamma, \tilde{\sigma}^V]\Big), \tag{8.8}$$

where $\tilde{\sigma}^C$ is the complete lift, as defined in Section 7.7. It is easy to see that this behaves tensorially for multiplication by basic functions and that $\tilde{\sigma}^H$ projects onto $\sigma$. Hence, extending the horizontal lift to $\mathrm{Sec}(\pi^*\tilde{\pi})$ by imposing linearity for multiplication by functions on $E$, we obtain a splitting of the short exact sequence (8.1) for the present situation. This in fact concludes the proof of the existence of a $\tilde{\rho}$-connection, but it is interesting to verify further the explicit formula for $P_H$. One can, for example, compute the Lie algebroid brackets $[\Gamma, \overline{\sigma}^V]$ and $[\Gamma, \overline{\sigma}^H]$ for $\overline{\sigma} \in \mathrm{Sec}(\overline{\pi})$, from which it then easily follows (using also the properties $S(\overline{\sigma}^H) = \overline{\sigma}^V$ and $S(\overline{\sigma}^V) = 0$), that $d_\Gamma S(\overline{\sigma}^V) = \overline{\sigma}^V$, $d_\Gamma S(\overline{\sigma}^H) = -\overline{\sigma}^H$ and $d_\Gamma S(\Gamma) = 0$. The verification that $P_H$ is a projection operator and that $P_H(\tilde{\sigma}^H) = \tilde{\sigma}^H$ then is immediate. □

The connection coefficients for the pseudo-Sode (7.18) are

$$\begin{aligned} \Gamma_0^\alpha &= -f^\alpha + \tfrac{1}{2} y^\beta \Big( \tfrac{\partial f^\alpha}{\partial y^\beta} + C_{0\beta}^\alpha \Big) = -f^\alpha - y^\beta \Gamma_\beta^\alpha \\ \Gamma_\beta^\alpha &= -\tfrac{1}{2} \Big( \tfrac{\partial f^\alpha}{\partial y^\beta} + y^\gamma C_{\gamma\beta}^\alpha + C_{0\beta}^\alpha \Big). \end{aligned} \tag{8.9}$$

**Proposition 8.6.** *A $\tilde{\rho}$-connection on $\pi$ is associated with a pseudo-*Sode *(by means of (8.7)) if and only if its torsion vanishes.*

Proof: We give a short coordinate proof. Substituting connection coefficients of the form (8.9), one finds that indeed $T = 0$ holds. On the other hand, if $T = 0$, then $\frac{\partial \Gamma_\alpha^\gamma}{\partial y^\beta} + \frac{1}{2} C_{\beta\alpha}^\gamma = \frac{\partial \Gamma_\beta^\gamma}{\partial y^\alpha} + \frac{1}{2} C_{\alpha\beta}^\gamma$ and thus $\frac{\partial}{\partial y^\beta} \Big( \Gamma_\alpha^\gamma + \frac{1}{2} y^\mu C_{\mu\alpha}^\gamma \Big) = \frac{\partial}{\partial y^\alpha} \Big( \Gamma_\beta^\gamma + \frac{1}{2} y^\mu C_{\mu\beta}^\gamma \Big)$. This means that there exist functions $g^\gamma \in C^\infty(E)$, such that

$$\Gamma_\alpha^\gamma + \frac{1}{2} y^\mu C_{\mu\alpha}^\gamma = \frac{\partial g^\gamma}{\partial y^\alpha}. \tag{8.10}$$

We will use this observation in the other part of the (zero) torsion: the coefficient of $\mathcal{X}^0 \wedge \mathcal{X}^\alpha \otimes \mathcal{V}_\gamma$ can be rewritten as $\frac{\partial}{\partial y^\alpha} \Big( \Gamma_0^\gamma - y^\mu C_{0\mu}^\gamma - 2g^\gamma + y^\beta \frac{\partial g^\gamma}{\partial y^\beta} \Big) = 0$, thus

$$\Gamma_0^\gamma - y^\mu C_{0\mu}^\gamma - 2g^\gamma + y^\beta \frac{\partial g^\gamma}{\partial y^\beta} = h^\gamma \tag{8.11}$$

for some functions $h^\gamma$ on $M$. Let us introduce now the functions $f^\gamma = -2g^\gamma - h^\gamma - C_{0\mu}^\gamma y^\mu$. Expressions (8.10,8.11) for $g^\gamma$ and $h^\gamma$ can be rewritten in terms of $f^\alpha$, leading indeed to connection coefficients of the form (8.9). □

## 8.4 Linear $\varrho$-connections on a vector bundle

After these side-observations on pseudo-Sodes, we can come back to the general picture of $\varrho$-connections (as in the diagram on page 129). This time we will assume that, next to $\tau$, also $\mu : P \to M$ is a *vector* bundle. Linearity of a connection is characterised in [11] by an invariance property of the map $h$ under the flow of the dilation field on $P$. A more direct characterisation of linearity is the following. Let $\Sigma_\lambda : P \times_M P \to P$ (for $\lambda \in \mathbb{R}$) denote the linear combination map: $\Sigma_\lambda(p_1, p_2) = p_1 + \lambda p_2$. A $\varrho$-connection on $\mu$ is said to be *linear* if the map $h : \mu^*\mathsf{V} \to TP$ has the property

$$h(p_1 + \lambda p_2, \mathsf{v}) = T_{(p_1,p_2)}\Sigma_\lambda\left(h(p_1, \mathsf{v}), h(p_2, \mathsf{v})\right), \tag{8.12}$$

for all $(p_1, p_2) \in P \times_M P$, $\lambda \in \mathbb{R}$ and $\mathsf{v} \in \mathsf{V}$.

As is shown in [11], any operator $\nabla : \mathrm{Sec}(\tau) \times \mathrm{Sec}(\mu) \to \mathrm{Sec}(\mu)$ which is $\mathbb{R}$-bilinear and has the properties

$$\nabla_{f\mathsf{s}}\sigma = f\nabla_{\mathsf{s}}\sigma, \qquad \nabla_{\mathsf{s}}(f\sigma) = f\nabla_{\mathsf{s}}\sigma + \varrho(\mathsf{s})(f)\sigma, \tag{8.13}$$

for all $\mathsf{s} \in \mathrm{Sec}(\tau)$, $\sigma \in \mathrm{Sec}(\mu)$ and $f \in C^\infty(M)$, defines a unique *linear* $\varrho$-connection on $\mu$. Linear $\varrho$-connections on $\mu$ were also called *pseudo-connections on* $\mu$ in [86]. As usual, the linearity of the covariant derivative operator $\nabla$ in its first argument, implies that the value of $\nabla_{\mathsf{s}}\sigma$ at a point $m \in M$, only depends on the value of $\mathsf{s}$ at $m$ and thus gives rise to an operator $\nabla_{\mathsf{v}} : \mathrm{Sec}(\mu) \to P_{\tau(\mathsf{v})}$, for each $\mathsf{v} \in \mathsf{V}$, determined by

$$\nabla_{\mathsf{v}}\eta := \nabla_{\mathsf{s}}\eta(m), \qquad \text{with} \quad \mathsf{s}(m) = \mathsf{v}.$$

In order to come to a covariant derivative along curves and a rule of parallel transport, we make the following preliminary observation. Going back to the overall diagram on p. 129, we see two ways to go from $T^\varrho P$ to $TP$, namely the direct map $\varrho^1$ and $h \circ j$. By definition, the image of a point of $T^\varrho P$ under both maps projects under $T\mu$ onto the same $\varrho(\mathsf{v})$, so that the difference is a vertical vector at some point $p \in P$ which, when $P$ is a vector bundle, can be identified with an element of $P_{\mu(p)}$. With these identifications understood, we eventually get a map from $T^\varrho P$ to $P$ which is called the connection map in [11] (by analogy with the connection map in [89]). Let us summarise this by writing simply

$$K := \varrho^1 - h \circ j : T^\varrho P \to P \tag{8.14}$$

(read: $K$ is $\varrho^1 - h \circ j$, when regarded as map from $T^\varrho P$ into $P$). The following side observation is worth being made here. In the alternative concept of $\varrho$-connections, as established by Proposition 8.2, it is clear that the connection map $K$ is nothing but the vertical projector $P_V$, with a similar identification being understood (to be precise: the isomorphism between $\mathcal{V}^\varrho_p P$ and $V_p P$, followed by the identification with $P_{\mu(p)}$ again). In fact this illustrates that the alternative view is superior to the one expressed by Definition 8.1, in the following sense. Once the importance of the space $T^\varrho P$ is recognised, one can (in the present case where $P$ is a vector bundle) define a vertical lift operation from $P_{\mu(p)}$ to $\mathcal{V}^\varrho_p P$ in the usual way; it extends to sections of bundles over $P$, i.e. yields a vertical lift from sections of $\mu^* P \to P$ to $\mathrm{Sec}(\mu^1)$. So, it is a matter of developing first these tangent bundle like features of the $\varrho$-prolongation, after which all tools are available to discuss $\varrho$-connections without ever needing the map $h$. For the sake of further unifying both pictures, however, we will continue here to take advantage of the insight which is being offered by our overall diagram.

Let now $\mathsf{c} : I \to \mathsf{V}$ be a *$\varrho$-admissible curve*, which means that $\dot{c}_M = \varrho \circ \mathsf{c}$, where $c_M = \tau \circ \mathsf{c}$ is the projected curve in $M$. Consider further a curve $\psi : I \to P$ in $P$ which projects onto $c_M$, i.e. such that $\psi_M := \mu \circ \psi = c_M$. It follows that $T\mu \circ \dot{\psi} = \varrho \circ \mathsf{c}$, so that such a $\psi$ actually gives rise to a curve in $T^\varrho P$: $t \mapsto (\mathsf{c}(t), \dot{\psi}(t))$. As a result, making use of the map $K$, we can obtain a new curve in $P$, which is denoted by $\nabla_{\mathsf{c}} \psi$:

$$\nabla_{\mathsf{c}} \psi(t) := K((\mathsf{c}(t), \dot{\psi}(t))) = \dot{\psi}(t) - h((\psi(t), \mathsf{c}(t))), \tag{8.15}$$

(the identification of $P$ with $VP$ being understood). If $\eta$ is a section of $\mu$ and $\mathsf{c}$ is an admissible curve, then denoting by $\psi$ the restriction of $\eta$ to that curve, $\psi(t) = \eta(c_M(t))$, one can show that

$$\nabla_{\mathsf{c}} \psi(t) = \nabla_{\mathsf{c}(t)} \eta. \tag{8.16}$$

As can be readily seen from (8.15), given an admissible curve $\mathsf{c}$ and a point $p \in P$, finding a curve $\psi$ in $P$ which starts at $p$ and makes $\nabla_{\mathsf{c}} \psi = 0$ is a well-posed initial value problem for a first-order ordinary differential equation, and hence gives rise to a unique solution. The solution is called the *horizontal lift* of $\mathsf{c}$ through $p$, denoted by $c^h_p$. Hence, we have

$$\nabla_{\mathsf{c}} c^h_p = 0, \tag{8.17}$$

and points in the image of $c^h_p$ are said to be obtained from $p$ by *parallel transport along* $\mathsf{c}$.

It is of some interest to rephrase what we have said at the beginning of the discussion on $\varrho$-admissible curves: if $\mathsf{c} : I \to \mathsf{V}$ is $\varrho$-admissible, then for every $\psi : I \to P$ which projects onto $c_M$, the curve $t \mapsto (\mathsf{c}(t), \dot{\psi}(t))$ in fact is a $\varrho^1$-admissible curve in $T^\varrho P$. This idea can be pushed a bit further. Indeed, when thinking of curves in the context of our alternative view on $\varrho$-connections, it is rather the following construction which looks like the natural thing to do.

Consider a curve $\gamma$ in $\mu^*\mathsf{V}$, i.e. $\gamma$ is of the form $\gamma : t \mapsto (\psi(t), \mathsf{c}(t))$, with $\mathsf{c} : I \to \mathsf{V}$ and $\psi : I \to P$, whereby the only assumption at the start is that $\psi_M = c_M$. Take its horizontal lift $\gamma^H : I \to T^\varrho P$ which is defined, according to (8.2), by

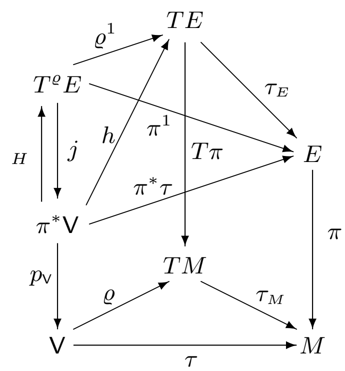
$$t \mapsto \gamma^H(t) = \big(\mathsf{c}(t), h(\psi(t), \mathsf{c}(t))\big). \tag{8.18}$$

Then, we could define $\psi$ to be $c^h$, the horizontal lift of $\mathsf{c}$, if $\gamma^H$ is a $\varrho^1$-admissible curve in $T^\varrho P$. Indeed, it is clear by construction that $\mu^1 \circ \gamma^H = \psi$, so that $\varrho^1$-admissibility requires that $\dot\psi = \varrho^1 \circ \gamma^H = h(\psi, \mathsf{c})$. Since $\psi_M = c_M$, this implies in particular that $\dot c_M = T\mu \circ \dot\psi = T\mu(h(\psi, \mathsf{c})) = \varrho \circ \mathsf{c}$. So, this alternative definition implies that $\mathsf{c}$ will necessarily have to be $\varrho$-admissible. Furthermore, from comparing what $\varrho^1$-admissibility means with (8.15) and (8.17), it is clear that we are talking then about the same concept of horizontal lift $c^h$.

Note, by the way, that this other way of defining $c^h$ by no means relies on the assumption of linearity of the $\varrho$-connection. So, it is perfectly possible to talk about parallel transport also in the context of non-linear connections. The difference then is, of course, that if we look at points of $P$ in the image of curves $c^h$ with different initial values in $P_m$, and look at this as a map between fibres of $P$, there need not be any special feature to talk about (compared to the fibre-wise linear action of this map we have in the case of a linear connection); also, if $\mathsf{c}$ has a given interval as domain, $c^h$ need not be defined over the same domain. Needless to say, one can introduce such a generalisation also within the more traditional approach described first. Indeed, the map $K$ makes sense for arbitrary $\varrho$-connections and as a result one can introduce an operation $\nabla_{\mathsf{s}}\sigma$ also in this more general situation. This then still depends on the section $\mathsf{s}$ of $\mathsf{V}$ in a $C^\infty(M)$-linear way, but the fact that such a $\nabla$ is not commonly used comes from the failure of having a derivation property with respect to the module structure of $\mathrm{Sec}(\mu)$.

## 8.5 Affine $\varrho$-connections on an affine bundle

Suppose that $\pi : E \to M$ is an affine bundle, modelled on a vector bundle $\overline{\pi} : \overline{E} \to M$. With reference to the second section of this chapter, the situation we will focus on now is the case where $\mu : P \to M$ is the affine bundle $\pi : E \to M$, whereas $\tau : \mathsf{V} \to M$ still is an arbitrary vector bundle.




Note that $\varrho$ can be regarded in an obvious way also as a map (denoted by the same symbol) from $\pi^*\mathsf{V}$ into $\pi^*TM$, by means of $\varrho(e, \mathsf{v}) = (e, \varrho(\mathsf{v}))$ for any $(e, \mathsf{v}) \in \pi^*\mathsf{V}$.

Suppose a $\varrho$-connection on $\pi$ is given. As always, we will use the same symbol for the extension of the maps $h$, $v$, ${}^H$ and ${}^V$ to sections of the corresponding bundles. As a consequence of the existence of a splitting, for any section $\mathcal{Z} \in \mathrm{Sec}(\pi^1)$, there exist uniquely determined sections $\mathsf{X}$ of $\pi^*\tau$ and $\overline{Y}$ of $\pi^*\overline{\pi}$ such that

$$\mathcal{Z} = \mathsf{X}^H + \overline{Y}^V. \tag{8.19}$$

In fact, if $\{\mathsf{s}_a\}$ is a local basis for $\mathrm{Sec}(\tau)$ and $\{\overline{\sigma}_\alpha\}$ a basis for $\mathrm{Sec}(\overline{\pi})$, and these are interpreted as sections of $\pi^*\mathsf{V} \to E$ and $\pi^*E \to E$, respectively, then $\{\mathsf{s}_a^H, \overline{\sigma}_\alpha^V\}$ provides a local basis for $\mathrm{Sec}(\pi^1)$.

Our main objective is to define and characterise $\varrho$-connections on $\pi$ which are *affine*. For that purpose, we will also need the overall diagram on page 129, with the vector bundles $\tilde{\pi} : \tilde{E} \to M$ and $\overline{\pi} : \overline{E} \to M$ in the role of $\mu : P \to M$.

**Definition 8.7.** *A $\varrho$-connection $h$ on the affine bundle $\pi : E \to M$ is said to be affine, if there exists a linear $\varrho$-connection $\tilde{h} : \tilde{\pi}^*\mathsf{V} \to T\tilde{E}$ on $\tilde{\pi} : \tilde{E} \to M$ such that,*

$$\tilde{h} \circ \iota = T\iota \circ h.$$

Both sides in the above commutative scheme of course are regarded as maps from $\pi^*\mathsf{V}$ to $T\tilde{E}$, which means that the $\iota$ on the left stands for the obvious extension $\iota : \pi^*\mathsf{V} \to \tilde{\pi}^*\mathsf{V}, (e, \mathsf{v}) \mapsto (\iota(e), \mathsf{v})$.

Probably the best way to see what this concept means is to look at a coordinate representation. In the current set-up, an arbitrary $h : \pi^*\mathsf{V} \to TE$ is of the form:

$$h(x^I, y^\alpha, \mathsf{v}^a) = (x^I, y^\alpha, \varrho^I_a(x)\mathsf{v}^a, -\Gamma^\alpha_a(x, y)\mathsf{v}^a), \tag{8.20}$$

Similarly, $\tilde{h} : \tilde{\pi}^*\mathsf{V} \to T\tilde{E}$, which is furthermore assumed to be linear, takes the form

$$\tilde{h}(x^I, y^A, \mathsf{v}^a) = (x^I, y^A, \varrho^I_a(x)\mathsf{v}^a, -\tilde{\Gamma}^A_{aB}(x)y^B\mathsf{v}^a). \tag{8.21}$$

Here, the indices $A, B$ start running from 0. We have

$$\tilde{h}(\iota(e), \mathsf{v}) = \Big(x^I, 1, y^\alpha, \varrho^I_a(x)\mathsf{v}^a, -(\tilde{\Gamma}^A_{a0}(x) + \tilde{\Gamma}^A_{a\beta}(x)y^\beta)\mathsf{v}^a)\Big),$$

whereas

$$T\iota \circ h(e, \mathsf{v}) = \Big(x^I, 1, y^\alpha, \varrho^I_a(x)\mathsf{v}^a, 0, -\Gamma^\alpha_a(x, y)\mathsf{v}^a\Big).$$

It follows that $\tilde{\Gamma}^0_{aB} = 0$ and, more importantly, that the connection coefficients of the affine $\varrho$-connection $h$ are of the form (omitting tildes)

$$\Gamma^\alpha_a(x, y) = \Gamma^\alpha_{a0}(x) + \Gamma^\alpha_{a\beta}(x)y^\beta. \tag{8.22}$$

Notice that $\overline{\pi} : \overline{E} \to M$ is a (proper) vector subbundle of $\tilde{\pi}$. With respect to the given anchor map, it of course also has its $\varrho$-prolongation $T^\varrho\overline{E}$. Taking the restriction of the linear $\varrho$-connection $\tilde{h}$ to $\overline{\pi}^*\mathsf{V}$, we get a linear $\varrho$-connection $\overline{h}$ on $\overline{\pi}$, meaning that $\tilde{h} \circ \overline{\iota} = T\overline{\iota} \circ \overline{h}$. The above coordinate expressions make this very obvious. Indeed, if $(x^I, w^\alpha)$ are the coordinates of an element $\overline{e} \in \overline{E}$, we have

$$\begin{aligned}
\overline{h}(x^I, \overline{y}^\alpha, \mathsf{v}^a) &= \tilde{h}(x^I, 0, w^\alpha, \mathsf{v}^a) \\
&= (x^I, 0, w^\alpha, \varrho^I_a\mathsf{v}^a, 0, -\Gamma^\alpha_{a\beta}w^\beta\mathsf{v}^a) \quad \text{as element of } T\tilde{E} \\
&= (x^I, w^\alpha, \varrho^I_a\mathsf{v}^a, -\Gamma^\alpha_{a\beta}w^\beta\mathsf{v}^a) \quad \text{as element of } T\overline{E}.
\end{aligned}$$

Note further that we can *formally* write for the coordinate expression of $h(e + \overline{e}, v)$:

$$\begin{aligned}
h(x^I, y^\alpha + w^\alpha, \mathsf{v}^a) &= \big(x^I, y^\alpha + w^\alpha, \varrho^I_a\mathsf{v}^a, -(\Gamma^\alpha_{a0} + \Gamma^\alpha_{a\beta}y^\beta)\mathsf{v}^a - \Gamma^\alpha_{a\beta}w^\beta\mathsf{v}^a\big) \\
&= h(x^I, y^\alpha, \mathsf{v}^a) + \overline{h}(x^I, w^\alpha, \mathsf{v}^a).
\end{aligned}$$

But this is more than just a formal way of writing: the following intrinsic construction which generalises (8.12) is backing it. Let $\Sigma$ denote the action of $\overline{E}$ on $E$ which defines the affine structure, i.e. $\Sigma(e,\overline{e}) = e + \overline{e}$ for $(e,\overline{e}) \in E \times_M \overline{E}$. Then the above formal relation expresses that we have:

$$h(e + \overline{e}, \mathsf{v}) = T_{(e,\overline{e})}\Sigma\left(h(e,\mathsf{v}), \overline{h}(\overline{e},\mathsf{v})\right) \tag{8.23}$$

In fact, by reading the above coordinate considerations backwards, roughly speaking, one can see that (8.23), for a given linear $\overline{h}$, will imply that the connection coefficients of the $\varrho$-connection $h$ have to be of the form (8.22). In other words, the following is an equivalent definition of affineness of $h$.

**Definition 8.8.** *A $\varrho$-connection $h$ on the affine bundle $\pi : E \to M$ is affine, if there exists a linear $\varrho$-connection $\overline{h} : \overline{\pi}^*\mathsf{V} \to T\overline{E}$ on $\overline{\pi} : \overline{E} \to M$, such that (8.23) holds for all $(e,\overline{e}) \in E \times_M \overline{E}$.*

One can then construct an extension $\tilde{h} : \tilde{\pi}^*\mathsf{V} \to T\tilde{E}$, which coincides with $\overline{h}$ when restricted to $\overline{\pi}^*\mathsf{V}$, by requiring that $\tilde{h}$ be linear and satisfy $\tilde{h}\circ\iota = T\iota\circ h$.

Let us repeat that, as a result of Proposition 8.2 and Definition 8.8, the existence of an affine $\varrho$-connection on $\pi$ is equivalent to the existence of a horizontal lift from $\mathrm{Sec}(\pi^*\tau)$ to $\mathrm{Sec}(\pi^1)$, giving rise to a direct sum decomposition (8.3), and which is such that, in coordinates, the connection coefficients are of the form (8.22).

We next turn our attention to the concept of connection map, and want to see for the particular case of an affine $\varrho$-connection, to what extent it gives rise also to a covariant derivative operator and a notion of parallel transport.

When considering the $\varrho$-prolongation of different bundles $P$, it is convenient to indicate the dependence on $P$ also in the map $\varrho^1$. Given a $\varrho$-connection $h$ on the affine bundle $\pi : E \to M$, the map $\varrho^1_E - h \circ j : T^\varrho E \to TE$ gives rise (as before) to a vertical tangent vector to $E$, at the point $e$ say. As such, this vector can be identified with an element of $\overline{E}$, the vector bundle on which $E$ is modelled, at the point $\pi(e)$. With the same notational simplification as before, we thus get a connection map

$$K := \varrho^1_E - h \circ j : T^\varrho E \to \overline{E} \tag{8.24}$$

(technically, $K = p_{\overline{E}} \circ {}_v \circ (\varrho^1 - h \circ j)$, where ${}_v : TE \to \pi^*\overline{E}$ is defined by (6.10) and $p_{\overline{E}}$ is the projection $\pi^*\overline{E} \to \overline{E}$). $K$ of course also extends to a map from $\mathrm{Sec}(\pi^1)$ to $\mathrm{Sec}(\overline{\pi})$. It follows directly from the definition that we have

$$K(\mathcal{H}_a) = 0, \qquad K(\mathcal{V}_\alpha) = \overline{e}_\alpha. \tag{8.25}$$

We wish to come back here in some more detail to the relation between the map $K$ and the vertical projector $P_V = id - P_H$, coming from the direct sum decomposition of $T^\varrho E$. In the present case of an affine bundle $\pi : E \to M$ over a vector bundle $\overline{\pi} : \overline{E} \to M$, we have shown in Section 6.1 that there is a natural vertical lift operation from $\overline{E}_m$ to $T_e E$ for each $e \in E_m$. It is determined by expression (6.7). This in turn extends to an operator $v : \pi^*\overline{E} \to T^\varrho E$, determined by $(e, \overline{e})^V = (0, v(e, \overline{e}))$, which defines an isomorphism between $\pi^*\overline{E}$ and $\mathrm{Im}\, P_V$. The short exact sequence (8.1) of which a $\varrho$-connection is a splitting, can thus be replaced by

$$0 \to \pi^*\overline{E} \xrightarrow{V} T^\varrho E \xrightarrow{j} \pi^*\mathsf{V} \to 0. \tag{8.26}$$

Within this picture of $\varrho$-connections, the connection map $K$ thus is essentially the co-splitting of the splitting ${}^H$, that is to say, we have $K \circ {}^V = id_{\pi^*\overline{E}}$ and ${}^V \circ K + {}^H \circ j = id_{T^\varrho E}$.

The map $K$ becomes more interesting when the connection is affine. Indeed, denoting the projection of $T^\varrho \tilde{E}$ onto $\tilde{\pi}^*\mathsf{V}$ by $\tilde{j}$, it then follows from Definition 8.8 that we also have a connection map

$$\tilde{K} := \varrho^1_{\tilde{E}} - \tilde{h} \circ \tilde{j} : T^\varrho \tilde{E} \to \tilde{E}. \tag{8.27}$$

The map $T\iota : TE \to T\tilde{E}$ extends to a map from $T^\varrho E$ to $T^\varrho \tilde{E}$ in the following obvious way: $T\iota : (\mathsf{v}, X_e) \mapsto (\mathsf{v}, T\iota(X_e))$. Indeed, we have $T\tilde{\pi}(T\iota(X_e)) = T(\tilde{\pi} \circ \iota)(X_e) = T\pi(X_e) = \varrho(\mathsf{v})$, as required.

**Proposition 8.9.** *For an affine $\varrho$-connection on $\pi$ we have*

$$\overline{\iota} \circ K = \tilde{K} \circ T\iota. \tag{8.28}$$

Proof: In coordinates, $K$ and $\tilde{K}$ are given by

$$\begin{aligned} K : &\quad (x^I, \mathsf{v}^a, y^\alpha, Z^\alpha) \mapsto (Z^\alpha + \Gamma^\alpha_a \mathsf{v}^a)\, \overline{e}_\alpha(x) \\ \tilde{K} : &\quad (x^I, \mathsf{v}^a, y^A, Z^A) \mapsto Z^0 e_0(x) + (Z^\alpha + \Gamma^\alpha_{aB} y^B \mathsf{v}^a)\, e_\alpha(x). \end{aligned}$$

Hence,

$$\begin{aligned} \tilde{K} \circ T\iota(x^I, \mathsf{v}^a, y^\alpha, Z^\alpha) &= \tilde{K}(x^I, \mathsf{v}^a, 1, y^\alpha, 0, Z^\alpha) \\ &= \Big(Z^\alpha + (\Gamma^\alpha_{a0} + \Gamma^\alpha_{a\beta} y^\beta)\mathsf{v}^a\Big)\, e_\alpha(x), \end{aligned}$$

from which the result follows in view of (8.22). □

Notice that $\overline{h}$ also has a corresponding connection map $\overline{K} : T^\varrho\overline{E} \to \overline{E}$, which obviously coincides with $\tilde{K}|_{T^\varrho\overline{E}}$, so that we also have

$$\bar{\iota} \circ \overline{K} = \tilde{K} \circ T\bar{\iota}. \tag{8.29}$$

Let now $\mathsf{s}$ be a section of $\tau$ and $\sigma$ a section of $\pi$. If we apply the tangent map $T\sigma : TM \to TE$ to $\varrho(\mathsf{s}(m))$, it is obvious by construction that $\big(\mathsf{s}(m), T\sigma(\varrho(\mathsf{s}(m)))\big)$ will be an element of $T^\varrho E$. The connection map $K$ maps this into a point of $\overline{E}|_m$. Hence, the covariant derivative operator of interest in this context is the map $\nabla : \mathrm{Sec}(\tau) \times \mathrm{Sec}(\pi) \to \mathrm{Sec}(\overline{\pi})$, defined by

$$\nabla_{\mathsf{s}}\sigma(m) = K\big(\mathsf{s}(m), T\sigma(\varrho(\mathsf{s}(m)))\big). \tag{8.30}$$

To discover the properties which uniquely characterise the covariant derivative associated to an affine $\varrho$-connection, we merely have to exploit the results of Proposition 8.9. In doing so, we will of course rely on the known properties (see [11]) of the covariant derivative $\tilde{\nabla}$, associated to the linear $\varrho$-connection $\tilde{h}$. We observe that $\nabla$ is manifestly $\mathbb{R}$-linear in its first argument and now further look at its behaviour with respect to the $C^\infty(M)$-module structure on $\mathrm{Sec}(\tau)$. From (8.28), it follows that for $f \in C^\infty(M)$,

$$\begin{aligned}
\bar{\iota}\big((\nabla_{f\mathsf{s}}\sigma)(m)\big) &= \bar{\iota}\Big(K\big(f\mathsf{s}(m), T\sigma(\varrho(f\mathsf{s}(m)))\big)\Big) \\
&= \tilde{K}\big(f\mathsf{s}(m), T(\iota\sigma)(\varrho(f\mathsf{s}(m)))\big) \\
&= \tilde{\nabla}_{f\mathsf{s}}(\iota\sigma)(m) = f(m)\,\tilde{\nabla}_{\mathsf{s}}(\iota\sigma)(m) \\
&= f(m)\tilde{K}\big(\mathsf{s}(m), T\iota \circ T\sigma(\varrho(\mathsf{s}(m)))\big) \\
&= f(m)\,\bar{\iota}\Big(K\big(\mathsf{s}(m), T\sigma(\varrho(\mathsf{s}(m)))\big)\Big) \\
&= \bar{\iota}\big(f(m)\,\nabla_{\mathsf{s}}\sigma(m)\big),
\end{aligned}$$

from which it follows that

$$\nabla_{f\mathsf{s}}\sigma = f\,\nabla_{\mathsf{s}}\sigma. \tag{8.31}$$

For the behaviour in the second argument, we replace $\sigma$ by $\sigma + f\overline{\eta}$, with $f \in C^\infty(M)$ and $\overline{\eta} \in \mathrm{Sec}(\overline{\pi})$. Denoting the linear covariant derivative coming from the restriction $\overline{K}$ by $\overline{\nabla}$, we compute in the same way, using (8.28) and

(8.29):

$$\begin{aligned}
\bar\iota\big(\nabla_{\mathsf{s}}(\sigma+f\overline{\eta})(m)\big) &= \bar\iota\Big(K\big(\mathsf{s}(m),T(\sigma+f\overline{\eta})(\varrho(\mathsf{s}(m)))\big)\Big)\\
&= \tilde K\big(\mathsf{s}(m),T(\iota\sigma+f\overline{\iota\eta})(\varrho(\mathsf{s}(m)))\big) = \tilde\nabla_{\mathsf{s}}(\iota\sigma+f\overline{\iota\eta})(m)\\
&= \tilde\nabla_{\mathsf{s}}\iota\sigma(m)+f(m)\big(\tilde\nabla_{\mathsf{s}}\overline{\iota\eta}\big)(m)+\varrho(\mathsf{s})(f)(m)\,\overline{\iota\eta}(m)\\
&= \tilde K\big(\mathsf{s}(m),T\iota\circ T\sigma(\varrho(\mathsf{s}(m)))\big)+f(m)\,\tilde K\big(\mathsf{s}(m),T\bar\iota\circ T\overline{\eta}(\varrho(\mathsf{s}(m)))\big)\\
&\quad +\varrho(\mathsf{s})(f)(m)\,\overline{\iota\eta}(m) \;=\; \bar\iota\Big(K\big(\mathsf{s}(m),T\sigma(\varrho(\mathsf{s}(m)))\big)\Big)\\
&\quad + f(m)\,\bar\iota\Big(\overline{K}\big(\mathsf{s}(m),T\overline{\eta}(\varrho(\mathsf{s}(m)))\big)\Big)+\varrho(\mathsf{s})(f)(m)\,\overline{\iota\eta}(m)\\
&= \bar\iota\Big(\nabla_{\mathsf{s}}\sigma(m)+f(m)\,\overline{\nabla}_{\mathsf{s}}\overline{\eta}(m)+\varrho(\mathsf{s})(f)(m)\overline{\eta}(m)\Big).
\end{aligned}$$

This expresses that we have the property:

$$\nabla_{\mathsf{s}}(\sigma+f\overline{\eta}) = \nabla_{\mathsf{s}}\sigma+f\,\overline{\nabla}_{\mathsf{s}}\overline{\eta}+\varrho(\mathsf{s})(f)\,\overline{\eta}. \tag{8.32}$$

In coordinates we have, for $\mathsf{s}=\mathsf{s}^a(x)\mathsf{e}_a$ and $\sigma = e_0+\sigma^\alpha(x)\overline{e}_\alpha$:

$$\nabla_{\mathsf{s}}\sigma = \left(\frac{\partial\sigma^\alpha}{\partial x^I}\varrho^I_a(x)+\Gamma^\alpha_{a0}(x)+\Gamma^\alpha_{a\beta}(x)\sigma^\beta(x)\right)\mathsf{s}^a(x)\,\overline{e}_\alpha. \tag{8.33}$$

As one can see, the linearity in $\mathsf{s}$ makes that the value of $\nabla_{\mathsf{s}}\sigma$ at a point $m$ only depends of the value of $\mathsf{s}$ at $m$, so that the usual extension works, whereby for any fixed $\mathsf{v}\in\mathsf{V}$, $\nabla_{\mathsf{v}}$ is a map from $\mathrm{Sec}(\pi)$ to $\overline{E}_m$, defined by $\nabla_{\mathsf{v}}\sigma=\nabla_{\mathsf{s}}\sigma(m)$, for any $\mathsf{s}$ such that $\mathsf{s}(m)=\mathsf{v}$.

**Proposition 8.10.** *An affine $\varrho$-connection $h$ on $\pi$ is uniquely characterised by the existence of an operator $\nabla : \mathrm{Sec}(\tau)\times\mathrm{Sec}(\pi)\to\mathrm{Sec}(\overline{\pi})$ and an associated $\overline{\nabla} : \mathrm{Sec}(\tau)\times\mathrm{Sec}(\overline{\pi})\to\mathrm{Sec}(\overline{\pi})$, such that $\nabla$ is $\mathbb{R}$-linear in its first argument, $\overline{\nabla}$ satisfies the requirements for the determination of a linear $\varrho$-connection on $\overline{\pi}$, and the properties (8.31) and (8.32) hold true.*

Proof: Given an affine $\varrho$-connection $h$ on $\pi$, the existence of operators $\nabla$ and $\overline{\nabla}$ with the required properties has been demonstrated above. Assume conversely that such operators are given. Then, there exists an extension $\tilde\nabla : \mathrm{Sec}(\tau)\times\mathrm{Sec}(\tilde\pi)\to\mathrm{Sec}(\tilde\pi)$, which is defined as follows. Every $\tilde\sigma\in\mathrm{Sec}(\tilde\pi)$ locally is either of the form $\tilde\sigma = f\,\iota(\sigma)$ for some $\sigma\in\mathrm{Sec}(\pi)$ or of the form $\tilde\sigma=\bar\iota(\overline{\eta})$ for some $\overline{\eta}\in\mathrm{Sec}(\overline{\pi})$. In the first case, we put

$$\tilde\nabla_{\mathsf{s}}\tilde\sigma = f\,\bar\iota(\nabla_{\mathsf{s}}\sigma)+\varrho(\mathsf{s})(f)\iota(\sigma); \tag{8.34}$$

in the second case, we put

$$\tilde{\nabla}_{\mathsf{s}}\tilde{\sigma} = \bar{\iota}(\overline{\nabla}_{\mathsf{s}}\overline{\eta}). \tag{8.35}$$

We further impose $\tilde{\nabla}$ to be $I\!R$-linear in its second argument. $I\!R$-linearity as well as $C^\infty(M)$-linearity in the first argument trivially follows from the construction. It is further easy to verify that for $g \in C^\infty(M)$: $\tilde{\nabla}_{\mathsf{s}}(g\tilde{\sigma}) = g\,\tilde{\nabla}_{\mathsf{s}}\tilde{\sigma} + \varrho(\mathsf{s})(g)\,\tilde{\sigma}$. Indeed, in the case that $\tilde{\sigma} = f\,\iota(\sigma)$, for example, we have

$$\begin{aligned}
\tilde{\nabla}_{\mathsf{s}}(g\tilde{\sigma}) &= gf\,\bar{\iota}(\nabla_{\mathsf{s}}\sigma) + \big(f\,\varrho(\mathsf{s})(g) + g\,\varrho(\mathsf{s})(f)\big)\iota(\sigma)\\
&= g\,\tilde{\nabla}_{\mathsf{s}}\tilde{\sigma} + \varrho(\mathsf{s})(g)\,\tilde{\sigma},
\end{aligned}$$

and likewise for the other case. Following [11] we thus conclude that $\tilde{\nabla}$ uniquely determines a linear $\varrho$-connection on $\tilde{\pi}$ by the following construction: for each $(\tilde{e},\mathsf{v}) \in \tilde{\pi}^*\mathsf{V}$, take any $\tilde{\psi} \in \mathrm{Sec}(\tilde{\pi})$ for which $\tilde{\psi}(\tau(\mathsf{v})) = \tilde{e}$, and put

$$\tilde{h}(\tilde{e},\mathsf{v}) = T\tilde{\psi}(\varrho(\mathsf{v})) - (\tilde{\nabla}_{\mathsf{v}}\tilde{\psi})^V_{\tilde{e}},$$

where the last term stands for the element $\tilde{\nabla}_{\mathsf{v}}\tilde{\psi}(\tau(\mathsf{v})) \in \tilde{E}_{\tau(\mathsf{v})}$, vertically lifted to a vector tangent to the fibre of $\tilde{E}$ at $\tilde{e}$.

Likewise, we define a fibre linear map $h : \pi^*\mathsf{V} \to TE$ by

$$h(e,\mathsf{v}) = T\psi(\varrho(\mathsf{v})) - (\nabla_{\mathsf{v}}\psi)^V_e,$$

which can be seen to be independent of the choice of a section $\psi$ for which $\psi(\tau(\mathsf{v})) = e$. It is obvious that $h$ satisfies the requirements of a $\varrho$-connection on $\pi$. It remains to show that $\tilde{h}\circ\iota = T\iota\circ h$. We have

$$\begin{aligned}
\tilde{h}(\iota(e),\mathsf{v}) &= T(\iota\psi)(\varrho(\mathsf{v})) - \big(\tilde{\nabla}_{\mathsf{v}}(\iota\circ\psi)\big)^V_{\iota(e)}\\
&= T(\iota\circ\psi)(\varrho(\mathsf{v})) - (\bar{\iota}\nabla_{\mathsf{v}}\psi)^V_{\iota(e)}\\
&= T\iota\circ T\psi(\varrho(\mathsf{v})) - T\iota\big((\nabla_{\mathsf{v}}\psi)^V_e\big)\\
&= T\iota(h(e,\mathsf{v})),
\end{aligned}$$

which completes the proof. □

Another interesting question one can raise in this context is about the circumstances under which a linear $\varrho$-connection $\tilde{h}$ on $\tilde{\pi}$ is associated to an affine $\varrho$-connection $h$ on $\pi$ in the sense of Definition 8.7. A simple look at coordinate expressions leads to the following result with a global meaning.

**Proposition 8.11.** *A linear $\varrho$-connection on $\tilde{\pi}$ is associated to an affine $\varrho$-connection on $\pi$ if and only if $e^0$ is parallel.*

PROOF: For the covariant derivative operator $\tilde{\nabla}$ associated to a linear $\tilde{h}$, we have for the local basis of $\mathrm{Sec}(\tilde{\pi})$:

$$\tilde{\nabla}_{\mathsf{s}} e_A = \mathsf{s}^a \tilde{\Gamma}^B_{aA}\, e_B,$$

($A$ stands for either 0 or $\alpha$) and by duality, for the basis of $\mathrm{Sec}(\pi^\dagger)$:

$$\tilde{\nabla}_{\mathsf{s}} e^A = -\mathsf{s}^a \tilde{\Gamma}^A_{aB}\, e^B.$$

It follows that $\tilde{\nabla}_{\mathsf{s}} e^0 = 0 \;\Leftrightarrow\; \tilde{\Gamma}^0_{aB} = 0$. The restriction of $\tilde{h}$ to $\iota(E)$ then defines an affine $\varrho$-connection on $\pi$. □

WARNING: Before continuing, we will adopt the following convention. In the following, we won't make a notational distinction anymore between a point in $E$ and its injection in $\tilde{E}$ (that is $e$ will stand for both $e \in E$ and $\iota(e) \in \tilde{E}$) and likewise for a vector in $\overline{E}$. Similarly, if for example, $\sigma$ denotes a section of the affine bundle, the same symbol will be used for its injection in $\mathrm{Sec}(\tilde{\pi})$ and even for the section $\sigma \circ \pi$ of $\pi^*\tilde{\pi}$ (if one looks at a section of $\pi^*\tilde{\pi}$ as a map $\tilde{X} : E \to \tilde{E}$ such that $\tilde{\pi} \circ \tilde{X} = \pi$). We trust that the meaning will be clear from the context.

## 8.6 Affine connections for time-dependent SODEs

We have focussed our main efforts on understanding in detail what affineness of a $\varrho$-connection means. Perhaps the simplest example of the natural appearance of an affine $\varrho$-connection (though for a trivial $\varrho$), is the following. Take $E$ to be the first-jet manifold $J^1M$ of a manifold $M$ which is fibred over $\mathbb{R}$, and $\mathsf{V} = TM$ with $\varrho = id_{TM}$. Then $T^\varrho E = TE$ and we are in the situation which we have been studying extensively in the first chapters. Every SODE on $J^1M$, say

$$\Gamma = \frac{\partial}{\partial t} + v^i \frac{\partial}{\partial x^i} + f^i(t,x,v) \frac{\partial}{\partial v^i}, \tag{8.36}$$

defines a non-linear connection on $\pi_M$ by means of the construction in Proposition 2.4. To say that the forces $f^i$ are quadratic in the velocities, i.e. are of the form

$$f^i = f^i_0(t,x) + f^i_j(t,x) v^j + f^i_{jk}(t,x) v^j v^k, \tag{8.37}$$

is an invariant condition and clearly gives rise then to a connection of affine type, as discussed in the previous section. Examples of such systems in classical mechanics are abundant. In fact, every holonomic Lagrangian system, for which the transformation formulas to generalised coordinates are time-dependent and non-linear, falls into this category, with non-constant coefficients in the quadratic expressions (8.37).

But there is more: even when the forces $f^i$ of the SODE are arbitrary, there are linear connections around. Indeed: it is time to come back to our Berwald-type connections, in particular in their appearance as linear connections on the pullback bundle $\pi_M^*\tau_M : \pi_M^*TM \to J^1M$. It is easy to see that the bundle $\pi_M^*\tau_M$ is the bi-dual $\tilde{\pi}$ of the affine bundle $\pi = \pi_M^*\pi_M : \pi_M^*J^1M \to J^1M$. $\pi_M^*\pi_M$ itself is modelled on the vector bundle $\overline{\pi} = \pi_M^*\overline{\pi}_M : \pi_M^*VM \to J^1M$. On the other hand, the tangent bundle $\tau_{J^1M} : TJ^1M \to J^1M$ takes the place of the arbitrary vector bundle $\tau : \mathsf{V} \to M$ in the general set-up of the previous section (while the anchor map still is the identity).

In our search for finding the most 'optimal' representative of the class of Berwald-type connections, we have made some successive restrictions on the set of linear connections on $\pi_M^*\tau_M$, which has led us finally to the connection $\hat{\mathrm{D}}$ given by expression (2.52). So, the natural question which arises is the following: Which linear connections in the class of Berwald-type connections are in fact derived from an affine connection on $\pi_M^*\pi_M$? Let us have a look at the statement in Proposition 8.11 in the current set-up.

Here, the one-form $dt$ (regarded as a basic one-form along $\pi_M$) is the global one-form $e^0$ which should be parallel for a linear connection D on $\pi_M^*\tau_M$ to be associated to an affine connection on $\pi_M^*\pi_M$. Using the decomposition (2.1) of $\mathcal{X}(\pi_M)$, the condition in Proposition 8.11 breaks down in two conditions: first, for any $\overline{X} \in \overline{\mathcal{X}}(\pi_M)$,

$$0 = (\mathrm{D}_\xi dt)(\overline{X}) = \mathrm{D}_\xi(\langle \overline{X}, dt\rangle) - \langle \mathrm{D}_\xi \overline{X}, dt\rangle = -\langle \mathrm{D}_\xi \overline{X}, dt\rangle, \qquad \forall \xi \in \mathcal{X}(J^1M).$$

Not surprisingly, we recover the condition (2.24) from which we started our characterisation of Finsler-type connections on $\pi_M^*\tau_M$. Recall that, within the framework of Finsler-type connections, the action of a Berwald-type connection on $\overline{\mathcal{X}}(\pi_M)$ is fixed by means of the explicit formula (2.34). The discussion about choosing an 'optimal' Berwald-type connection mainly dealt with the question how to fix the freedom in the time-component. For affine connections, we next find for the canonical vector field,

$$0 = (\mathrm{D}_\xi dt)(\mathbf{T}) = \mathrm{D}_\xi(\langle \mathbf{T}, dt\rangle) - \langle \mathrm{D}_\xi \mathbf{T}, dt\rangle = -\langle \mathrm{D}_\xi \mathbf{T}, dt\rangle, \qquad \forall \xi \in \mathcal{X}(J^1M),$$

or

$$\mathrm{D}_\xi \mathbf{T} \in \overline{\mathcal{X}}(\pi_M), \qquad \forall \xi \in \mathcal{X}(J^1 M). \tag{8.38}$$

In Section 2.8, we found that, among others, condition (2.51) was necessary to ensure that covariant derivation preserves the characteristics of the lifts (2.46). The two conditions (2.51) and (8.38) together are thus contradictory unless the linear connection satisfies $\mathrm{D}_\xi \mathbf{T} = 0$. This last requirement completed the definition of the Berwald-type connection (2.52). Remark that also the linear connection (2.35) of Crampin *et al* satisfies (8.38) and can thus be associated to an affine connection on $\pi_M^* \pi_M$. In the next chapter, within the more general set-up of $\varrho$-connections on $\pi$, we will explain the difference between these two affine connections by means of their rule of parallel transport. Before doing so, we should discuss first the relation between parallel and Lie transport for affine connections on $\pi$.

## 8.7 Parallel transport and Lie transport for affine generalised connections

A curve in $\pi^*\mathsf{V}$ is a couple $(\psi, \mathsf{c})$, where $\psi$ is a curve in $E$ and $\mathsf{c}$ a curve in $\mathsf{V}$ with the properties that the projected curves on $M$ coincide: $\psi_M = c_M$ (in taking a curve in $\pi^*\mathsf{V}$ we will suppose that $I = [a,b] \subset Dom((\psi, \mathsf{c}))$ is an interval in $Dom(\psi) \cap Dom(\mathsf{c})$). The *horizontal lift of the curve* $(\psi, \mathsf{c})$ is a curve $(\psi, \mathsf{c})^H$ in $T^\varrho E$, determined by

$$(\psi, \mathsf{c})^H : u \mapsto (\mathsf{c}(u), h(\psi(u), \mathsf{c}(u))) \qquad \text{for all } u \in I. \tag{8.39}$$

We also say that $\psi$ in $E$ is *a horizontal lift of* $\mathsf{c}$, and write $\psi = c^h$, if $(\psi, \mathsf{c})^H$ is a $\varrho^1$-admissible curve. Since by construction $\pi^1 \circ (\psi, \mathsf{c})^H = \psi$, this means that

$$\dot\psi(u) = \varrho^1 \circ (\psi, \mathsf{c})^H(u) = h(\psi(u), \mathsf{c}(u)), \tag{8.40}$$

which automatically implies that $\mathsf{c}$ must be $\varrho$-admissible, since $\dot c_M = \dot\psi_M = T\pi \circ \dot\psi = T\pi \circ h(\psi, \mathsf{c}) = \varrho \circ \mathsf{c}$. Given $\mathsf{c}$, with $c_M(a) = m$ say, we will write $c_e^h$ for the unique solution $\psi$ of (8.40) passing through the point $e \in E_m$ at $u = a$ (i.e. $c_e^h(a) = e$) and denote the lift $(c_e^h, \mathsf{c})^H$ by $\dot c_e^H$, where, of course, the 'dot' merely refers to the fact that $\varrho^1(\dot c_e^H) = \dot c_e^h$. Curves of the form $\dot c_e^H$ are $\varrho^1$-admissible by construction.

Let $\mathsf{s}$ be a section of $\mathsf{V} \to M$, which we regard as section of $\pi^*\mathsf{V} \to E$ via the composition with $\pi$. In that sense, we can talk about the vector field $h(\mathsf{s}) \in \mathcal{X}(E)$, which has the following interesting characteristics.

**Lemma 8.12.** *For any* $\mathsf{s} \in \mathrm{Sec}(\tau)$*, the vector field* $h(\mathsf{s})$ *on* $E$ *(with* $\mathsf{s}$ *here regarded as a section of* $\pi^*\mathsf{V} \to E$*) has the property that all its integral curves are horizontal lifts of* $\varrho$*-admissible curves in* $\mathsf{V}$*.*

Proof: Let $\gamma_e$ denote an integral curve of $h(\mathsf{s})$ through the point $e$. Since $h(\mathsf{s})$ is $\pi$-related to $\varrho(\mathsf{s}) \in \mathcal{X}(M)$, $\pi \circ \gamma_e$ then is an integral curve of $\varrho(\mathsf{s})$ through $m = \pi(e)$, which we shall call $c_m$. Obviously, $\mathsf{c} = \mathsf{s}(c_m)$ now is a $\varrho$-admissible curve in $\mathsf{V}$ and we have for all $u$ in its domain,

$$h(\mathsf{s})(c_e^h(u)) = h\Big(c_e^h(u), \mathsf{s}\big(\pi(c_e^h(u))\big)\Big) = h\big(c_e^h(u), \mathsf{c}(u)\big) = \dot{c}_e^h(u), \quad (8.41)$$

which shows that $\gamma_e = c_e^h$. □

So far, the above characterisation of a horizontal lift applies to any $\varrho$-connection on $\pi$. If, in particular, the connection is *affine*, then we can express the definition of a horizontal lift also in the more standard approach, i.e. in terms of the operator $\nabla$. Indeed, for any $\varrho$-admissible curve $\mathsf{c}$ in $\mathsf{V}$ and any curve $\psi$ with $c_M = \psi_M$, we can define a new curve $\nabla_{\mathsf{c}}\psi$ by a formula which is formally identical to (8.15).

$$\nabla_{\mathsf{c}}\psi(u) = K((\mathsf{c}(u), \dot{\psi}(u))) = \Big(\dot{\psi}(u) - h\big((\psi(u), \mathsf{c}(u))\big)\Big)_v. \quad (8.42)$$

Note, however, that $\nabla_{\mathsf{c}}\psi$ is a curve in $\overline{E}$ now. Nevertheless, it makes perfect sense to say that $\psi$ in $E$ is a horizontal lift of $\mathsf{s}$ if the associated curve $\nabla_{\mathsf{c}}\psi$ in $\overline{E}$ is zero for all $u$. Obviously, $\psi = c_e^h$ iff $\psi(a) = e$ and $\nabla_{\mathsf{c}}\psi = 0$ for all $u \in I$.

Putting $c_M(b) = m_b$, the point $c_e^h(b) \in E_{m_b}$ is called the *parallel translate of* $e$ *along* $\mathsf{c}$. It is instructive to see in detail how for different initial values in a fixed fibre of $E$, one can get an affine action between the affine fibres of $E$. Take $e_1, e_2 \in E_m$ and consider the horizontal lifts $c_{e_1}^h$ and $c_{e_2}^h$. Denote the difference $e_1 - e_2$ by $\overline{e} \in \overline{E}_m$ and put $\overline{\eta}_{\overline{e}} := c_{e_1}^h - c_{e_2}^h$. As the subscript indicates, $\overline{\eta}_{\overline{e}}$ is a curve in $\overline{E}$ starting at $\overline{e}$. From the action of $\nabla$ on curves (8.42) and the affineness-property (8.23), it easily follows that

$$\nabla_{\mathsf{c}} c_{e_1}^h(u) - \nabla_{\mathsf{c}} c_{e_2}^h(u) = \overline{\nabla}_{\mathsf{c}}\overline{\eta}_{\overline{e}}(u), \qquad \text{for all } u \in I.$$

Since both $c_{e_1}^h$ and $c_{e_2}^h$ are solutions of the equation $\nabla_{\mathsf{c}}\psi = 0$, $\overline{\eta}_{\overline{e}}$ must be the unique solution of the initial value problem $\overline{\nabla}_{\mathsf{c}}\overline{\eta} = 0$, $\overline{\eta}(a) = \overline{e}$, i.e. the unique $\overline{h}$-horizontal lift through $\overline{e}$. Therefore, *the difference between the* $\nabla$*-parallel translates of* $e_1$ *and* $e_2$ *along* $\mathsf{c}$ *is the* $\overline{\nabla}$*-parallel translated of* $e_1 - e_2$ *along* $\mathsf{c}$. In fact, this property is necessary and sufficient for the connection to be *affine*. From now on we will use the notation $c_{\overline{e}}^{\overline{h}}$ for $c_{e_1}^h - c_{e_2}^h$.

**Proposition 8.13.** *A $\varrho$-connection $h$ on $\pi$ is affine if and only if there exists a linear $\varrho$-connection $\overline{h}$ on $\overline{\pi}$, such that for all admissible curves $\mathsf{c}$ and for any two points $e_1, e_2 \in E_m$ ($m = c_M(a)$) the difference $c^h_{e_1}(u) - c^h_{e_2}(u)$ is the $\overline{\nabla}$-parallel translate of $e_1 - e_2$ along $\mathsf{c}$.*

Proof: The proof in one direction has already been given. For the converse, suppose that a *linear* connection $\overline{h}$ exists, having the above properties. It suffices to show that $\overline{h}$ is related to $h$ by means of (8.23). Choosing a $\mathsf{v} \in \mathsf{V}_m$ arbitrarily, we take a $\varrho$-admissible curve $\mathsf{c}$, that passes through it (for $u = a$) and consider its $h$-horizontal lift $c^h_e$ through $e$ and its $\overline{h}$-horizontal lift through $\overline{e}$. By assumption we know that $c^h_{e+\overline{e}} - c^h_e$ is the $\overline{h}$-horizontal lift of $\mathsf{c}$ through $\overline{e}$, i.e.

$$\Sigma(c^h_e(u), c^{\overline{h}}_{\overline{e}}(u)) = c^h_{e+\overline{e}}(u) \qquad \text{for all } u \in I.$$

Taking the derivative of this expression at $u = a$, we get

$$T\Sigma_{(e,\overline{e})}\big(\dot{c}^h_e(a), \dot{c}^{\overline{h}}_{\overline{e}}(a)\big) = \dot{c}^h_{e+\overline{e}}(a). \tag{8.43}$$

In view of (8.40) and its analogue for $\overline{h}$, this is indeed what we wanted to show. □

Let us introduce corresponding 'flow-type' maps. For that purpose, it is convenient to use (temporarily) the more accurate notation $c^h_{a,e}$ for the horizontal lift which passes through $e$ at $u = a$. Putting

$$\phi^h_{u,a}(e) = c^h_{a,e}(u), \qquad u \in [a, b],$$

the result of the preceding proposition can equivalently be expressed as,

$$\phi^h_{u_2,u_1}(e + \overline{e}) = \phi^h_{u_2,u_1}(e) + \phi^{\overline{h}}_{u_2,u_1}(\overline{e}),$$

i.e. $\phi^h_{u_2,u_1} : E_{c_M(u_1)} \to E_{c_M(u_2)}$ is an affine map with linear part $\phi^{\overline{h}}_{u_2,u_1}$. Its tangent map, therefore, can be identified with its linear part. As a result, when we consider Lie transport of vertical vectors in $TE$ along the curve $c^h_{a,e}$, in the case of an affine connection, the image vectors will come from the parallel translate associated to the linear connection $\overline{h}$. Indeed, putting $Y_a = v(e, \overline{e})$ and defining its *Lie translate* from $a$ to $b$ as $Y_b = T\phi^h_{b,a}(Y_a)$, we have for each $g \in C^\infty(E)$,

$$\begin{aligned} Y_b(g) = Y_a(g \circ \phi^h_{b,a}) &= \frac{d}{dt}\big(g \circ \phi^h_{b,a}(e + t\overline{e})\big)_{t=0} \\ &= \frac{d}{dt}\big(g(\phi^h_{b,a}(e) + t\phi^{\overline{h}}_{b,a}(\overline{e}))\big)_{t=0} = v\big(c^h_{a,e}(b), c^{\overline{h}}_{a,\overline{e}}(b)\big)g, \end{aligned}$$

where the transition to the last line requires affineness of the connection. It follows that in the affine case, the Lie translate of $Y_a = v(e, \overline{e})$ to $b$ is given by

$$Y_b = v\big(c^h_{a,e}(b), c^{\overline{h}}_{a,\overline{e}}(b)\big). \tag{8.44}$$

At this point, it is appropriate to make a few more comments about the general idea of Lie transport. If (on an arbitrary manifold) $Y$ is a vector field along an integral curve of some other vector field $X$, and we therefore have the genuine (local) flow $\phi_s$ of $X$ at our disposal, then the Lie derivative of $Y$ with respect to $X$ is defined to be

$$\mathcal{L}_X Y(u) = \frac{d}{ds}\big(T\phi_{-s}(Y(s+u))\big)_{s=0}.$$

As shown for example in [24] (p. 68), the requirement $\mathcal{L}_X Y = 0$, subject to some initial condition, $Y(0) = Y_0$ say, then uniquely determines a vector field $Y$ along an integral curve of $X$ in such a way that $Y(u)$ is obtained by Lie transport of $Y_0$. The description of Lie transport, therefore, is more direct when we are in the situation of an integral curve of a vector field.

Lemma 8.12, unfortunately, does not create such a situation for us because, when an admissible curve $\mathsf{c}$ is given, together with one of its horizontal lifts $c^h_e$, it does not provide us with a way of constructing a vector field which has $c^h_e$ as one of its integral curves. The complication for constructing such a vector field primarily comes from the fact that the differential equations (8.40) which define $c^h_e$ are non-autonomous. The usual way to get around this problem is to make the system autonomous by adding an extra dimension. A geometrical way of achieving this here, which takes into account that $c^h_e$ in the first place has to be a curve projecting onto $c_M$, is obtained by passing to the pullback bundle $c^*_M E \to I \subset I\!R$. We introduce the notation $c^1_M : c^*_M E \to E, (u,e) \mapsto (c_M(u), e)$, and likewise $\overline{c}^1_M : c^*_M \overline{E} \to \overline{E}, (u, \overline{e}) \mapsto (c_M(u), \overline{e})$. With the help of these maps, we can single out vector fields $\Lambda_c \in \mathcal{X}(c^*_M E)$ and $\overline{\Lambda}_c \in \mathcal{X}(c^*_M \overline{E})$ as follows.

**Proposition 8.14.** *For any $\varrho$-connection $h$ on $\pi$ and given $\varrho$-admissible curve $\mathsf{c}$ in $\mathsf{V}$, there exists a unique vector field $\Lambda_c$ on $c^*_M E$ that projects on the coordinate vector field on $I\!R$ and is such that*

$$Tc^1_M(\Lambda_c(u,e)) = h\big(c^1_M(u,e), \mathsf{c}(u)\big), \tag{8.45}$$

*for all $(u,e) \in c^*_M E$. Likewise, if the connection is affine with linear part $\overline{h}$, there exists a unique vector field $\overline{\Lambda}_c$ on $c^*_M \overline{E}$ that projects on the coordinate* *vector field on* $\mathbb{R}$ *and is such that*

$$T\overline{c}^1_M(\overline{\Lambda}_c(u,\overline{e})) = \overline{h}\big(\overline{c}^1_M(u,\overline{e}), \mathsf{c}(u)\big), \tag{8.46}$$

*for all* $(u,\overline{e}) \in c_M^*\overline{E}$.

Proof: The proof is analogous for both cases; we prove only the first. Let $u$ denote the coordinate on $\mathbb{R}$ and $y^\alpha$ the fibre coordinates of some $e \in (c_M^*E)_u$. Representing the given curve as $\mathsf{c} : u \mapsto (x^I(u), \mathsf{c}^a(u))$ and putting $\Lambda_c(u,e) = U(u,e)\frac{\partial}{\partial u}|_{(u,e)} + Y^\alpha(u,e)\frac{\partial}{\partial y^\alpha}|_{(u,e)}$, one finds that

$$Tc^1_M(\Lambda_c(u,e)) = (x^I(u), y^\alpha; U(u,e)\dot{x}^I(u), Y^\alpha(u,e)).$$

On the other hand

$$h\big(c^1_M(u,e), \mathsf{c}(u)\big) = \big(x^I(u), y^\alpha; \mathsf{c}^a(u)\varrho^I_a(c_M(u)), -\mathsf{c}^a(u)\Gamma^\alpha_a(c_M(u),e)\big).$$

Identification of the two expressions gives that $Y^\alpha(u,e) = -\mathsf{c}^a(u)\Gamma^\alpha_a(c_M(u),e)$, and that $U\dot{x}^I = \varrho^I_a\mathsf{c}^a$. At points $u$ where $\mathsf{c}(u)$ does not lie in the kernel of $\varrho$, the latter equality would by itself determine $U$ to be 1. The extra projectability requirement ensures that this will hold also when $\dot{x}^I = \varrho^I_a\mathsf{c}^a = 0$. □

An interesting point here is that the complication about ensuring separately that $U = 1$ in some sense disappears when we look at horizontality on $T^\varrho E$ rather than on $TE$, i.e. horizontality in the sense of (8.39). To see this, observe first that we can use the linear bundle map $\mathsf{f} : T\mathbb{R} \to \mathsf{V}, U\frac{d}{du}|_u \mapsto U\mathsf{c}(u)$ over $c_M : I \to M$, to obtain, in accordance with (7.7), the following extended notion of tangent map:

$$T^\varrho c^1_M : T(c_M^*E) \to T^\varrho E, \lambda \mapsto (\mathsf{f}(T\tau_{\boldsymbol{R}}\lambda), Tc^1_M(\lambda)).$$

Here $\tau_{\boldsymbol{R}}$ is the bundle projection of $c_M^*E \to \mathbb{R}$. $T^\varrho c^1_M$ is well defined, since $T\tau_{\boldsymbol{R}}\lambda$ is of the form $U\frac{d}{du}|_u$ and $T\pi \circ Tc^1_M(\lambda) = T(\pi\circ c^1_M)(\lambda) = T(c_M \circ \tau_{\boldsymbol{R}})(\lambda) = Tc_M \circ T\tau_{\boldsymbol{R}}(\lambda) = Tc_M(U\frac{d}{du}|_u) = U\dot{c}_M(u) = U\varrho(\mathsf{c}(u)) = \varrho(\mathsf{f}(U\frac{d}{du}|_u)) = \varrho(\mathsf{f}(T\tau_{\boldsymbol{R}}\lambda))$. Notice that this remains true also at points where $\mathsf{c}$ lies in the kernel of $\varrho$. Now, $\Lambda_c$ can be defined as the unique vector field on $c_M^*E$ for which

$$T^\varrho c^1_M(\Lambda_c(u,e)) = \big(c^1_M(u,e), \mathsf{c}(u)\big)^H \qquad \text{for all } (u,e) \in c_M^*E.$$

Indeed, the second component of this equality is just the condition (8.45) again, whereas the first component says that $U\mathsf{c}(u) = \mathsf{c}(u)$ and thus implies

$U = 1$ (except in the very special case of a constant curve $\mathsf{c}(u) = (m_0, 0) \in \mathsf{V}$).

We will look now at the integral curves of $\Lambda_c$ and $\overline{\Lambda}_c$. Since $\Lambda_c$ projects on the coordinate field on $I\!R$, the integral curves are essentially sections of $c_M^* E \to I\!R$. Let $\gamma_e$ denote the integral curve going through $(a, e)$ at time $u = a$, so that

$$\dot{\gamma}_e(u) = \Lambda_c(\gamma_e(u)) \qquad \forall u \in I', \tag{8.47}$$

where $I'$ is some interval, possibly smaller than the domain $I$ of $\mathsf{c}$. In a similar way, we will write $\overline{\gamma}_{\overline{e}}$ for the integral curve of $\overline{\Lambda}_c$ through $(a, \overline{e})$.

**Proposition 8.15.** *For any $\varrho$-connection on $\pi$, the curve $c_M^1 \circ \gamma_e$ is the $h$-horizontal lift of $\mathsf{c}$ through $e$. Likewise, if the connection is affine with linear part $\overline{h}$, the curve $\overline{c}_M^1 \circ \overline{\gamma}_{\overline{e}}$ is the $\overline{h}$-horizontal lift of $\mathsf{c}$ through $\overline{e}$.*

Proof: Again, we will prove only the first statement. $c_M^1 \circ \gamma_e$ is a curve in $E$ projecting on the curve $c_M$ in $M$. Using (8.47) and (8.45) we find:

$$\frac{d}{du}(c_M^1 \circ \gamma_e)(u) = Tc_M^1(\dot{\gamma}_e(u)) = Tc_M^1(\Lambda_c(\gamma_e(u))) = h(c_M^1 \circ \gamma_e(u), c(u)),$$

which shows that $c_M^1 \circ \gamma_e = c_e^h$. □

We now proceed to look at Lie transport along integral curves of the vector field $\Lambda_c$ on $c_M^* E$. It is clear that $c_M^* E \to I\!R$ is an affine bundle modelled on the vector bundle $c_M^* \overline{E} \to I\!R$. As for any affine bundle, there is a vertical lift map, which we will denote by $v_{c_M^* E}$ which maps elements of $(c_M^* E \times_{I\!R} c_M^* \overline{E})$ to vertical vectors of $T(c_M^* E)$. We will consider Lie transport of such vertical vectors along integral curves of $\Lambda_c$. Starting from $e \in E_{c_M(a)}$, $\overline{e} \in \overline{E}_{c_M(a)}$ and putting $\Upsilon_{e,\overline{e}}(a) = v_{c_M^* E}\big((a, e), (a, \overline{e})\big)$, we know that the condition $\mathcal{L}_{\Lambda_c} \Upsilon_{e,\overline{e}} = 0$ uniquely defines a vector field along the integral curve $\gamma_{(a,e)}$ of $\Lambda_c$ ($\gamma_{(a,e)}(a) = (a, e)$) which takes the (vertical) value $\Upsilon_{e,\overline{e}}(a)$ at the point $(a, e)$. As said before, the value of $\Upsilon_{e,\overline{e}}$ at any later $u$ is the Lie translate of $\Upsilon_{e,\overline{e}}(a)$, that is to say, we have $\Upsilon_{e,\overline{e}}(u) = T\phi_{u,a}\big(\Upsilon_{e,\overline{e}}(a)\big)$, where $\phi_{u,a}$ refers to the flow of $\Lambda_c$ ($\phi_{a,a}$ is the identity). It is interesting to observe here that $\Upsilon_{e,\overline{e}}$ is directly related to the Lie translate we discussed before, of vertical vectors on $E$ along the horizontal lift $c_{a,e}^h$. To be precise, with $Y_{e,\overline{e}}(a) = Tc_M^1\big(\Upsilon_{e,\overline{e}}(a)\big) = v(e, \overline{e})$ and defining $Y_{e,\overline{e}}(u)$ to be $T\phi_{u,a}^h\big(Y_{e,\overline{e}}(a)\big)$ as before, we have (at any later time $u$ in the domain of $\gamma_{(a,e)}$)

$$Tc_M^1\big(\Upsilon_{e,\overline{e}}(u)\big) = Y_{e,\overline{e}}(u). \tag{8.48}$$

Indeed, it follows from Proposition 8.15 (using here again the somewhat more accurate notations which take the "initial time" $a$ into account), that $c^1_M \circ \phi_{u,a} = \phi^h_{u,a} \circ c^1_M$. Therefore, we have

$$\begin{aligned} Tc^1_M\big(\Upsilon_{e,\overline{e}}(u)\big) &= Tc^1_M\big(T\phi_{u,a}(\Upsilon_{e,\overline{e}}(a))\big) = T\phi^h_{u,a}\big(Tc^1_M(\Upsilon_{e,\overline{e}}(a))\big) \\ &= T\phi^h_{u,a}\big(Y_{e,\overline{e}}(a)\big) = Y_{e,\overline{e}}(u). \end{aligned}$$

The case of an *affine* $\varrho$-connection is of special interest. The affine nature of the maps $c^h_e(u)$, for fixed $u$, implies via Proposition 8.15 that the flow maps of $\Lambda_c$ are also affine or, expressed differently, that:

$$\gamma_{e+\overline{e}}(u) = \gamma_e(u) + \overline{\gamma}_{\overline{e}}(u). \tag{8.49}$$

We have shown already that in the affine case: $Y_{e,\overline{e}}(u) = v\big(c^h_e(u), c^{\overline{h}}_{\overline{e}}(u)\big)$. The translation of this result via the relation (8.48) means that we have

$$\Upsilon_{e,\overline{e}}(u) = v_{c^*_M E}\big(\gamma_e(u), \overline{\gamma}_{\overline{e}}(u)\big). \tag{8.50}$$

Summarising the more interesting aspects of what we have observed above, we can make the following statement

**Proposition 8.16.** *For an arbitrary $\varrho$-connection on $\pi$, Lie transport of vertical vectors on $E$ along the horizontal lift $c^h_e$ is equivalent to Lie transport of vertical vectors on $c^*_M E$ along integral curves of the vector field $\Lambda_c$. In the particular case that the connection is affine, the translates in both cases are compatible with the affine nature of the flow maps between fixed fibres.*

The next result concerns an important property of the covariant derivative operators $\nabla$ and $\overline{\nabla}$ which become available when the $\varrho$-connection is affine. The preceding considerations about vector fields $\Lambda_c$ on $c^*_M E$ will help to prove it in a purely geometrical way. The reader may wish to skip this rather technical proof and pass to the remark immediately following it.

**Proposition 8.17.** *Let $h$ be an affine $\varrho$-connection. For all $\mathsf{s} \in \mathrm{Sec}(\tau)$, $\overline{\sigma} \in \mathrm{Sec}(\overline{\pi})$ and $\sigma \in \mathrm{Sec}(\pi)$, the brackets $[h\mathsf{s}, v\overline{\sigma}]$ and $[h\mathsf{s}, v\sigma]$ of vector fields on $E$ are vertical and we have*

$$\overline{\nabla}_{\mathsf{s}}\overline{\sigma} = [h\mathsf{s}, v\overline{\sigma}]_v, \qquad \nabla_{\mathsf{s}}\sigma = [h\mathsf{s}, v\sigma]_v. \tag{8.51}$$

Proof: We start with the bracket $[h\mathsf{s}, v\overline{\sigma}]_v$. Since the vector fields under consideration are $\pi$-related to $\varrho(\mathsf{s})$ and the zero vector field on $M$, respectively, their Lie bracket is $\pi$-related to $[\varrho(\mathsf{s}), 0] = 0$ and is therefore vertical. If we project it down to $\overline{E}$, strictly speaking by taking $([h\mathsf{s}, v\overline{\sigma}](e))_v \in$

$E\times_M \overline{E}$ with $e\in E_m$ and looking at the second component, we obtain an element of $\overline{E}_m$ which does not depend on the fibre coordinates of $e$ (as we can see instantaneously by thinking of coordinate expressions). In other words, $[h\mathsf{s}, v\overline{\sigma}]_v$ gives rise to a section of $\overline{\pi}$ which has the following properties: for all $f\in C^\infty(M)$

$$[h(f\mathsf{s}), v\overline{\sigma}]_v = f[h\mathsf{s}, v\overline{\sigma}]_v, \qquad [h\mathsf{s}, v(f\overline{\sigma})]_v = f[h\mathsf{s}, v\overline{\sigma}]_v + \varrho(\mathsf{s})(f)\overline{\sigma}.$$

These are precisely the properties of the covariant derivative operator $\overline{\nabla}_{\mathsf{s}}\overline{\sigma}$, from which it follows that $L(\mathsf{s},\overline{\sigma}) = \overline{\nabla}_{\mathsf{s}}\overline{\sigma} - [h\mathsf{s}, v\overline{\sigma}]_v$ is tensorial in $\mathsf{s}$ and $\overline{\sigma}$.

To prove that $L$ is actually zero, we will use a rather subtle argument, which is based on the following considerations of a quite general nature. If $L(\mathsf{s},\overline{\sigma})$ is tensorial, so that for all $m$, $L(\mathsf{s},\overline{\sigma})(m)$ depends on $\mathsf{s}(m)$ and $\overline{\sigma}(m)$ only, then for each curve $c_M$ in $M$, there exists a corresponding operator $l(\mathsf{r},\overline{\eta})$, acting on arbitrary sections along the curve $c_M$, which is completely determined by the property: if $\mathsf{s}\in \mathrm{Sec}(\tau)$, $\overline{\sigma}\in\mathrm{Sec}(\overline{\pi})$ and putting $\mathsf{r}=\mathsf{s}|_{c_M}$, $\overline{\eta}=\overline{\sigma}|_{c_M}$, then $l(\mathsf{r},\overline{\eta})(u) = L(\mathsf{s},\overline{\sigma})(c_M(u))$. In turn, the value of $L(\mathsf{s},\overline{\sigma})(m)$ at an arbitrary point $m$ can be computed by choosing an arbitrary curve $c_M$ through $m$ ($c_M(0)=m$ say), selecting sections $\mathsf{r}$ and $\overline{\eta}$ along $c_M$ for which $\mathsf{r}(0)=\mathsf{s}(m)$ and $\overline{\eta}(0)=\overline{e}=\overline{\sigma}(m)$, and then computing $l(\mathsf{r},\overline{\eta})(0)$.

We apply this general idea in the following way. Starting from an arbitrary $e\in E_m$, we know from Lemma 8.12 that the integral curve of $h(\mathsf{s})$ through $e$ is the horizontal lift $c_e^h$ of some admissible curve $\mathsf{c}$ through $\mathsf{s}(m)$ (here the "initial time" $a$ is taken to be zero). Take $\mathsf{r}$ to be this curve $\mathsf{c}$ (with projection $c_M$) and choose $\overline{\eta}$ to be the curve $c_{\overline{e}}^{\overline{h}}$. Then,

$$l(\mathsf{r},\overline{\eta})(0) = \overline{\nabla}_{\mathsf{c}} c_{\overline{e}}^{\overline{h}}(0) - \big(\mathcal{L}_{h\mathsf{s}} v(c_e^h, c_{\overline{e}}^{\overline{h}})\big)_v(0).$$

But $\overline{\nabla}_{\mathsf{c}} c_{\overline{e}}^{\overline{h}}$ is zero by construction. Concerning the second term, we observe that $h\mathsf{s}$ is $c_M^1$-related to $\Lambda_c$, by definition of $\Lambda_c$. Also $v(c_e^h, c_{\overline{e}}^{\overline{h}})$ is $c_M^1$-related to the vector field $v_{e,\overline{e}}(u)$ along the integral curve $\gamma_e$ of $\Lambda_c$ through the point $(0,e)\in c_M^*E$ (see (8.48)). But we know that $\mathcal{L}_{\Lambda_c}\Upsilon_{e,\overline{e}}=0$, so that in particular $\big(\mathcal{L}_{h\mathsf{s}} v(c_e^h, c_{\overline{e}}^{\overline{h}})\big)_v(0)=0$. It follows that $\overline{\nabla}_{\mathsf{s}}\overline{\sigma} = [h\mathsf{s}, v\overline{\sigma}]_v$.

For the second part, we should specify in the first place what is meant by $v(\sigma)$: any $\sigma\in\mathrm{Sec}(\pi)$ can be thought of as a section of $\tilde{\pi}$ and then $v(\sigma)(e) = v(e, \vartheta_e(\sigma(\pi(e))))$. Making use of the canonical section $\mathbf{T}$ of $\pi^*\tilde{\pi}$, we can write in fact that $v(\sigma) = v(\sigma - \mathbf{T})$, where $\sigma - \mathbf{T}\in\mathrm{Sec}(\pi^*\overline{\pi})$. It is clear that $[h\mathsf{s}, v\sigma]$ is vertical again, and we find the properties: $\forall f \in C^\infty(M)$,

$$\begin{aligned} [h(f\mathsf{s}), v\sigma]_v &= f[h\mathsf{s}, v\sigma]_v, \\ [h\mathsf{s}, v(\sigma + f\overline{\sigma})]_v &= [h\mathsf{s}, v(\sigma - \mathbf{T}) + fv(\overline{\sigma})]_v \\ &= [h\mathsf{s}, v\sigma]_v + f[h\mathsf{s}, v\overline{\sigma}]_v + \varrho(\mathsf{s})(f)\overline{\sigma} \\ &= [h\mathsf{s}, v\sigma]_v + f\overline{\nabla}_{\mathsf{s}}\overline{\sigma} + \varrho(\mathsf{s})(f)\overline{\sigma}. \end{aligned}$$

Again, these are the characterising properties of the covariant derivative $\nabla_{\mathsf{s}}\sigma$. It follows that the operator $L(\mathsf{s}, \sigma) = \overline{\nabla}_{\mathsf{s}}\sigma - [h\mathsf{s}, v\sigma]_v$ is linear in $\mathsf{s}$ and affine in $\sigma$. This is the analogue, when there are affine components involved, of an operator $L$ being tensorial. The rest of the reasoning follows the same pattern as before. This time, starting from an arbitrary $e \in E_m$ and an integral curve $c_e^h$ of $h(\mathsf{s})$ through $e$, we put $\overline{e} = \sigma(\pi(e)) - e$ and choose the curves $\overline{\eta} = c_{\overline{e}}^{\overline{h}}$ in $\overline{E}$ and $\eta = c_e^h + \overline{\eta}$ in $E$ to obtain a section of $\pi$ along $c_e^h$ which has $\sigma(\pi(e))$ as initial value. □

*Remark.* A more direct, but perhaps geometrically less appealing proof, consists in verifying the statements of Proposition 8.17 by a coordinate calculation. We have, for $\mathsf{s} = \mathsf{s}^a(x)\mathsf{e}_a$ and $\sigma = e_0 + \sigma^\alpha(x)\overline{e}_\alpha$,

$$\nabla_{\mathsf{s}}\sigma = [h\mathsf{s}, v\sigma]_v = \big(\rho_a^i \frac{\partial \sigma^\alpha}{\partial x^i} + \Gamma_{a0}^\alpha(x) + \Gamma_{a\beta}^\alpha(x)\sigma^\beta\big)\mathsf{s}^a(x)\overline{e}_\alpha,$$

and similarly, for $\overline{\sigma} = \overline{\sigma}^\alpha\overline{e}_\alpha$,

$$\overline{\nabla}_{\mathsf{s}}\overline{\sigma} = [h\mathsf{s}, v\overline{\sigma}]_v = \big(\rho_a^i \frac{\partial \overline{\sigma}^\alpha}{\partial x^i} + \Gamma_{a\beta}^\alpha(x)\overline{\sigma}^\beta\big)\mathsf{s}^a(x)\overline{e}_\alpha.$$

**Corollary 8.18.** *If* $(e, \mathsf{v}) \in \pi^*\mathsf{V}$ *and* $\mathsf{s} \in \mathrm{Sec}(\tau)$, $\sigma \in \mathrm{Sec}(\pi)$ *are sections passing through* $\mathsf{v}$ *and* $e$ *respectively, then we have the following relation*

$$h(e, \mathsf{v}) = T\sigma(\varrho(\mathsf{v})) - [h\mathsf{s}, v\sigma](e). \tag{8.52}$$

Proof: It was shown in Proposition 8.10 that a pair of operators having the properties of covariant derivatives $(\nabla, \overline{\nabla})$ uniquely define an affine $\varrho$-connection on $\pi$. The brackets $[h\mathsf{s}, v\sigma]_v$ and $[h\mathsf{s}, v\overline{\sigma}]_v$ constitute such a pair (as shown above) and according to the proof of Proposition 8.10, the right-hand side of (8.52) would then define the associated affine $\varrho$-connection. A priori, however, there is no reason why this would be the $h$ we started from. But the proof of Proposition 8.17 precisely guarantees now that it must be the $h$ we started from, and hence we have (8.52). □

There is of course a similar formula for $\overline{h}$, which reads,

$$\overline{h}(\overline{e}, \mathsf{v}) = T\overline{\sigma}(\varrho(\mathsf{v})) - [h\mathsf{s}, v\overline{\sigma}](\overline{e}). \tag{8.53}$$

As an immediate benefit of the formulas (8.51), we can obtain an explicit defining relation now for the extension of the operators $(\nabla, \overline{\nabla})$ to a covariant derivative $\tilde{\nabla}$ on $\mathrm{Sec}(\tilde{\pi})$. Each $\tilde{\sigma} \in \mathrm{Sec}(\tilde{\pi})$ is either of the form $f\sigma$, with $f \in C^\infty(M)$ and $\sigma \in \mathrm{Sec}(\pi)$, or of the form $\overline{\sigma}$, for some $\overline{\sigma} \in \mathrm{Sec}(\overline{\pi})$. Then, $\tilde{\nabla}_{\mathsf{s}}\tilde{\sigma}$ is defined either by (8.34) or by (8.35).

**Corollary 8.19.** *A unifying formula for the computation of* $\tilde{\nabla}_{\mathsf{s}}\tilde{\sigma}$ *is given by*

$$\tilde{\nabla}_{\mathsf{s}}\tilde{\sigma} = [h\mathsf{s}, v\tilde{\sigma}]_v + \varrho(\mathsf{s})(\langle\tilde{\sigma}, e^0\rangle)\mathcal{I}. \tag{8.54}$$

PROOF: In the first case we have $\tilde{\nabla}_{\mathsf{s}}\tilde{\sigma} = f[h\mathsf{s}, v\sigma]_v + \varrho(\mathsf{s})(f)\sigma$. Using the properties that $\sigma$, regarded as section of $\pi^*\tilde{\pi}$, can be written as $\sigma = \vartheta(\sigma) + \mathcal{I}$, and that $\big(v(\sigma)\big)_v = \vartheta(\sigma)$, this expression can be rewritten as $\tilde{\nabla}_{\mathsf{s}}\tilde{\sigma} = [h\mathsf{s}, v(f\sigma)]_v + \varrho(\mathsf{s})(f)\mathcal{I}$, which is of the form (8.54) since $\langle\tilde{\sigma}, e^0\rangle = f$ in this case. In the second case, we have $\tilde{\nabla}_{\mathsf{s}}\tilde{\sigma} = [h\mathsf{s}, v\overline{\sigma}]_v$, which is immediately of the right form since $\langle\tilde{\sigma}, e^0\rangle = 0$ now. □

The representation (8.54) of $\tilde{\nabla}_{\mathsf{s}}\tilde{\sigma}$ is exactly the decomposition (6.4) of $\tilde{\nabla}_{\mathsf{s}}\tilde{\sigma}$, regarded as section of $\pi^*\tilde{\pi}$. One should not forget, of course, that such a decomposition somehow conceals part of the information in case the section under consideration, as is the case with $\tilde{\nabla}_{\mathsf{s}}\tilde{\sigma}$ here, is basic, in the sense that it is actually a section of $\tilde{\pi} : \tilde{E} \to M$.

# Chapter 9

# *Berwald-type connections – affine generalised case*

## 9.1 Generalised connections of Berwald type

In Section 8.6, we have explained how the Berwald-type connections (2.35) and (2.52) of the time-dependent model fit within the present scheme. We shall now explore to what extent an arbitrary $\varrho$-connection on an affine bundle, in the general picture of the diagram of Section 8.5, has a kind of induced linearisation of Berwald-type, and we intend to unravel in that process the origin of the two specific choices for fixing the connection.

To define $\mathrm{D}_\xi X$, it suffices to specify separately the action of horizontal and vertical vector fields, where "horizontality" is defined of course via the non-linear connection one starts from. In the more general situation of a $\varrho$-connection, however, horizontality of vector fields on $E$ is not an unambiguous notion, in the sense that $\mathrm{Im}\, h$ may not provide a full complement of the set of vertical vectors and may even have a non-empty intersection with this set (see [11]). As said in Section 8.2, we do have a direct complement for the vertical sections of $\pi^1 : T^\varrho E \to E$. So the right way to look here for a linear connection on $\pi^*\tilde{\pi}$ is as a $\varrho^1$-connection.

The linear $\varrho^1$-connection on $\pi^*\tilde{\pi}$ will actually be generated by an affine $\varrho^1$-connection on $\pi^*E \to E$. However, as long as we let $\mathsf{V} \to M$ be any vector bundle, not related to $E \to M$ and without the additional structure of a Lie algebroid, there is no bracket of sections of $\tau$ or $\pi^1$ available. We should, therefore, not expect to discover immediately direct defining relations of the kind (2.35) or (2.52). Instead, we shall approach the problem of detecting corresponding $\varrho^1$-connections on $\pi^*E$ via their covariant derivative operators, for which we will use the results of Proposition 8.17 as one of the sources of inspiration.

An other source of inspiration has already been mentioned briefly in Section 1.2. In [20], in the context of non-linear connections on a tangent bundle, Crampin gives an interesting geometrical characterisation of a Berwald-type connection by fixing the rules of parallel transport it should satisfy. Following the lead of Crampin's approach, we will enter the subject of understanding the details of possible rules of parallel transport first.

Recall that the concept of parallel transport in $E$, i.e. the construction of the horizontal lift $c_e^h$ of a $\varrho$-admissible curve in $\mathsf{V}$, exists for any $\varrho$-connection $h$. If $h$ is affine, we know that for horizontal lifts which start at $e$ and $e_1 = e+\overline{e}$ at an initial time $a$, we have at any later time $b$ that $c_{e+\overline{e}}^h(b) = c_e^h(b) + c_{\overline{e}}^{\overline{h}}(b)$. In addition, $v(e,\overline{e})$ identifies the couple $(e,\overline{e}) \in \pi^*\overline{E}$ with a vertical tangent vector to $E$, and we have seen that the evolution to the vector $v(c_e^h, c_{\overline{e}}^{\overline{h}})$ is just Lie transport along $c_e^h$. If $h$ is *not* affine, Lie transport of a vertical vector along $c_e^h$ still exists and one could somehow reverse the order of thinking to use that for defining an affine action on fibres of $E$. To be specific, writing $Y_{e,\overline{e}}(a) = v(e,\overline{e})$ for the initial vertical vector and considering its Lie transport, defined as before by $Y_{e,\overline{e}}(u) = T\phi_{u,a}^h\big(Y_{e,\overline{e}}(a)\big)$, we get the following related actions on $\pi^*\overline{E}$ and $\pi^*E$:

$$\begin{aligned}
(e,\overline{e}) &\mapsto \big(c_e^h, p_{\overline{E}}\big((Y_{e,\overline{e}})_v\big)\big),\\
(e,e_1) &\mapsto \big(c_e^h, c_e^h + p_{\overline{E}}\big((Y_{e,e_1-e})_v\big)\big).
\end{aligned} \tag{9.1}$$

We will refer to this as the *affine action on $\pi^*E$ by Lie transport along horizontal curves*. Obviously, when $h$ is not affine, the image of $(e,e_1)$ under this affine action will not be $(c_e^h, c_{e_1}^h)$.

The question which arises now is whether there are natural ways also to define an affine action on $\pi^*E$ along vertical curves, i.e. curves in a fixed fibre $E_m$ of $E$. Let $c_e^v$ denote an arbitrary curve through $e$ in the fibre $E_m$ ($m = \pi(e)$). It projects onto the constant curve $c_m : u \mapsto c_m(u) = m$ in $M$. A curve in $\mathsf{V}$ which has the same projection (and actually is $\varrho$-admissible) can be taken to be $\mathsf{o}_m : u \mapsto \mathsf{o}_m(u) = (m, \mathsf{o}_m)$. $\dot{c}_e^v$ is a curve in $TE$ which projects onto $c_e^v$ and has the property $T\pi(\dot{c}_e^v) = 0$. By analogy with earlier constructions, we define a new curve $\dot{c}_e^V$ in $T^\varrho E$, determined by

$$\dot{c}_e^V := (\mathsf{o}_m, \dot{c}_e^v). \tag{9.2}$$

Obviously, by construction, we have that $\pi^1 \circ \dot{c}_e^V = c_e^v$ and $\varrho^1 \circ \dot{c}_e^V = \dot{c}_e^v$, so that $\dot{c}_e^V$ is $\varrho^1$-admissible.

Let us now address the problem of defining a transport rule in $\pi^*E$ along such a curve $c_e^v$. Remember that for the horizontal curves, we described such a transport rule by looking first at the way vertical tangent vectors can be transported. For the transport of vertical tangent vectors within a fixed fibre, the usual procedure is to take simple translation (this is sometimes called *complete parallelism*). Thus, starting from a point $(e, e_1) \in \pi^*E$, to which we want to associate first a vertical tangent vector, we think of $(e, e_1)$ as belonging to $\pi^*\tilde{E}$ and consider $v(e, e_1) = v(e, \vartheta_e(e_1)) = v(e, e_1 - e)$. Its parallel translate along a curve $c_e^v$ is $v(c_e^v, e_1 - e)$ which can be identified with $(c_e^v, e_1 - e) \in \pi^*\overline{E}$. But it makes sense to associate with this a new element of $\pi^*E$ as well, in exactly the same way as we did it for horizontal curves. We thus arrive at the following action on $\pi^*\overline{E}$ and $\pi^*E$

$$\begin{aligned} (e, \overline{e}) &\mapsto (c_e^v, \overline{e}), \\ (e, e_1) &\mapsto (c_e^v, c_e^v + e_1 - e). \end{aligned} \tag{9.3}$$

It could be described as a *vertical affine action by translation in* $\pi^*\overline{E}$.

There is, however, another way of transporting points in $\pi^*E$ along a curve of type $c_e^v$, which is in fact the most obvious one if one does insist on having a link with a transport rule of vertical tangent vectors via the vertical lift operator. It is obtained by looking at the action

$$\begin{aligned} (e, \overline{e}) &\mapsto (c_e^v, \overline{e}), \\ (e, e_1) &\mapsto (c_e^v, e_1), \end{aligned} \tag{9.4}$$

and could be termed as a *vertical affine action by translation in* $\pi^*E$.

Given an arbitrary $\varrho$-connection $h$ on the affine bundle $\pi : E \to M$, we now want to construct an induced $\varrho^1$-connection $h^1$ on the affine bundle $\pi^*\pi : \pi^*E \to E$ through the identification of suitable covariant derivative operators $\mathrm{D}$ and $\overline{\mathrm{D}}$. That is to say, we should give a meaning to things like $\mathrm{D}_{\mathcal{Z}}X$ and $\overline{\mathrm{D}}_{\mathcal{Z}}\overline{X}$, for $\mathcal{Z} \in \mathrm{Sec}(\pi^1)$, $X \in \mathrm{Sec}(\pi^*\pi)$, $\overline{X} \in \mathrm{Sec}(\pi^*\overline{\pi})$. As explained in Section 8.5, every $\mathcal{Z}$ has a unique decomposition in the form $\mathcal{Z} = \mathsf{X}^H + \overline{Y}^V$, with $\mathsf{X} \in \mathrm{Sec}(\pi^*\tau)$, $\overline{Y} \in \mathrm{Sec}(\pi^*\overline{\pi})$. These in turn are finitely generated (over $C^\infty(E)$) by *basic sections*, i.e. sections of $\tau$ and of $\overline{\pi}$, respectively. The same is true for the sections $X$ or $\overline{X}$ on which $\mathrm{D}_{\mathcal{Z}}$ and $\overline{\mathrm{D}}_{\mathcal{Z}}$ operate. This means that, for starting the construction of covariant derivatives, we must think of a defining relation for $\mathrm{D}_{\mathsf{s}^H}\sigma$, $\mathrm{D}_{\overline{\eta}^V}\sigma$, $\overline{\mathrm{D}}_{\mathsf{s}^H}\overline{\sigma}$, $\overline{\mathrm{D}}_{\overline{\eta}^V}\overline{\sigma}$, with $\mathsf{s} \in \mathrm{Sec}(\tau)$, $\sigma \in \mathrm{Sec}(\pi)$, $\overline{\eta}, \overline{\sigma} \in \mathrm{Sec}(\overline{\pi})$. The expectation is, since we look for a $\mathrm{D}$ and $\overline{\mathrm{D}}$, that $h^1$, as a kind of linearisation of $h$, will be an affine connection and so, in the particular case that the given $h$ is affine, it should essentially reproduce a copy of itself. Therefore, the first idea which presents itself is to set

$$\mathrm{D}_{\mathsf{s}^H}\sigma = [h\mathsf{s}, v\sigma]_v, \quad \overline{\mathrm{D}}_{\mathsf{s}^H}\overline{\sigma} = [h\mathsf{s}, v\overline{\sigma}]_v, \quad \mathrm{D}_{\overline{\eta}^V}\sigma = \overline{\mathrm{D}}_{\overline{\eta}^V}\overline{\sigma} = 0. \tag{9.5}$$

The first point in the proof of Proposition 8.17 did not rely on the assumption of $h$ being affine, so we know that these formulas at least are consistent with respect to the module structure over $C^\infty(M)$. We then extend the range of the operators D and $\overline{\mathrm{D}}$ in the obvious way, by the following three rules: for every $F \in C^\infty(E)$, we put

$$\begin{aligned} \mathrm{D}_{F\mathsf{s}^H}\sigma &= F\mathrm{D}_{\mathsf{s}^H}\sigma = F[h\mathsf{s}, v\sigma]_v, & (9.6)\\ \overline{\mathrm{D}}_{F\mathsf{s}^H}\overline{\sigma} &= F\overline{\mathrm{D}}_{\mathsf{s}^H}\overline{\sigma} = F[h\mathsf{s}, v\overline{\sigma}]_v, & (9.7)\\ \mathrm{D}_{F\overline{\eta}^V}\sigma &= \overline{\mathrm{D}}_{F\overline{\eta}^V}\overline{\sigma} = 0, & (9.8) \end{aligned}$$

which suffices to know what $\mathrm{D}_{\mathcal{Z}}\sigma$ and $\overline{\mathrm{D}}_{\mathcal{Z}}\overline{\sigma}$ mean for arbitrary $\mathcal{Z} \in \mathrm{Sec}(\pi^1)$, and finally we put

$$\begin{aligned} \overline{\mathrm{D}}_{\mathcal{Z}}(F\overline{\sigma}) &= F\overline{\mathrm{D}}_{\mathcal{Z}}\overline{\sigma} + \varrho^1(\mathcal{Z})(F)\overline{\sigma}, & (9.9)\\ \mathrm{D}_{\mathcal{Z}}(\sigma + F\overline{\sigma}) &= \mathrm{D}_{\mathcal{Z}}\sigma + F\overline{\mathrm{D}}_{\mathcal{Z}}\overline{\sigma} + \varrho^1(\mathcal{Z})(F)\overline{\sigma}, & (9.10) \end{aligned}$$

which suffices to give a meaning to all $\mathrm{D}_{\mathcal{Z}}X$ and $\overline{\mathrm{D}}_{\mathcal{Z}}\overline{X}$. Our operators satisfy by construction all the necessary requirements for defining an affine $\varrho^1$-connection $h^1$.

It is worthwhile to observe that for the covariant derivatives of general $X \in \mathrm{Sec}(\pi^*\pi)$ and $\overline{X} \in \mathrm{Sec}(\pi^*\overline{\pi})$, we still have an explicit formula at our disposal when $\mathcal{Z}$ is of the form $\mathsf{s}^H$, with $\mathsf{s}$ basic. This follows from the fact that $\varrho^1(\mathsf{s}^H) = h(\mathsf{s})$, so that

$$\begin{aligned} \mathrm{D}_{\mathsf{s}^H}(\sigma + F\overline{\sigma}) &= [h\mathsf{s}, v\sigma]_v + F[h\mathsf{s}, v\overline{\sigma}]_v + \varrho^1(\mathsf{s}^H)(F)\overline{\sigma},\\ &= [h\mathsf{s}, v(\sigma + F\overline{\sigma})]_v, & (9.11) \end{aligned}$$

and likewise for $\mathrm{D}_{\mathsf{s}^H}\overline{X}$.

The next point on our agenda is to understand what parallel transport means for the affine connection $(\mathrm{D}, \overline{\mathrm{D}})$, or even better, to show that it is uniquely characterised by certain features of its parallel transport. The general idea of parallel transport is clear, of course: starting from any $\varrho^1$-admissible curve $c^1$ in $T^\varrho E$, its horizontal lift is a curve $\psi^1$ in $\pi^*E$ having the same projection $\psi^1_E = c^1_E$ in $E$ and satisfying $\mathrm{D}_{c^1}\psi^1 = 0$; image points of $\psi^1$ then give parallel translation by definition. Now, $\psi^1$ is essentially a pair of curves in $E$ having the same projection in $M$, so the determination of $\psi^1$ is a matter of constructing a second curve in $E$ having the same projection in $M$ as $c^1_E$. It will be sufficient to focus on $\varrho^1$-admissible curves of the form $\dot{c}^H_e$ and $\dot{c}^V_e$, for which the corresponding projections on $E$ are curves of the form $c^h_e$ and $c^v_e$, respectively, and to consider curves $\psi^1$ which come from the restriction of sections of $\pi^*\pi$ to $c^h_e$ or $c^v_e$. To simplify matters even further, we can use basic sections $\mathsf{s} \in \mathrm{Sec}(\tau)$ to generate horizontal curves, because we know from Lemma 8.12 that the integral curves of $h(\mathsf{s}) \in \mathcal{X}(E)$ are horizontal lifts. Vertical curves, of course, can be generated as integral curves of vertical vector fields.

**Proposition 9.1.** *Let $\mathsf{s} \in \mathrm{Sec}(\tau)$, $\overline{Y} \in \mathrm{Sec}(\pi^*\overline{\pi})$ be arbitrary. Denote the integral curves of $h(\mathsf{s})$ and $v\overline{Y}$ through a point $e$ by $c^h_e$ and $c^v_e$ and consider their lifts to $\varrho^1$-admissible curves $\dot{c}^H_e$ and $\dot{c}^V_e$ in $T^\varrho E$. $(\mathrm{D}, \overline{\mathrm{D}})$ is the unique affine $\varrho^1$-connection on $\pi^*\pi$ with the properties*

*(i) Parallel transport along $\dot{c}^H_e$ is the affine action on $\pi^*E$ by Lie transport along horizontal curves.*

*(ii) Parallel transport along $\dot{c}^V_e$ is the vertical affine action by translation in $\pi^*E$.*

Proof: Recall that $\mathsf{s}^H \in \mathrm{Sec}(\pi^1)$ is defined at each $e \in E$ by $\mathsf{s}^H(e) = \big(\mathsf{s}(\pi(e)), h(\mathsf{s})(e)\big)$, so that at each point along an integral curve $c^h_e$ of $h(\mathsf{s})$, we have

$$\mathsf{s}^H(c^h_e(u)) = \big(\mathsf{s} \circ \pi \circ c^h_e(u), \dot{c}^h_e(u)\big) = \dot{c}^H_e(u).$$

Let now $X$ be an arbitrary section of $\pi^*\pi$ and put $\psi^1(u) = X(c^h_e(u))$, which defines a curve in $\pi^*E$ projecting onto $c^h_e$ in $E$. We have

$$\big(\mathrm{D}_{\dot{c}^H_e}\psi^1\big)(u) = \mathrm{D}_{\dot{c}^H_e(u)}X = \mathrm{D}_{\mathsf{s}^H(c^h_e(u))}X = \big(\mathrm{D}_{\mathsf{s}^H}X\big)\big(c^h_e(u)\big).$$

If such curve is required to govern parallel transport in $\pi^*E$, we must have $\big(\mathrm{D}_{\dot{c}^H_e}\psi^1\big)(u) = 0$, $\forall u$. This implies that $\forall \mathsf{s} \in \mathrm{Sec}(\tau)$, $\forall X \in \mathrm{Sec}(\pi^*\pi)$, $\mathrm{D}_{\mathsf{s}^H}X$ should be zero along integral curves of $h(\mathsf{s}) \in \mathcal{X}(E)$. By the remark about the explicit formula (9.11) for $\mathrm{D}_{\mathsf{s}^H}X$ and with $v(X)(c^h_e(u)) = v(\psi^1(u))$, which defines a vertical vector field along $c^h_e$, this requirement is further equivalent to $\mathcal{L}_{h(\mathsf{s})}v(\psi^1) = 0$, which is precisely the characterisation of Lie transport. The same arguments apply to $\overline{\mathrm{D}}$ and show that our $(\mathrm{D},\overline{\mathrm{D}})$ has the property (i).

With $\overline{Y} \in \mathrm{Sec}(\pi^*\overline{\pi})$, $\overline{Y}^V \in \mathrm{Sec}(\pi^1)$ is such that $\varrho^1(\overline{Y}^V)$ is a vertical vector field on $E$. Hence, its integral curves are curves of the form $c_e^v$ in a fixed fibre $E_{\pi(e)}$ and we have from (9.2):

$$\overline{Y}^V\big(c_e^v(u)\big) = \big(\mathsf{o}_{\pi(e)}, \dot{c}_e^v(u)\big) = \dot{c}_e^V(u).$$

Let $X$ again be an arbitrary section of $\pi^*\pi$ and put this time $\psi^1(u) = X(c_e^v(u))$, which defines a curve in $E \times_M E$ projecting onto $c_e^v$ for its first component. We wish to show that if $\psi^1(u)$ rules parallel transport along $\dot{c}_e^V$, it is necessarily a curve which is constant in its second component. We have

$$\big(\mathrm{D}_{\dot{c}_e^V}\psi^1\big)(u) = \mathrm{D}_{\dot{c}_e^V(u)}X = \mathrm{D}_{\overline{Y}^V(c_e^v(u))}X = \big(\mathrm{D}_{\overline{Y}^V}X\big)\big(c_e^v(u)\big).$$

There is no explicit formula available for $\mathrm{D}_{\overline{Y}^V}X$. However, $X$ is locally of the form $X = \sigma + F_i\overline{\sigma}_i$, with $\sigma$, $\overline{\sigma}_i$ basic sections and $F_i \in C^\infty(E)$. It then follows that $\mathrm{D}_{\overline{Y}^V}X = \varrho^1(\overline{Y}^V)(F_i)\overline{\sigma}_i$ and the requirement $\mathrm{D}_{\overline{Y}^V}X(c_e^v(u)) = 0$ implies that the $F_i$ must be first integrals of $\varrho^1(\overline{Y}^V)$. In turn this means that the value $X(c_e^v(u))$ is constant. This way we see that the affine connection $(\mathrm{D},\overline{\mathrm{D}})$ also has property (ii).

That properties (i) and (ii) uniquely fix the connection is easy to see, because the above arguments show that they impose in particular that $\mathrm{D}_{\mathsf{s}^H}\sigma = [h\mathsf{s}, v\sigma]_v$ and $\mathrm{D}_{\overline{\eta}^V}\sigma = 0$ (and similarly for $\overline{\mathrm{D}}$), for basic $\mathsf{s},\sigma$ and $\overline{\eta}$. And these are exactly the defining relations (9.5) from which our couple $(\mathrm{D},\overline{\mathrm{D}})$ was constructed. □

We have seen earlier on that there is a second interesting transport rule along vertical curves and would like to discover now what modifications to the affine connection must be made to have this other rule as vertical parallel transport. We are referring here to the action (9.3) for which the curve starting at some $(e, e_1) \in E_m \times E_m$ is of the form

$$u \overset{\psi^1}{\mapsto} \big(c_e^v(u), e_1 + c_e^v(u) - e\big) = \big(c_e^v(u), c_e^v(u) + e_1 - e\big).$$

Now, it is easy to identify a section of $\pi^*\pi$ which along $c_e^v$ coincides with this curve. Indeed, choosing a basic section $\overline{\sigma} \in \mathrm{Sec}(\overline{\pi})$ which at $m = \pi(e)$ takes the value $e_1 - e$, we are simply looking at the restriction to $c_e^v$ of $\mathcal{I} + \overline{\sigma}$, where $\mathcal{I}$ here denotes the identity map on $E$ (but regarded as a section of $\pi^*\tilde{\pi}$).

Let $(\hat{\mathrm{D}}, \overline{\hat{\mathrm{D}}})$ denote the affine connection we are looking for now and which clearly will coincide with $(\mathrm{D}, \overline{\mathrm{D}})$ for its "horizontal action". If, as before, $\overline{Y} \in \mathrm{Sec}(\pi^*\overline{\pi})$ generates the vertical vector field $v\overline{Y}$ whose integral curves are the $c_e^v$, the above $\psi^1$ will produce parallel transport, provided we have

$$\big(\hat{\mathrm{D}}_{\dot{c}_e^V}\psi^1\big)(u) = \hat{\mathrm{D}}_{\overline{Y}^V}(\mathcal{I} + \overline{\sigma})(c_e^v(u)) = 0.$$

Since this must hold for each $\overline{Y}^V$ and, for every fixed $\overline{Y}^V$ also for all $\overline{\sigma}$, this is equivalent to requiring that $\hat{\mathrm{D}}_{\overline{Y}^V}\mathcal{I} = 0$ and $\overline{\hat{\mathrm{D}}}_{\overline{Y}^V}\overline{\sigma} = 0$, $\forall \overline{Y}, \overline{\sigma}$. In fact, in view of the linearity in $\overline{Y}$, we actually obtain the conditions

$$\hat{\mathrm{D}}_{\overline{\eta}^V}\mathcal{I} = 0 \qquad \text{and} \qquad \overline{\hat{\mathrm{D}}}_{\overline{\eta}^V}\overline{\sigma} = 0, \tag{9.12}$$

for all basic $\overline{\sigma}$ and $\overline{\eta}$. It is interesting to characterise this completely by properties on basic sections, because the extension to a full affine connection on $\pi^*\pi$ then follows automatically. If $\sigma$ is an arbitrary basic section of $\pi^*\pi$, it can be decomposed (see(6.4)) in the form $\sigma = \mathcal{I} + \vartheta(\sigma)$, whereby $\vartheta(\sigma)(e) = \big(e, \sigma(\pi(e)) - e\big)$. Clearly, $\sigma(\pi(e)) - e$, as an element of $\overline{E}$, has components which are linear functions of the fibre coordinates of $e$, in such a way that when acted upon by the vector field $\varrho^1(\overline{\eta}^V)$, we will obtain $-\overline{\eta}$. It follows that $\overline{\hat{\mathrm{D}}}_{\overline{\eta}^V}\vartheta(\sigma) = -\overline{\eta}$ and therefore that

$$\hat{\mathrm{D}}_{\overline{\eta}^V}\mathcal{I} = 0 \qquad \Longleftrightarrow \qquad \hat{\mathrm{D}}_{\overline{\eta}^V}\sigma = -\overline{\eta},\ \forall \sigma \in \mathrm{Sec}(\pi). \tag{9.13}$$

This way, we have detected an alternative way for defining an affine $\varrho^1$-connection $(\hat{\mathrm{D}}, \overline{\hat{\mathrm{D}}})$ on $\pi^*\pi$. Compared to (9.5), its defining relations are

$$\hat{\mathrm{D}}_{\mathsf{s}^H}\sigma = [h\mathsf{s}, v\sigma]_v, \quad \overline{\hat{\mathrm{D}}}_{\mathsf{s}^H}\overline{\sigma} = [h\mathsf{s}, v\overline{\sigma}]_v, \quad \hat{\mathrm{D}}_{\overline{\eta}^V}\sigma = -\overline{\eta}, \quad \overline{\hat{\mathrm{D}}}_{\overline{\eta}^V}\overline{\sigma} = 0. \tag{9.14}$$

We can further immediately draw the following conclusion about its characterisation

**Proposition 9.2.** *With the same premises as in Proposition 9.1,* $(\hat{\mathrm{D}}, \overline{\hat{\mathrm{D}}})$ *is the unique affine* $\varrho^1$*-connection on* $\pi^*\pi$ *with the properties*

*(i) Parallel transport along* $\dot{c}_e^H$ *is the affine action on* $\pi^*E$ *by Lie transport along horizontal curves.*

*(ii) Parallel transport along* $\dot{c}_e^V$ *is the vertical affine action by translation in* $\pi^*\overline{E}$*.*

We will refer to the connections $(\mathrm{D}, \overline{\mathrm{D}})$ and $(\hat{\mathrm{D}}, \overline{\hat{\mathrm{D}}})$, as well as their extensions $\tilde{\mathrm{D}}$ and $\tilde{\hat{\mathrm{D}}}$ as *Berwald-type connections*. For completeness, we list their defining relations here in coordinates (with $\mathcal{H}_0$, $\mathcal{H}_\alpha$ and $\mathcal{V}_\alpha$ as defined in Section 8.2).

$$\mathrm{D}_{\mathcal{H}_a} e_0 = \left(\Gamma^\gamma_a - y^\beta \frac{\partial \Gamma^\gamma_a}{\partial y^\beta}\right) \overline{e}_\gamma, \qquad \mathrm{D}_{\mathcal{V}_\alpha} e_0 = 0,$$

$$\overline{\mathrm{D}}_{\mathcal{H}_a} \overline{e}_\beta = \frac{\partial \Gamma^\gamma_a}{\partial y^\beta} \overline{e}_\gamma, \qquad \overline{\mathrm{D}}_{\mathcal{V}_\alpha} \overline{e}_\beta = 0$$

and

$$\hat{\mathrm{D}}_{\mathcal{H}_a} e_0 = \left(\Gamma^\gamma_a - y^\beta \frac{\partial \Gamma^\gamma_a}{\partial y^\beta}\right) \overline{e}_\gamma, \qquad \hat{\mathrm{D}}_{\mathcal{V}_\alpha} e_0 = -\overline{e}_\alpha,$$

$$\overline{\hat{\mathrm{D}}}_{\mathcal{H}_a} \overline{e}_\beta = \frac{\partial \Gamma^\gamma_a}{\partial y^\beta} \overline{e}_\gamma, \qquad \overline{\hat{\mathrm{D}}}_{\mathcal{V}_\alpha} \overline{e}_\beta = 0.$$

## 9.2 The case of affine Lie algebroids and the canonical connection associated to a pseudo-SODE

It is now time to relate the quite general results of the preceding section to the earlier chapters. Recall that our interest in affine bundles comes in the first place from the geometrical study of time-dependent second-order equations and the analysis of Berwald-type connections in that context. Secondly, we have explored a time-dependent generalisation of Lagrangian systems on Lie algebroids and thus arrived at the introduction and study of affine Lie algebroids. Notice that Lagrangian systems on algebroids are particular cases of pseudo-SODEs, but the concept of a pseudo-SODE in itself, strictly speaking, does not require the full structure of a Lie algebroid.

Pseudo-SODEs on the affine bundle $E$ are essentially vector fields with the property that all the integral curves are $\rho$-admissible. In saying that, we are in fact assuming that the anchor map has $E$ in its domain. In this section, therefore, the starting point is that we have an affine bundle map $\rho : E \to TM$ at our disposal. In what follows $\tilde{\rho}$ plays the role of the anchor map $\varrho$ we had before. This means in particular that the vector bundle $\tau : \mathsf{V} \to M$ from now on is taken to be the bundle $\tilde{\pi} : \tilde{E} \to M$. We have seen in Section 7.8 that pseudo-SODEs can be regarded as sections of the prolonged bundle $\pi^1 : T^{\tilde{\rho}}E \to E$ (rather than as vector fields on $E$). Further, if $\pi$ is assumed to be an affine Lie algebroid, Proposition 8.5 shows a way of pinning down a $\tilde{\rho}$-connection on $\pi$. The particular case of a Lagrangian system on the affine Lie algebroid $\pi$ can be find in Section 7.9. In that case, the functions $f^\alpha$ which determine the connection coefficients are given by (7.20).

We come back to the construction of Berwald-type connections associated to arbitrary $\tilde{\rho}$-connections, in the case of an affine Lie algebroid. So assume we have a $\tilde{\rho}$-connection on the affine Lie algebroid $\pi$ (not necessarily of pseudo-Sode type). It is then appropriate to work with the adapted basis $\{\mathcal{H}_0, \mathcal{H}_\alpha, \mathcal{V}_\alpha\}$ for $\mathrm{Sec}(\pi^1)$, rather than the "coordinate basis" $\{\mathcal{X}_0, \mathcal{X}_a, \mathcal{V}_\alpha\}$. In (8.6) we have list the most useful bracket relations. It will further be appropriate to write now ${}_H$ for the projection $T^{\tilde{\rho}}E \to \pi^*\tilde{E}$ and likewise define the map ${}_V : T^{\tilde{\rho}}E \to \pi^*\overline{E} \subset \pi^*\tilde{E}$ by: $\mathcal{Z}_V = \big(\tilde{\rho}^1(P_V\mathcal{Z})\big)_v$. The reason is that this will bring us in line with notations used in the second chapter to which the next proposition strongly relates. Combining the horizontal and vertical lift operations with the direct sum decomposition (6.5) of $\mathrm{Sec}(\pi^*\tilde{\pi})$, it is more convenient now to think of the following threefold decomposition of $\mathrm{Sec}(\pi^1)$:

$$\mathrm{Sec}(\pi^1) = \langle \mathcal{I}^H \rangle \oplus \mathrm{Sec}(\pi^*\overline{\pi})^H \oplus \mathrm{Sec}(\pi^*\overline{\pi})^V. \tag{9.15}$$

Note that, in the particular case of a pseudo-Sode connection, we have $\mathcal{I}^H = \Gamma$.

We know that any $\tilde{\rho}$-connection generates Berwald-type connections. The strong point of the next result, however, is that if we assume that $\pi$ is an affine Lie algebroid, there is a direct defining formula for the two Berwald-type connections discussed in the preceding section.

**Proposition 9.3.** *If the affine bundle $\pi$ carries an affine Lie algebroid structure, the Berwald-type connections $\tilde{\mathrm{D}}$ and $\hat{\tilde{\mathrm{D}}}$ are determined by the following direct formulae:*

$$\tilde{\mathrm{D}}_{\mathcal{Z}}\tilde{X} = [P_H\mathcal{Z}, \tilde{X}^V]_V + [P_V\mathcal{Z}, \tilde{X}^H]_H + \tilde{\rho}^1(P_H\mathcal{Z})\big(\langle \tilde{X}, e^0\rangle\big)\mathcal{I}, \tag{9.16}$$

$$\hat{\tilde{\mathrm{D}}}_{\mathcal{Z}}\tilde{X} = [P_H\mathcal{Z}, \tilde{X}^V]_V + [P_V\mathcal{Z}, \overline{X}^H]_H + \tilde{\rho}^1\mathcal{Z}\big(\langle \tilde{X}, e^0\rangle\big)\mathcal{I}, \tag{9.17}$$

*with $\overline{X} := \tilde{X} - \langle \tilde{X}, e^0\rangle\mathcal{I}$.*

Proof: Using the properties ${}_V \circ P_H = 0$, ${}_H \circ P_V = 0$, $h \circ {}_H = \tilde{\rho}^1 \circ P_H$, it is easy to verify that the above expressions satisfy the appropriate rules when the arguments are multiplied by a function on $E$. Hence, both operators define a linear $\tilde\rho^1$-connection on the vector bundle $\pi^*\tilde\pi$. Next, we verify that this connection comes from an affine $\tilde\rho^1$ connection on $\pi^*\pi$. For that, according to Proposition 8.11, it is necessary and sufficient that $e^0$ (here regarded as basic section of $\pi^*\tilde\pi$) is parallel. We have

$$(\tilde{\mathrm{D}}_{\mathcal{Z}} e^0)(\tilde X) = \tilde\rho^1 \mathcal{Z}(\langle \tilde X, e^0\rangle) - \langle \tilde{\mathrm{D}}_{\mathcal{Z}} \tilde X, e^0\rangle, \tag{9.18}$$

and similarly for $\tilde{\hat{\mathrm{D}}}$. In the case of $\tilde{\mathrm{D}}$, we have

$$\langle \tilde{\mathrm{D}}_{\mathcal{Z}} \tilde X, e^0\rangle = \langle [P_V\mathcal{Z}, \tilde X^H]_H, e^0\rangle + \tilde\rho^1(P_H\mathcal{Z})\big(\langle \tilde X, e^0\rangle\big).$$

Making use of the bracket relations (8.6), it is straightforward to verify that the first term on the right is equal to $\tilde\rho^1(P_V\mathcal{Z})\big(\langle \tilde X, e^0\rangle\big)$, so that the sum of both terms indeed makes the right-hand side of (9.18) vanish. The computation for $\tilde{\hat{\mathrm{D}}}$ is similar.

It remains now to check that the restrictions to $\mathrm{Sec}(\pi^*\pi)$ and $\mathrm{Sec}(\pi^*\overline\pi)$ of (9.16) and (9.17) verify, respectively, the defining relations (9.5) and (9.14) for $(\mathrm{D}, \overline{\mathrm{D}})$ and $(\hat{\mathrm{D}}, \overline{\hat{\mathrm{D}}})$. If we take $\mathcal{Z} = \tilde\sigma^H$ and $\tilde X = \eta$, for basic $\tilde\sigma \in \mathrm{Sec}(\tilde\pi)$ and $\eta \in \mathrm{Sec}(\pi)$, then we know from Proposition 8.17 that the bracket $[h\tilde\sigma, v\eta]$ is vertical in $TE$. As a consequence $(0, [h\tilde\sigma, v\eta])$ is vertical in $T^{\tilde\rho}E$. But this is precisely $[\tilde\sigma^H, \eta^V]$, because the bracket of the two projectable sections $\tilde\sigma^H$ and $\eta^V$ is by construction (see Section 7.3) the section $\Big([\tilde\sigma, 0], [\tilde\rho^1\tilde\sigma^H, \tilde\rho^1\eta^V]\Big)$ of $\pi^1$. Therefore, $\mathrm{D}_{\tilde\sigma^H}\eta = [\tilde\sigma^H, \eta^V]_V = \big(\tilde\rho^1(P_V[\tilde\sigma^H, \tilde\eta^V])\big)_v = \big(\tilde\rho^1[\tilde\sigma^H, \tilde\eta^V]\big)_v = [h\tilde\sigma, v\tilde\eta]_v$, where the Lie algebra homomorphism provided by the anchor map $\tilde\rho^1$ has been used. Similar arguments apply for the other operators $\overline{\mathrm{D}}$, $\hat{\mathrm{D}}$ and $\overline{\hat{\mathrm{D}}}$ when $\mathcal{Z}$ is horizontal. In remains to look at the case $\mathcal{Z} = \overline\sigma^V$ ($\overline\sigma \in \mathrm{Sec}(\overline\pi)$). Since $[\overline\sigma^V, \overline\eta^H]$ is vertical, it follows that $\overline{\mathrm{D}}_{\overline\sigma^V}\overline\eta = \overline{\hat{\mathrm{D}}}_{\overline\sigma^V}\overline\eta = 0$. For $\tilde X = \eta$, since then $\langle \eta, e^0\rangle = 1$, we find for the first connection $\mathrm{D}_{\overline\sigma^V}\eta = 0$. For the second connection, it suffices to check (see (9.12)) that $\hat{\mathrm{D}}_{\overline\sigma^V}\mathcal{I} = 0$, and this is trivial. □

## 9.3 Conclusions

In the last two chapters, two main objectives have been attained: we have unravelled the mechanism by which a generalised connection over an anchored bundle leads to a linearised connection over an appropriate prolonged anchored bundle; we have at the same time focussed on the special features of connections on an affine bundle, in general, and on an affine Lie algebroid in particular. The latter subject ties up with the first issue, as a generalisation of the study of Berwald-type connections in Chapter 2, where we dealt, so to speak, with the prototype of an affine Lie algebroid, namely the first-jet extension of a bundle fibred over $\mathbb{R}$, this being the geometrical arena for time-dependent mechanics.

What are such Berwald-type connections good for? We have already mentioned in Chapter 3 that the covariant derivative operators associated to (classical) Berwald-type connections have proved to be very useful tools in a number of applications concerning qualitative features of SODEs, such as, for example, in the characterisation of linearisability [52, 23] and of separability [55, 12] of SODEs; in the inverse problem of Lagrangian mechanics [27]; in the study of Jacobi fields and Raychaudury's equation [39]. There is little doubt that there are similar applications ahead for the qualitative study of pseudo-SODEs on Lie algebroids. However, to reward their endurance in getting this far, we allow the readers a small break. We therefore postpone these investigations for a next occasion.

# *Bibliography*


[1] M. Abate and G. Patrizio, *Finsler Metrics - A Global Approach* (Springer-Verlag, Berlin) (1994).

[2] M. Anastasiei, Finsler connections in generalized Lagrange spaces, *Balkan J. of Geom. and its Appl.* **1** (1996) 1–9.

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

[4] P.L. Antonelli, R.S. Ingarden and M. Matsumoto, *The theory of sprays and Finsler spaces with applications in physics and biology*, Fundamental Theories of Physics Vol. 58, (Kluwer Academic Publishers, Dordrecht) (1993).

[5] D. Bao and S.S. Chern, On a notable connection in Finsler geometry, *Houston J. Math.* **19** (1993) 138–180.

[6] D. Bao, S.-S. Chern and Z. Shen, *An Introduction to Riemann-Finsler Geometry*, Graduate Texts in Mathematics Vol. 200 ( Springer-Verlag, New York) (2000).

[7] A. Bejancu, *Finsler geometry and applications*, Ellis Horwood series in Math. and its Appl., (Ellis Horwood Limited, New York) (1990).

[8] L. Berwald, Untersuchung der Krümmung allgemeiner metrischen Räume auf Grund des in ihnen herrschenden Parallelismus, *Math. Zeit.* **25** (1925) 40–73.

[9] G. Byrnes, A complete set of Bianchi identities for tensor fields along the tangent bundle projection, *J. Phys. A: Math. Gen.* **27** (1994) 6617–6632.

[10] A. Cannas da Silva and A. Weinstein, *Geometric models for noncommutative algebras*, Berkeley Mathematics Lecture Notes Vol. 10 (Am. Math. Society, Providence) (1999).

[11] F. Cantrijn and B. Langerock, Generalised connections over a bundle map, *Diff. Geom. Appl.* **18** (2003) 295–317.

[12] F. Cantrijn, W. Sarlet, A. Vandecasteele and E. Martínez, Complete separability of time-dependent second-order ordinary differential equations, *Acta Appl. Math.* **42** (1996) 309–334.

[13] J.F. Cariñena, Lie groupoids and algebroids in classical and quantum mechanics, In: *Symmetries in quantum mechanics and quantum optics* (University of Burgos, Spain) (1999) 67–81.

[14] J.F. Cariñena and E. Martínez, Lie algebroid generalization of geometric mechanics, In: *Lie algebroids and related topics in differential geometry*, J. Kubarski, P. Urbański and R. Wolak, eds., Banach Center Publications Vol. 54 (Pol. Acad. Sci., Warsaw) (2001) 201–215.

[15] E. Cartan, *Les espaces de Finsler*, Actualités Vol. 79 (Paris) (1934), (2nd edit, 1971).

[16] J. Clemente-Gallardo, Applications of Lie algebroids in mechanics and control theory, In: *Nonlinear control in the new millenium*, F. Lamnabhi-Lagarrigur, W. Respondek and A. Isidori, eds., (Springer Verlag, London) (2000) 299–313.

[17] T. Courant, Dirac Manifolds, *Trans. Amer. Math. Soc.* **319** (1990), 631–661.

[18] T. Courant, Tangent Lie algebroids, *J. Phys. A: Math. Gen.* **27** (1994), 4527–4536.

[19] M. Crampin, Tangent bundle geometry for Lagrangian dynamics, *J. Phys. A* **16** (1983) 3755–3772.

[20] M. Crampin, *Jet bundle techniques in analytical mechanics*, Quaderni del Consiglio Nazionale delle Ricerche Vol. **47** (Firenze) (1995).

[21] M. Crampin, Connections of Berwald type, *Publ. Math. (Debrecen)* **57** (2000) 455–473.

[22] M. Crampin, The second variation formula in Lagrange and Finsler geometry, *Houston J. Math.* **26** (2000) 255–276.

[23] M. Crampin, E. Martínez and W. Sarlet, Linear connections for systems of second-order ordinary differential equations, *Ann. Inst. H. Poincaré, Phys. Théor.* **65** (1996) 223–249.

[24] M. Crampin and F.A.E. Pirani, *Applicable Differential Geometry*, London Math. Soc. Lect. Notes Series Vol. 59 (Camebridge Univ. Press, Camebridge) (1986).

[25] M. Crampin, G.E. Prince, W. Sarlet and G. Thompson, The inverse problem of the calculus of variations: separable systems, *Acta Appl. Math.* **57** (1999) 239–254.

[26] M. Crampin, G.E. Prince and G. Thompson, A geometrical version of the Helmholtz conditions in time-dependent Lagrangian dynamics, *J. Phys. A* **17** (1984) 1437–1447.

[27] M. Crampin, W. Sarlet, E. Martínez, G.B. Byrnes and G.E. Prince, Towards a geometrical understanding of the inverse problem of the calculus of variations, *Inverse problems* **10** (1994) 256–260.

[28] R.L. Fernandes, Connections in Poisson geometry I: holonomy and invariants, *J. Diff. Geom.* **54** (2000) 303–365.

[29] R.L. Fernandes, Lie algebroids, holonomy and characteristic classes, *Adv. Math.* **170** (2002) 119–170.

[30] P. Finsler, *Über Kurven und Flächen in allgemeinen Räumen, 1918* (Verlag Birkhuser, Basel) (1951).

[31] R. Fritsch, Georg Pick und Ludwig Berwald, Zwei Mathematiker an der Deutschen Universität Prag, *Schr. Sudet.dtsch. Akad. Wiss. Künste Forsch.beitr. Nat.wiss. Kl.* **22** (2001) 9–16, see also: www.mathematik.uni-muenchen.de/∼fritsch/biographien.html.

[32] A. Frölicher and A. Nijenhuis, Theory of vector-valued differential forms, *Proc. Ned. Acad. Wetensch. Sér. A* **59** (1956) 338–359.

[33] J. Grabowski and P. Urbañski, Tangent and cotangent lifts and graded Lie algebras associated with Lie algebroids, *Ann. Global An. Geom.* **15** (1997) 447–486.

[34] J. Grabowski and P. Urbañski, Lie algebroids and Poisson-Nijenhuis structures, *Rep. Math. Phys.* **40** (1997) 195–208.

[35] J. Grifone, Structure presque tangente et connexions I, *Ann. Inst. Fourier, Grenoble* **22** (no. 1) (1972) 287–334.

[36] J. Grifone, Structure presque tangente et connexions II, *Ann. Inst. Fourier, Grenoble* **22** (no. 3) (1972) 291–338.

[37] P.J. Higgins and K. Mackenzie, Algebraic constructions in the category of Lie algebroids, *J. of Algebra* **129** (1990) 194–230.

[38] N. Jacobson, *Lie algebras* (Dover Publ., Inc., New York) (1962).

[39] M. Jerie and G. Prince, Jacobi fields and linear connections for arbitrary second order ODE's , *J. Geom. Phys.* **43** (2002) 351–370.

[40] A. Kawaguchi, On the theory of non-linear connections, *Convegno Internazionale di Geometria Differenziale* (Italy) (1954) 27–32.

[41] J. Klein, Espaces variationnels et mécanique, *Ann. Inst. Fourier (Grenoble)* **12** (1962) 1–124.

[42] B. Langerock, Nonholonomic mechanics and connections over a bundle map, *J. Phys. A: Math. Gen.* **34** (2001) L609–L615.

[43] B. Langerock, A connection theoretic approach to sub-Riemannian geometry, *J. Geom. Phys.* **46** (2003) 203–230.

[44] B. Langerock, Geometric aspects of the Maximum Principle and lifts over a bundle map, *Acta Appl. Math.* **77** (2003) 71–104.

[45] P. Libermann, Lie algebroids and Mechanics, *Arch. Math. (Brno)* **32** (1996) 147–162.

[46] K. Mackenzie, *Lie groupoids and Lie algebroids in differential geometry*, London Math. Soc. Lect. Note Series Vol. 124 (Cambridge Univ. Press, Camebridge) (1987).

[47] L. Mangiarotti and G. Sardanashvily, *Connections in classical and quantum field theory* (World Scientific Publishing, River Edge, NJ) (2000).

[48] E. Martínez, Lagrangian Mechanics on Lie algebroids, *Acta. Appl. Math.* **67**, 295–320 (2001).

[49] E. Martínez, Geometric Formulation of Mechanics on Lie algebroids, In: *Procs. of the VIII Fall Workshop on Geometry and Physics, Medina del Campo 1999* (Publicaciones de las RSME) (2001).

[50] E. Martínez, Hamiltonian Mechanics on Lie algebroids, preprint (2000).

[51] E. Martínez, Connection theory on Lie algebroids, preprint (2002).

[52] E. Martínez and J.F. Cariñena, Geometric characterization of linearizable second-order differential equations, *Math. Proc. Camb. Phil. Soc.* **119** (1996) 373–381.

[53] E. Martínez, J.F. Cariñena and W. Sarlet, Derivations of differential forms along the tangent bundle projection, *Diff. Geometry and its Applications* **2** (1992) 17–43.

[54] E. Martínez, J.F. Cariñena and W. Sarlet, Derivations of differential forms along the tangent bundle projection II, *Diff. Geometry and its Applications* **3** (1993) 1–29.

[55] E. Martínez, J.F. Cariñena and W. Sarlet, Geometric characterization of separable second-order equations, *Math. Proc. Camb. Phil. Soc.* **113** (1993) 205–224.

[56] E. Martínez, T. Mestdag and W. Sarlet, Lie algebroid structures and Lagrangian systems on affine bundles, *J. Geom. Phys.* **44** (2002) 70–95.

[57] E. Massa and E. Pagani, Jet bundle geometry, dynamical connections, and the inverse problem of Lagrangian mechanics, *Ann. Inst. H. Poincaré Phys. Théor.* **61** (1994) 17–62.

[58] M. Matsumoto, *History of Finsler Geometry*, Lecture notes from a talk held at the Department of Geometry, University of Debrecen (1979).

[59] M. Matsumoto, *Foundations of Finsler Geometry and special Finsler spaces* (Kaisheisha Press, Otsu) (1986).

[60] T. Mestdag, Generalised connections on affine Lie algebroids, *Rep. on Math. Phys.* **51** (2003) 297–305.

[61] T. Mestdag and W. Sarlet, The Berwald connection for time-dependent second-order differential equations and its applications in theoretical mechanics, In: *Proceedings of the 5th National Congress on Theoretical and Applied Mechanics, Louvain-la-Neuve (Belgium), May 23-24, 2000*, 71–74.

[62] T. Mestdag and W. Sarlet, The Berwald-type connection associated to time-dependent second-order differential equations, *Houston J. Math.* **27** (2001) 763–797.

[63] T. Mestdag and W. Sarlet, The Berwald-type linearisation of generalised connections, J. Phys. A: Math. Gen. (2003) to appear.

[64] T. Mestdag, W. Sarlet and E. Martínez, Note on generalized connections and affine bundles, *J. Phys. A: Math. Gen.* **35** (2002) 9843–9856.

[65] T. Mestdag, J. Szilasi and V. Tóth, On the geometry of generalized metrics, *Publ. Math. (Debrecen)* **62** (2003) 511–545.

[66] T. Mestdag and V. Tóth, On the geometry of Randers manifolds, *Rep. Math. Phys.* **50** (2002) 167–193.

[67] R. Miron and M. Anastasiei, *The Geometry of Lagrange Spaces: Theory and Applications*, Fundamental Theories of Physics Vol. 59 (Kluwer Academic Publishers, Dordrecht) (1994).

[68] R. Miron, T. Aikou, M. Hashiguchi, On minimality of axiomatic systems of remarkable Finsler connections, *Rep. Fac. Sci. Kagoshima Univ. (Math. Phys. Chem. )* **26** (1993) 41–51.

[69] A. Nijenhuis, Vector form brackets in Lie algebroids, *Arch. Math. (Brno)* **32** (1996) 317–323.

[70] T. Okada, Minkowskian product of Finsler spaces and Berwald connections, *J. Math. Kyoto Univ.* **22** (1982) 323-332.

[71] M. Pinl, In memory of Ludwig Berwald, *Scripta Math.* **27** (1964) 193–203.

[72] M. Popescu and P. Popescu, Geometric objects defined by almost Lie structures, In: *Lie Algebroids and related topics in differential geometry*, J. Kubarski, P. Urbański and R. Wolak, eds., Banach Center Publications Vol. 54 (Pol. Acad. Sci., Warsaw) (2001) 217–233.

[73] B. Riemann, *Über die Hypothesen, welche der Geometrie zu Grunde liegen*, edited by H. Weyl, Springer (1921).

[74] H. Rund, *The Differential Geometry of Finsler Spaces*, Grundlehren der Mathematischen Wissenschaffen Vol. 101, (Springer-Verlag, Berlin) (1959).

[75] W. Sarlet, M. Crampin and E. Martínez, The integrability conditions in the inverse problem of the calculus of variations for second-order ordinary differential equations, *Acta Appl. Math.* **54** (1998) 233–273.

[76] W. Sarlet and T. Mestdag, Aspects of time-dependent second-order differential equations: Berwald-type connections, In: *L. Kozma, P.*

*T. Nagy and L. Tamássy (eds.), Steps in Differential Geometry, Proceedings of the Colloquium on Differential Geometry, July 25-30, 2000, Debrecen, Hungary* (University of Debrecen, Debrecen) (2001) 283–293.

[77] W. Sarlet, T. Mestdag and E. Martínez, Lie algebroid structures on a class of affine bundles, *J. Math. Phys.* **43** (2002) 5654–5674.

[78] W. Sarlet, T. Mestdag and E. Martínez, Lagrangian equations on affine Lie algebroids, In: *O. Kowalski, D. Krupka, J. Slovak (eds.), Differential Geometry and its Applications, Proc. Conf., Opava (Czech Republic), August 27-31, 2001* (Silesian University, Opava) (2002) 461-472.

[79] W. Sarlet, G. Thompson and G.E. Prince, The inverse problem of the calculus of variations: the use of geometrical calculus in Douglas's analysis, *Trans. Amer. Math. Soc.* **354** (2002) 2897–2919.

[80] W. Sarlet, A. Vandecasteele, F. Cantrijn and E. Martínez, Derivations of forms along a map: the framework for time-dependent second-order equations, *Diff. Geometry and its Applications* **5** (1995) 171–203.

[81] D. Saunders, *The geometry of jet bundles*, London Math. Soc. Lect. Note Series Vol. 142 (Cambridge University Press, Cambridge) (1989).

[82] M. Spivak, *A comprehensive introduction to differential geometry.*, Vol. II., Brandeis Univ., Waltham, Mass. (1970).

[83] J.L. Synge, A generalisation of the Riemannian line-element, *Trans. Amer. Math. Soc.* **27** (1925) 61–67.

[84] J. Szilasi, Notable Finsler connections on a Finsler manifold, *Lect. Matematicas* **19** (1998) 7–34.

[85] J. Szilasi, *A setting for Spray and Finsler Geometry*, Handbook of Finsler Geometry (ed. P.L. Antonelli), to be published by Kluwer Academic Publishers, Dordrecht.

[86] J. Szilasi and Z. Kovács, Pseudoconnections and Finsler-type connections, In: *Topics in Differential Geometry*, Colloquia Math. Soc. János Bolyai Vol. 46 (North-Holland, Amsterdam) (1984) 1165–1184.

[87] J.H. Taylor, A generalization of Levi-Civita's parallelism and the Frenet formulas, *Trans. Amer. Math. Soc.* **27** (1925) 246–264.

[88] I. Vaisman, *Lectures on the Geometry of Poisson manifolds*, Progress in Math. Vol. 118 (Birkhäuser, Basel) (1994).

[89] J. Vilms, Curvature of nonlinear connections, *Proc. Amer. Math. Soc.* **19** (1968) 1125–1129.

[90] J. Vilms, On the existence of Berwald connections, *Tensor, N.S.* **26** (1972) 118–120.

[91] A. Weinstein, Lagrangian mechanics and groupoids, In: *Mechanics day (Waterloo, ON, 1992)*, Fields Institute Communications Vol. 7 (Am. Math. Society, Providence) (1996) 207–231.

# *Index*

# *Samenvatting*

In algemene bewoordingen kan men de constructie van de Berwaldconnectie in Finslermeetkunde als volgt omschrijven: de energiefunctie van een Finslervariëteit $M$ geeft aanleiding tot een stelsel autonome tweede-orde differentiaalvergelijkingen op $M$. Dit stelsel bepaalt een niet-lineaire connectie op de raakbundel $TM \to M$, die op haar beurt 'gelineariseerd' kan worden tot een lineaire connectie, met name de Berwaldconnectie.

In de literatuur is het algemeen bekend dat een aantal van de hoger vermelde stappen onafhankelijk van elkaar betekenis hebben. In het EERSTE HOOFDSTUK van deze verhandeling herhalen we bijvoorbeeld hoe met elk autonoom tweede-orde systeem een niet-lineaire connectie geassocieerd kan worden, ongeacht of dit stelsel afkomstig is van een energiefunctie. Bovendien is het ook mogelijk om, startend van een willekeurige niet-lineaire connectie op $TM$, een aanverwante lineaire connectie te construeren. De connectie die via dit 'linearisatieproces' ontstaat zullen we in het vervolg aanduiden als de lineaire connectie 'van het Berwaldtype'.

De uitbreiding van deze constructie naar de context van tijdsafhankelijke tweede-orde systemen

$$\begin{aligned} \dot{x}^i &= v^i, \\ \dot{v}^i &= f^i(t,x,v), \end{aligned}$$

brengt enkele complicaties met zich mee. Om ook tijdsafhankelijke coördinatentransformaties $\hat{x}^i = \hat{x}^i(t,x)$ in het model toe te laten moeten we voor zulke systemen een 'ruimte-tijd'-punt $(t,x)$ opvatten als een element van een variëteit $M$ die gevezeld is over de reële rechte. De bijhorende snelheidsruimte wordt vervolgens gecreëerd door de eerste-jetextensie $\pi_M : J^1M \to M$ van $M \to I\!R$. In de eerste paragrafen van het TWEEDE HOOFDSTUK brengen we twee belangrijke bundels (en enkele van de canonische structuren die op deze bundels leven) onder de aandacht: de pullbackbundel $\pi_M^*\tau_M : \pi_M^*TM \to J^1M$ en de raakbundel $\tau_{J^1M} : TJ^1M \to J^1M$ van $J^1M$. De aanwezigheid van de extra tijdsparameter $t$ zorgt er onrechtstreeks voor dat $\pi_M^*\tau_M$ een bijzondere sectie $\mathbf{T}$ bezit. Elke andere sectie $X$ (waarvoor we in het vervolg de benaming 'vectorveld langs $\pi_M$' gebruiken) heeft bijgevolg een component die parallel is met dat vectorveld.

$$X = X^0\mathbf{T} + \overline{X}.$$

De component $X^0\mathbf{T}$ zullen we gemeenzaam aanduiden met 'de tijdscomponent'. Typisch voor de tijdsafhankelijke uitbreiding van meetkundige structuren is de volgende methodiek: men vertaalt de autonome structuur naar de component $\overline{X}$ en men tracht op een natuurlijke manier de overgebleven vrijheid voor $X^0\mathbf{T}$ in te vullen.

Elk tijdsafhankelijk stelsel tweede-orde differentiaalvergelijkingen genereert op een natuurlijke manier een niet-lineaire connectie op $J^1M \to M$. Men kan zich dus de vraag stellen of, in het algemeen, ook niet-lineaire connecties op $J^1M$ kunnen gelineariseerd worden tot lineaire connecties. In de context van tijdsafhankelijke systemen zijn er in de literatuur drie verschillende constructies bekend die een lineaire connectie associëren aan een niet-lineaire connectie op $J^1M$. De drie constructies verschillen van elkaar in twee opzichten. Eerst en vooral wordt er op een verschillende manier gebruik gemaakt van de vrijheid, veroorzaakt door de extra tijdsparameter. Bovendien is de 'draagruimte' van de drie constructies verschillend: in twee gevallen leeft de lineaire connectie op de raakbundel $\tau_{J^1M}$, in het ander geval op de pullbackbundel $\pi_M^*\tau_M$. In het TWEEDE HOOFDSTUK onderzoeken we of er redenen zijn om één van de drie constructies boven de andere te verkiezen.

We stellen eerst een algemeen schema op waarin de drie constructies onderling vergeleken kunnen worden. Trouw aan het oorspronkelijke idee van een Berwald-typeconnectie, eisen we enkel de aanwezigheid van een niet-lineaire connectie op $J^1M$ als input. Een lineaire connectie op $\pi_M^*\tau_M$ wordt met D aangeduid; een lineaire connectie op $\tau_{J^1M}$ met $\nabla$. In ons vergelijkend schema laten we voorlopig de vrijheid in de tijdscomponent open. Dit betekent dat we, zowel op $\pi_M^*\tau_M$ als $\tau_{J^1M}$, klassen van lineaire connecties introduceren, waarin twee elementen van eenzelfde klasse slechts in de tijdscomponent van elkaar kunnen verschillen. Daarnaast geven we aan hoe men, startend van een lineaire connectie D op $\pi_M^*\tau_M$, een klasse van 'gelifte' connecties $\nabla$ op $\tau_{J^1M}$ kan vinden. Ook omgekeerd zullen we met elke lineaire connectie $\nabla$ op $\tau_{J^1M}$ een klasse van lineaire connecties D op $\pi_M^*\tau_M$ associëren. Het spreekt nu vanzelf dat als we beide procedures na elkaar toepassen op bijvoorbeeld een lineaire connectie op $\pi_M^*\tau_M$, we een nieuwe lineaire connectie op $\pi_M^*\tau_M$ vinden. De voorwaarden die zorgen dat de nieuwe lineaire connectie tot dezelfde klasse als de oorspronkelijke behoort, leggen een aantal restricties op aan die klassen. Lineaire connecties, zowel op $\pi_M^*\tau_M$ als op $\tau_{J^1M}$, die aan deze restricties voldoen worden 'van het Finslertype' genoemd. Men kan nu gemakkelijk binnen de klasse van lineaire connecties van het Finslertype de klasse van het Berwaldtype onderscheiden.

Alhoewel de drie voormelde constructies oorspronkelijk opgesteld werden binnen het kader van tijdsafhankelijke tweede-ordesystemen is het mogelijk om een uitbreiding uit te werken waarin ze ook toepasbaar worden voor een willekeurige niet-lineaire connectie. We tonen aan dat op die manier alle constructies (één op $\pi_M^*\tau_M$ en twee op $\tau_{J^1M}$) leiden tot een lineaire connectie van het Berwaldtype. We analyseren hun onderlinge verschil aan de hand van het al dan niet verdwijnen van hun torsiecomponenten.

Men kan zich de vraag stellen of er 'optimale' connectie van het Berwaldtype bestaat. Hiertoe onderzoeken we eerst enkele aspecten betreffende covariante afleiding van tensorvelden $\mathcal{U}$ op $\tau_{J^1M}$. Gebruik makend van de niet-lineaire connectie kan men zo'n tensorveld ontbinden in tensorvelden langs $\pi_M$. Een natuurlijke verwachting lijkt de volgende: als $\nabla$ een lineaire connectie van het Finslertype is en $\mathcal{U}$ is een (1,1)-tensorveld op $\tau_{J^1M}$, dan verdwijnt $\nabla\mathcal{U}$ op $\tau_{J^1M}$ als en slechts als al zijn componenten langs $\pi_M$ verdwij-nen na covariante afleiding d.m.v. een geassocieerde lineaire connectie D van het Finslertype. Deze eigenschap is echter slechts geldig voor een aantal lineaire connecties en legt bijgevolg opnieuw restricties op aan de klasse van aanvaardbare lineaire connecties, in het bijzonder op de actie van de connecties op de tijdscomponent van vectorvelden langs $\pi_M$. Gesterkt door bovenstaande analyse voeren we een nieuwe lineaire connectie op $\pi_M^*\tau_M$ van het Berwaldtype in die in alle opzichten de eenvoudigste kandidaat is die aan alle restricties voldoet. Daarom gebruiken we in het vervolg het adjectief 'optimaal' voor deze lineaire connectie. Merk op dat deze nieuwe lineaire connectie verschilt van de eerder in de literatuur besproken constructie op $\pi_M^*\tau_M$. We concluderen dus dat, wat de lineaire connecties op $\pi_M^*\tau_M$ betreft, er twee belangrijke connecties van het Berwaldtype zijn.

De Berwaldconnectie is slechts één van de vele connecties die in Finslermeetkunde worden aangewend. In de laatste paragrafen van het TWEEDE HOOFDSTUK bespreken we hoe enkele andere Finslerconnecties naar de huidige context kunnen vertaald worden. Ten slotte gaan we de invloed na van onze nieuwe lineaire connectie op de classificatie van afleidingsoperatoren op vormen langs $\pi_M$.

In het DERDE HOOFDSTUK vermelden we enkele toepassingsgebieden waarin men gebruik maakt van aan de Berwald-type connectie geassocieerde operatoren. We schetsen hoe deze operatoren leiden tot een meetkundige karakterisatie van lineariseerbare en scheidbare tweede-ordesystemen. We vermelden tot slot het inverse vraagstuk van de variatierekening. Een (reguliere) tijdsafhankelijke Lagrangiaan is een functie op $J^1M$ en bepaalt een tijdsafhankelijk tweede-ordesysteem. In het inverse vraagstuk gaat men na wanneer een gegeven tweede-ordesysteem afkomstig is van zo'n Lagrangiaan. De voorwaarden die deze eigenschap op het systeem leggen kan in termen van de hogervermelde operatoren uitdrukken.

De Lagrangiaanse mechanica vormt ook de link met het tweede deel van de verhandeling. In het VIERDE HOOFDSTUK herhalen we eerst Weinstein's formalisme voor Lagrangiaanse systemen op Lie-algebroïden. Een Lie-algebroïde is een vectorbundel $\tau : \mathsf{V} \to M$ waarvan de secties een (reële) Lie algebra vormen. Verder wordt elke sectie van een Lie-algebroïde 'vastgeankerd' aan een vectorveld op $M$ d.m.v. een Lie-algebrahomomorfisme $\varrho : \mathrm{Sec}(\tau) \to \mathcal{X}(M)$. Op deze manier worden de structuurfuncties $\mathsf{C}^{\,c}_{ab}$ van het Liehaakje aan de ankerafbeelding gerelateerd. Elke Lie-algebroïde genereert een Poissonstructuur op de duale bundel. De vezelafgeleiden $\frac{\partial L}{\partial \mathsf{v}^a}$ van een (reguliere) functie $L$ op $\mathsf{V}$ vormen een afbeelding tussen $\mathsf{V}$ en $\mathsf{V}^*$ waarmee men de Poissonstructuur op $\mathsf{V}^*$ tot een Poissonstructuur op $\mathsf{V}$ kan terugtrekken. Weinstein's Lagrangiaanse vergelijkingen komen dan, ruwweg, overeen met de Hamiltonvergelijkingen, met betrekking tot de Poissonstructuur op $\mathsf{V}$, voor de energie $\mathsf{v}^a \frac{\partial L}{\partial \mathsf{v}^a} - L$. Men vindt

$$\begin{aligned} \dot{x}^I &= \varrho^I_a(x)\mathsf{v}^a, \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \mathsf{v}^a}\right) &= \varrho^I_a \frac{\partial L}{\partial x^I} - \mathsf{C}^{\,c}_{ab}\mathsf{v}^b \frac{\partial L}{\partial \mathsf{v}^c}. \end{aligned}$$

Martínez toonde aan dat er een alternatief formalisme bestaat waarin deze vergelijkingen kunnen gemodelleerd worden. Dit formalisme lijkt sterk op de standaardtheorie voor autonome Lagrangiaanse systemen, zij het dat het eerder steunt op een 'geprolongeerde bundel' (We komen later nog op deze bundel terug).

Als in het bijzonder de Lie-algebroïde een raakbundel is met de standaard Lie-algebroïdestructuur, dan reduceren Weinstein's vergelijkingen zich duidelijk tot de Lagrangiaanse vergelijkingen voor autonome Lagrangianen. Een tijdsafhankelijke Lagrangiaan is echter een functie op $J^1M$, een affiene bundel, en dus geen subgeval van Weinstein's formalisme. Steunend op enkele aanwijzingen vanuit een 'rudimentaire' variatierekening, poneren we in het tweede deel van het VIERDE HOOFDSTUK hoe, binnen elke lokale kaart, de uitbreiding van het begrip Lie-algebroïde naar een affiene bundel $\pi : E \to M$

met affiene ankerafbeelding $\rho : E \to TM$ er moet uitzien. De bijhorende Lagrangiaanse vergelijkingen op $\pi$ (met in het bijzonder ook op $\pi_M : J^1M \to M$) nemen dan de volgende vorm

$$\begin{aligned} \dot{x}^I &= \rho^I_0 + \rho^I_\alpha y^\alpha, \\ \frac{d}{dt}\Big(\frac{\partial L}{\partial y^\alpha}\Big) &= \rho^I_\alpha \frac{\partial L}{\partial x^I} - (C^\gamma_{\alpha\beta} y^\beta - C^\gamma_{0\alpha})\frac{\partial L}{\partial y^\gamma}. \end{aligned}$$

Lagrangiaanse systemen op een affiene Lie-algebroïde kunnen binnen een ruimere klasse van dynamische systemen bekeken worden, de klasse van pseudo-SODEs.

$$\begin{aligned} \dot{x}^I &= \rho^I_\alpha(x) y^\alpha + \rho^I_0(x) \\ \dot{y}^\alpha &= f^\alpha(x, y). \end{aligned}$$

In het VIJFDE en het ZESDE HOOFDSTUK stellen we een meetkundig model op voor Lie-algebroïden op een affiene bundel, pseudo-SODEs en Lagrangiaanse vergelijkingen op een affiene Lie-algebroïde. In het VIJFDE HOOFDSTUK hebben we enkele algemene concepten betreffende affiene ruimtes samengebracht. De affiene functies op een affiene ruimte vormen een vectorruimte, de (uitgebreide) duale ruimte genaamd. Op zijn beurt heeft deze vectorruimte een duale die zowel een kopie van de affiene ruimte als van de onderliggende vectorruimte bevat. In het vervolg zullen we naar deze vectorruimte refereren als de biduale ruimte. Net zoals voor vectorruimtes is het mogelijk om een uitwendige algebra van vormen te ontwikkelen. Een vorm op een affiene ruimte kan opgebouwd worden uit een vorm op de onderliggende vectorruimte en een 'affien stuk'. Zo'n vorm kan bijvoorbeeld gevormd worden door het antisymmetrisch product van elementen van de duale ruimte. We tonen aan hoe het antisymmetrisch product van twee vormen op een affiene ruimte binnen de definitie past. Lie algebra's vormen de kern van het begrip Lie-algebroïde. We geven betekenis aan een 'Lie algebra over een affiene ruimte' en we tonen aan hoe ze in verband kunnen gebracht worden met Lie algebra's over de biduale vectorruimte.

In HOOFDSTUK 6 bespreken we eerst enkele algemeenheden over een affiene bundel $\pi : E \to M$, zoals bijvoorbeeld haar inbedding in een biduale bundel $\tilde{\pi} : \tilde{E} \to M$. We vermelden verder een globale één-vorm $e^0$, een canonische sectie $\mathcal{I}$ van $\mathrm{Sec}(\pi^*\tilde{\pi})$ en de verticale lift $v : \mathrm{Sec}(\pi^*\tilde{\pi}) \to \mathcal{X}(E)$. Indien we de extra aanwezigheid van een affiene afbeelding $\rho : E \to TM$ veronderstellen, dan kunnen we nu, steunend op onze definitie van een affiene Lie algebra, ook voor een affiene bundel het begrip 'Lie-algebroïde' invoeren. We tonen eerst aan dat affiene Lie-algebroïden een subklasse van Lie-algebroïden op de biduale bundel genereren, en omgekeerd. De ankerafbeelding $\tilde{\rho} : \tilde{E} \to TM$ voor de Lie-algebroïde op $\tilde{\pi}$ is een uitbreiding van de affiene ankerafbeelding $\rho$.

Eén van de aspecten die het idee ondersteunen dat de veralgemening van Lie algebra tot Lie-algebroïde zinvol is, is het volgende: de axioma's van een Lie-algebroïde zijn zo geconcipieerd dat het mogelijk wordt om op de verzameling van vormen een uitwendige afleidingsoperator $d$ te definiëren die aan de eigenschap $d^2 = 0$ voldoet. We controleren, in het affiene geval, dat er een uitwendige afgeleide voor vormen op een affiene bundel bestaat die een soortgelijke eigenschap vertoont. We tonen aan dat deze uitwendige afgeleide nauw verbonden is met de uitwendige afgeleide op de biduale Lie-algebroïde en dat de eigenschap $de^0 = 0$ nodig en voldoende is om te kunnen besluiten dat een Lie-algebroïde op de biduale afkomstig is van een affiene Lie-algebroïde. Verder tonen we aan dat er voor iedere affiene Lie-algebroïde een Poisson structuur op de uitgebreide duale bundel bestaat. We eindigen het ZESDE HOOFDSTUK met enkele voorbeelden.

Het model van Martínez voor Lagrangiaanse systemen op een (vector) Lie-algebroïde steunde op een zekere 'geprolongeerde' bundel. Alvorens terug te keren tot de hoger vermelde dynamische systemen, werken we eerst het prolongatie-idee in zijn meest algemene vorm uit. Veronderstel dat $\mu : P \to M$ een willekeurige bundel is en $\tau : \mathsf{V} \to M$ een vectorbundel met lineaire ankerafbeelding $\varrho : \mathsf{V} \to TM$. De totale variëteit van de $\varrho$-prolongatie van $\mu$ is de pullbackvariëteit $T^\varrho P = \varrho^* TP$. We zijn niet zo zeer geïnteresseerd in de natuurlijke projecties $\mu^2$ en $\varrho^1$ van deze variëteit op respectievelijk $\mathsf{V}$ en $TP$, maar in de vezeling van $T^\varrho P$ over $P$. De projectie $\mu^1 : T^\varrho P \to P$ van de geprolongeerde bundel kan als volgt gevonden worden. Eerst wordt een element van $T^\varrho P$ via $\varrho^1$ op $TP$ geprojecteerd, waarna de raakprojectie $\tau_P : TP \to P$ volgt. Elementen in $T^\varrho P$ die via $\mu^2$ op de nulvector in $\mathsf{V}$ projecteren worden 'verticaal' genoemd en de verzameling van verticale elementen wordt aangeduid met $\mathcal{V}^\varrho P$. Tot slot bestaat er een natuurlijke projectie van $T^\varrho P$ op $\mu^*\mathsf{V}$, die de volgende canonische exacte rij vervolledigt.

$$0 \to \mathcal{V}^\varrho P \to T^\varrho P \to \mu^*\mathsf{V} \to 0.$$

Verder in het hoofdstuk specificeren we de structuur van de bundels $\mu$ en $\tau$:

1. Indien $\tau$ een Lie-algebroïde is, dan is ook elke $\varrho$-prolongatie een Lie-algebroïde.

2. Indien $\tau$ de biduale vectorbundel is van een affiene bundel $\pi$ en indien de Lie-algebroïde op $\tau$ afkomstig is van een affiene Lie-algebroïde, dan is ook de Lie-algebroïde op de geprolongeerde bundel afkomstig van een affiene Lie-algebroïde.

3. Indien $\mu$ een affiene bundel $\pi$ is, maar $\tau$ willekeurig, dan is het mogelijk om een 'verticale' liftprocedure te definiëren die elementen in $\pi^*\tilde{\pi}$ afbeeldt op verticale elementen in $T^\varrho E$.

4. Indien $\mu$ een affiene Lie-algebroïde $\pi$ is en $\tau$ zijn biduale bundel, dan bestaat er 'complete' liftprocedure van een sectie van $\pi^*\tilde{\pi}$ naar een sectie van de geprolongeerde bundel (dat opnieuw op de oorspronkelijke sectie projecteert). Verder is er een globale één-vorm $\mathcal{X}^0$ en een 'verticaal endomorfisme' $S$ voor secties van de geprolongeerde bundel.

In de laatste stap bevinden we ons in de situatie waarin pseudo-SODEs kunnen gedefinieerd worden. Een pseudo-SODE kan opgevat worden als een sectie $\Gamma$ van de geprolongeerde bundel waarvoor $S(\Gamma) = 0$ en $\langle \Gamma, \mathcal{X}^0 \rangle = 1$. Om pseudo-SODEs te definiëren hebben we strikt genomen geen Lie-algebroïdestructuur nodig. Veronderstel nu dat $\pi$ een affiene Lie-algebroïde is. Neem de uitwendige afgeleide $d$ op de geprolongeerde bundel en bepaal voor $L \in C^\infty(E)$ de twee-vorm $\omega_L = d\theta_L$, $\theta_L = dL \circ S + L\mathcal{X}^0$. De Lagrangiaanse pseudo-SODE is dan de unieke pseudo-SODE $\Gamma$ waarvoor $i_\Gamma \omega_L = 0$. Op het einde van HOOFDSTUK 7 brengen we deze Lagrangiaanse systemen in verband met een vraagstuk van variatierekening.

In de volgende stap bouwen we voor de hogervermelde dynamische systemen opnieuw een theorie van connecties op. In het bijzonder willen we nagaan of er voor elke pseudo-SODE een Berwald-type linearisatieproces bestaat. Hiertoe moeten we natuurlijk eerst weten hoe we het begrip 'connectie' kunnen veralgemenen naar het kader van Lie-algebroïden. In de literatuur kent men vele versies van 'veralgemeende connecties'. Met het prolongatie-idee in het achterhoofd, is het mogelijk om een alternatieve kijk te krijgen op veralgemeende connecties in hun meest algemene gedaante. In HOOFDSTUK 8 zullen we opnieuw veronderstellen dat $\mu : P \to M$ een willekeurige bundel is, $\tau : \mathsf{V} \to M$ een vectorbundel en $\varrho : \mathsf{V} \to TM$ een lineaire afbeelding. Veralgemeende connecties kunnen nu opgevat worden als een splijting van de hoger vermelde canonische exacte rij. Als zodanig genereren ze een 'horizontaal' direct complement van de verticale subbundel van de prolongatiebundel. Daarnaast bestaat er ook een 'horizontale lift' $h : \mu^*\mathsf{V} \to TP$.

We zijn in het bijzonder geïnteresseerd in twee deelgevallen. Veronderstel dat $\mu$ een affiene Lie-algebroïde is en $\tau$ zijn corresponderende biduale Lie-algebroïde. We keren dus terug naar de situatie waarin we eerder pseudo-SODEs bespraken. We tonen aan dat de aanwezigheid van de Lie-algebroïde er voor zorgt dat er een natuurlijke constructie bestaat die met elke pseudo-SODE een veralgemeende connectie associeert.

Voor het tweede bijzondere geval is $\mu$ een affiene bundel (niet noodzakelijk met een Lie-algebroïdestructuur) terwijl $\tau$ een willekeurige vectorbundel mag blijven. Ruwweg gesproken, wanneer alle connectiecoëfficiënten affiene functies zijn, dan noemen we de veralgemeende connectie 'affien'. We tonen aan dat er in dat geval een lineaire veralgemeende connectie op de biduale bundel bestaat en we geven een globale karakterisatie voor affiene veralgemeende connecties aan de hand van hun verband met deze lineaire veralgemeende connectie. Een tweede equivalente voorstelling geeft het verband tussen de affiene veralgemeende connectie op de affiene bundel en een geassocieerde lineaire veralgemeende connectie op de onderliggende vectorbundel. Vervolgens definiëren we voor elke affiene veralgemeende connectie drie 'veralgemeende covariante afleidingsoperatoren': $\nabla$ op de affiene bundel, $\overline{\nabla}$ op de onderliggende vectorbundel en $\tilde{\nabla}$ op de biduale bundel. Tot slot kunnen we het 'affien' zijn van een veralgemeende connectie karakteriseren d.m.v. de eigenschap dat de globale één-vorm parallel is: $\tilde{\nabla}e^0 = 0$.

Een belangrijk concept met betrekking tot covariante afgeleiden betreft parallel transport langs toelaatbare krommen. In deze situatie noemt men een kromme in $\mathsf{V}$ toelaatbaar indien de raakvector van haar projectie op $M$ in het beeld van de kromme onder $\varrho$ ligt. Een horizontale gelifte kromme is een kromme in $E$ met de eigenschap dat haar raakvector in het beeld van $h$ ligt. We tonen het verband aan tussen parallel transport langs toelaatbare krommen en Lie transport langs horizontale gelifte krommen voor affiene veralgemeende connecties.

In het LAATSTE HOOFDSTUK komen we opnieuw terug op de Berwald-type connecties: deze keer als gelineariseerde versies van veralgemeende connecties. In de algemene situatie waarbij $\tau$ een willekeurige vectorbundel is en $\mu$ de affiene bundel $\pi$, is het niet mogelijk om een compacte directe bepalingsformule voor Berwald-type connecties te vinden. We definiëren daarom de Berwald-type connecties door vast te leggen op welke wijze parallel transport gebeurt langs toelaatbare banen. We tonen ook aan welk mechanisme aan de grondslag ligt van de twee keuzes voor de tijdsafhankelijke connecties op $\pi_M^*\tau_M$ uit het TWEEDE HOOFDSTUK. In de laatste paragraaf beschouwen tot slot het bijzonder geval van een veralgemeende connectie op een affiene Lie-algebroïde (al dan niet afkomstig van een pseudo-Sode die zelf al dan niet van het Lagrangiaanse type kan zijn). We tonen aan dat in dit geval er wel directe bepalingsformules bestaan voor de eerder besproken Berwald-type connecties.