The Berwald connection for time-dependent second-order differential equations and its applications in theoretical mechanics

T. Mestdag and W. Sarlet
Ghent University, Department of Mathematical Physics and Astronomy
Krijgslaan 281, 9000 Ghent

1 Introduction

The general concept of a Berwald connection relates to a kind of linearization of an arbitrary non-linear connection on a vector bundle. Yet, one will encounter this notion almost exclusively in the literature on Finsler geometry, which is a generalization of Riemannian geometry. Recently, a number of authors have extensively used a coordinate free calculus of derivation operators in the study of a variety of questions concerning the equations modelling Newtonian and Lagrangian mechanics (see the final section for a couple of examples). The calculus in question appears to be closely related to a Berwald-type connection. Motivated by this, we have carried out an extensive study [5] of the freedom in constructing a Berwald-type connection in time-dependent mechanics, leading to an optimal way of restricting this freedom. The results of this study will be briefly outlined here.

Recall first the definition of a linear connection on a bundle $\pi : A \to B$. It is an $\mathbb{R}$-bilinear map $\mathrm{D} : \mathcal{X}(B) \times Sec(\pi) \to Sec(\pi)$, which satisfies the following requirements with respect to multiplication by functions f on B:

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

[$\mathcal{X}(B)$ denotes the set of vector fields on B and elements of $Sec(\pi)$, called sections of π, are maps $\sigma : B \to A$ with the property $\pi \circ \sigma = I_B$ (the identity map on B).]

The kernel of the tangent map $T\pi$ is known as the vertical distribution $V\pi$. A non-linear connection on π is a horizontal distribution $H\pi$, that is to say a pointwise construction of a complementary space, yielding a direct sum decomposition of the tangent bundle TA of A: $TA \equiv V\pi \oplus H\pi$. The projection operator P_H of tangent vectors to A onto their horizontal component, contains all information which defines $H\pi$. If (x^i) are coordinates on B and (x^i, v^α) coordinates on A, P_H is of the form $P_H = dx^i \otimes (\frac{\partial}{\partial x^i} - \Gamma^\alpha_i \frac{\partial}{\partial v^\alpha})$, where the functions $\Gamma^\alpha_i(x, v)$ are called *connection coefficients*.

The equations of Newtonian mechanics are second-order differential equations (Sodes). In the autonomous case, $\dot{x}^i = v^i, \dot{v}^i = f^i(x, v)$ say, their geometrical representation is a vector field $\Gamma = v^i \frac{\partial}{\partial x^i} + f^i \frac{\partial}{\partial v^i}$ on the tangent bundle $\tau : TM \to M$. The reason why the theory of connections is important for the study of Sodes is that each Sode on TM comes naturally equiped with a horizontal distribution, defined by $P_H = \frac{1}{2}(I_{TM} - \mathcal{L}_\Gamma S)$, where $S = dx^i \otimes \frac{\partial}{\partial v^i}$ is the canonical vertical endomorphism on TM. The connection coefficients of a Sode-connection are given by: $\Gamma^i_j = -\frac{1}{2}\frac{\partial f^i}{\partial v^j}$.

The main part of this note is concerned with time-dependent Sodes, for which the natural geometric framework is that of a vector field

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

on the first jet bundle $J^1\pi$ of a bundle $\pi : E \to \mathbb{R}$. Again, Γ determines a non-linear connection, this time on the bundle $\pi^0_1 : J^1\pi \to E$, with projector $P_H = dt \otimes (\frac{\partial}{\partial t} - \Gamma^i_0 \frac{\partial}{\partial v^i}) + dx^i \otimes (\frac{\partial}{\partial x^i} - \Gamma^j_i \frac{\partial}{\partial v^j})$, where $\Gamma^i_j = -\frac{1}{2}\frac{\partial f^i}{\partial v^j}$ and $\Gamma^i_0 = -(f^i + \Gamma^i_j v^j)$.

2 Autonomous systems

Given a horizontal distribution on TM, the direct sum decomposition of each tangent space enables us to write for each $\xi = \xi^i_1 \frac{\partial}{\partial x^i} + \xi^i_2 \frac{\partial}{\partial v^i} \in \mathcal{X}(TM)$:

$$\xi = \xi^i_1 \left(\frac{\partial}{\partial x^i} - \Gamma^j_i \frac{\partial}{\partial v^j} \right) + \left(\xi^i_2 + \xi^j_1 \Gamma^i_j \right) \frac{\partial}{\partial v^i}.$$

Considering the pullback bundle $\tau^*\tau : \tau^*(TM) \to TM$, whose sections are called vector fields along τ and constitute a $C^\infty(TM)$-module denoted by $\mathcal{X}(\tau)$, we see that ξ can be regarded as originating from two vector fields along τ, via a process of horizontal and vertical lift, respectively. Explicitly: $\xi = {\xi_H}^H + {\xi_V}^V$, with $\xi_H = \xi^i_1 \frac{\partial}{\partial x^i}$ and $\xi_V = \left(\xi^i_2 + \xi^j_1 \Gamma^i_j \right) \frac{\partial}{\partial x^i}$. The *Berwald-type connection* on the bundle $\tau^*\tau$, associated to the given horizontal distribution, is the linear connection defined by: $\forall \xi \in \mathcal{X}(TM)$, $\forall X \in \mathcal{X}(\tau)$,

$$(1) \quad \mathrm{D}_\xi X = [P_H(\xi), X^V]_V + [P_V(\xi), X^H]_H.$$

As Crampin [2] has shown, this is the unique connection which has the following properties: (i) the restriction to fibres T_xM is the canonical complete parallelism; (ii) parallel translation along a horizontal curve is given by a rule of Lie transport.

A point to be observed here is that most other accounts of Berwald-type connections are situated within the context of Finsler-type connections (see e.g. [7]). These consist of a pair (P_H, ∇), where ∇ is a linear connection on $T\tau$, also called linear connection on TM, i.e. ∇_ξ acts on vector fields $\eta \in \mathcal{X}(TM)$, rather than on elements of $\mathcal{X}(\tau)$. The pair is said to be a Finsler-type connection if $\nabla P_H = \nabla J = 0$, where J is the almost complex structure related to the given horizontal distribution. Given a horizontal distribution and a linear connection D on $\tau^*\tau$, a Finsler pair (P_H, ∇) can be constructed by putting:

$$(2) \ \nabla_\xi X^H = (\mathrm{D}_\xi X)^H, \qquad \nabla_\xi X^V = (\mathrm{D}_\xi X)^V.$$

If D is the Berwald-type connection on $\tau^*\tau$ as defined by (1), the 'doubling procedure' (2) gives rise to the corresponding Berwald-type connection on $T\tau$. The torsion T of a connection on $T\tau$ is the skew-symmetric type (1,2) tensor field defined by: $T(\xi,\eta) = \nabla_\xi \eta - \nabla_\eta \xi - [\xi,\eta]$. Via the decomposition into horizontal and vertical lifts, it gives rise to six tensor fields along τ. Another characteristic feature of the Berwald-type connection coming from a SODE horizontal distribution is that it has maximally vanishing torsion: all of these torsion tensors vanish except for one, namely $\mathcal{R}(X,Y) = T(X^H, Y^H)_V$, which is in fact related to the curvature of the horizontal distribution.

3 The time-dependent case

With $\pi_1^0 : J^1\pi \to E$ as described before, and $\tau_E : TE \to E$, the pullback bundle $\pi_1^{0*}(\tau_E)$ takes over the role of $\tau^*\tau$ in the previous section. The module $\mathcal{X}(\pi_1^0)$ of sections of this bundle has a canonically defined element, namely the 'total time derivative' operator $\mathbf{T} = \frac{\partial}{\partial t} + v^i \frac{\partial}{\partial x^i}$. We have the decomposition:

$$(3) \qquad \mathcal{X}(\pi_1^0) \equiv \overline{\mathcal{X}}(\pi_1^0) \oplus \langle \mathbf{T} \rangle,$$

where sections in $\overline{\mathcal{X}}(\pi_1^0)$ are annihilated by dt. Likewise, if a horizontal distribution on $J^1\pi$ is given, we have

$$(4) \quad \mathcal{X}(J^1\pi) \equiv \overline{\mathcal{X}}(\pi_1^0)^H \oplus \overline{\mathcal{X}}(\pi_1^0)^V \oplus \langle \mathbf{T}^H \rangle.$$

Every vector field $\xi \in \mathcal{X}(J^1\pi)$ can then be decomposed in one of the following ways:

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

with $\overline{\xi}_H, \overline{\xi}_V \in \overline{\mathcal{X}}(\pi_1^0)$. The one-dimensional distribution spanned by $\mathbf{T}$, respectively $\mathbf{T}^H$, is the source of a certain freedom or ambiguity, when one tries to carry over the constructions of the autonomous picture to the present framework. In fact, one can find three different constructions of a linear connection associated to a time-dependent SODE in the literature. They were independently derived by Massa and Pagani [4], Byrnes [1] and Crampin *et al* [3]. The first two live on $J^1\pi$, whereas the third one is a connection on $\pi_1^{0*}(\tau_E)$. In [5] we have investigated in great detail all aspects of the differences between these constructions. We will not enter into all these aspects here, but merely highlight a few features of what the above mentioned freedom has to offer when one constructs connections of Finsler or Berwald type in the time-dependent set-up.

Let D be a linear connection on $\pi_1^{0*}(\tau_E)$, which we assume to have the property:

$$(5) \quad \mathrm{D}_\xi(\overline{\mathcal{X}}(\pi_1^0)) \subset \overline{\mathcal{X}}(\pi_1^0) \quad \forall \xi \in \mathcal{X}(J^1\pi).$$

Let further P_H represent a given horizontal distribution on $J^1\pi$. The mechanism expressed by (2) for constructing a Finsler pair in the autonomous case, in a way uses the connection coefficients of D twice: once for vertical and once for horizontal vector fields. In the present framework, we can do the same, except that there is an extra dimension now, which leaves us the freedom of selecting an arbitrary type (1,1) tensor field K on $J^1\pi$. To be precise, the following rules define a linear connection ∇ on $J^1\pi$, associated to the given D and P_H (* stands for both ^{H} and ^{V}):

$$(6) \quad \nabla_\xi \overline{X}^\star = (\mathrm{D}_\xi \overline{X})^\star, \quad \nabla_\xi \mathbf{T}^H = K(\xi).$$

It follows from (5) that ∇ has the properties

$$(7) \qquad \nabla_\xi (\overline{\mathcal{X}}(\pi_1^0)^\star) \subset \overline{\mathcal{X}}(\pi_1^0)^\star,$$

and

$$(8) \ J(\nabla_\xi \overline{X}^H) = \nabla_\xi \overline{X}^V, \ J(\nabla_\xi \overline{X}^V) = -\nabla_\xi \overline{X}^H,$$

or equivalently $\nabla_\xi J|_{\overline{\mathcal{X}}(J^1\pi)} = 0$ (where $\overline{\mathcal{X}}(J^1\pi) \equiv \overline{\mathcal{X}}(\pi_1^0)^H \oplus \overline{\mathcal{X}}(\pi_1^0)^V$). Here, J is the degenerate almost complex structure defined by $J(\overline{X}^H) = \overline{X}^V$, $J(\overline{X}^V) = -\overline{X}^H$ and $J(\mathbf{T}^H) = 0$. A pair (P_H, ∇) consisting of a horizontal distribution P_H and a linear connection ∇ on $J^1\pi$ with the properties (7) and (8) is what we call a *connection of Finsler type* here. One can prove that equivalent characterizations of the properties (7) and (8) are:

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

where $P_{\overline{H}}$ is defined by $P_H = P_{\overline{H}} + dt \otimes \mathbf{T}^H$, and $S = (dx^i - v^i dt) \otimes \frac{\partial}{\partial v^i}$ is the canonical vertical endomorphism on $J^1\pi$.

Conversely, if (P_H, ∇) is a Finsler pair on $J^1\pi$, we can define a linear connection on $\pi_1^{0*}(\tau_E)$ by putting

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

where L is a $C^\infty(J^1\pi)$-linear map from vector fields on $J^1\pi$ to sections of $\mathcal{X}(\pi_1^0)$. Any such D will have the property (5). Apart from the arbitrariness expressed by the tensor fields K and L in (6) and (9), it is clear that the essence of a Finsler-type connection on $J^1\pi$ in fact comes from a connection on $\pi_1^{0*}(\tau_E)$. In [5], various aspects have been considered which can give an indication on natural ways to fix the remaining arbitrariness. We briefly sketch one here.

Tensor fields of type (1,1) on $J^1\pi$ can be constructed from tensor fields along π_1^0. If fact, all tensor fields of interest considered so far, such as $P_{\overline{H}}$, J and S, come from the identity tensor on $\overline{\mathcal{X}}(\pi_1^0)$ via appropriate lifting procedures. If U is an arbitrary type (1,1) tensor field along π_1^0, we distinguish the following four ways for lifting it to a tensor field on $J^1\pi$:

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

Then, any type (1,1) tensor field $\mathcal{U}$ on $J^1\pi$ has a unique decomposition of the form:

$$\mathcal{U} = U_1^{H;H} + U_2^{H;V} + U_3^{V;H} + U_4^{V;V}, \tag{10}$$

where U_1 is a general tensor field along π_1^0, U_2 has the property $U_2(\mathcal{X}(\pi_1^0)) \subset \overline{\mathcal{X}}(\pi_1^0)$, $U_3(\mathbf{T}) = 0$ and U_4 has the properties of both U_2 and U_3. If (P_H, ∇) is a Finsler pair and D is a corresponding connection on $\pi_1^{0*}(\tau_E)$, one would hope that ∇-invariance of $\mathcal{U}$ gets translated into D-invariance of the U_i. One can prove that

$$\nabla_\xi \mathcal{U} = 0 \Longleftrightarrow \left\{\begin{array}{ll}\mathrm{D}_\xi U_1(X) = 0, & \mathrm{D}_\xi U_2(X) = 0,\\ \mathrm{D}_\xi U_3(\overline{X}) = 0, & \mathrm{D}_\xi U_4(\overline{X}) = 0,\end{array}\right.$$

$\forall X \in \mathcal{X}(\pi_1^0)$, $\overline{X} \in \overline{\mathcal{X}}(\pi_1^0)$, provided that

$$\nabla_\xi \mathbf{T}^H = (\mathrm{D}_\xi \mathbf{T})^H. \tag{11}$$

A natural way to restrict the as yet arbitrary tensor field K in (6) therefore is to put $K(\xi) = (\mathrm{D}_\xi \mathbf{T})^H$. This has the additional effect that the lifting procedure (6) can be cast into one compact formula, namely: $\forall \xi, \eta \in \mathcal{X}(J^1\pi)$,

$$\nabla_\xi \eta = (\mathrm{D}_\xi \eta_H)^H + (\mathrm{D}_\xi \overline{\eta}_V)^V. \tag{12}$$

Observe, however, that due to the appearance of an $\overline{X}$ in the covariant derivatives of U_3 and U_4, the restriction (11) is not sufficient to ensure that ∇-invariance of $\mathcal{U}$ is equivalent to D-invariance of the U_i. For that, we need in addition that

$$\mathrm{D}_\xi \mathbf{T} \in \langle \mathbf{T} \rangle. \tag{13}$$

Again, this is a very natural additional restriction to impose: taken in conjunction with (5), it guarantees that the connection D preserves the decomposition (3) of $\mathcal{X}(\pi_1^0)$. Moreover, in view of (11), we then further have that ∇ also preserves the decomposition (4).

Let us now come to the issue of defining Berwald-type connections in the time-dependent framework. Our first preoccupation in [5] was to obtain a scheme in which the three existing versions [1, 3, 4] of a linear connection associated to a SODE, referred to before, can rightly be termed connections of Berwald type. This can be achieved if we do not impose restrictions such as (11) and (13) and in fact treat all connections which only differ in their selection of a tensor K or L, as belonging to the same equivalence class. Guided by the definition (1) for the autonomous case, we then come to the following concept.

Definition *A linear connection* D *on* $\pi_1^{0*}(\tau_E)$ *with the property (5) belongs to the class of Berwald-type connections with respect to a given horizontal distribution, if it satisfies*

$$\mathrm{D}_\xi \overline{X} = [P_H(\xi), \overline{X}^V]_V + [P_V(\xi), \overline{X}^H]_H, \tag{14}$$

for all $\overline{X} \in \overline{\mathcal{X}}(\pi_1^0)$. *A Finsler pair* (P_H, ∇) *on* $J^1\pi$ *is said to be of Berwald type if it is lifted via (6) from a connection on* $\pi_1^{0*}(\tau_E)$ *with the property (14).*

In [5], we have shown that this definition meets the purpose and that the terminology is justified by the fact that such connections all share the two properties which were characteristic for Berwald-type connections as analysed by Crampin [2] in the autonomous case (cf. Section 2). We further have explained from this perspective in what respect the constructions in [1, 3, 4] precisely differ. One of the distinguishing features, for example, has to do with different torsion properties, but we will not enter into such details here. Instead, let us come immediately to a kind of optimal selection of a representative of the class of Berwald-type connections.

The formula (14) says nothing about the action of D on $\mathbf{T}$, and the action of ∇ on $\mathbf{T}^H$. Being left with this freedom, the most obvious choice is to take

(15) $\qquad \mathrm{D}_\xi \mathbf{T} = 0$ and $\nabla_\xi \mathbf{T}^H = 0$.

This double choice clearly meets the two natural restrictions (11) and (13) identified before. It has the additional appealing feature that there exists a direct construction formula for the complete action of this D (which perfectly matches the definition (1) of the autonomous case). We have:

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

Note that the corresponding lifted connection ∇ on $J^1\pi$, in the particular case that the horizontal distribution is the one canonically associated to a SODE, is precisely the connection constructed by Massa and Pagani [4].

As a final remark, one can verify that the above optimal selection of a representative for the Berwald class does not insist on having maximally vanishing torsion tensor fields along π_1^0. In the special case of a SODE horizontal distribution, the torsion components which do not vanish are not only those related to the curvature of the horizontal distribution ($\mathcal{R}(\overline{X}, \overline{Y}) = T(\overline{X}^H, \overline{Y}^H)_V$ and $\mathcal{R}_\mathbf{T}(\overline{X}) = T(\mathbf{T}^H, \overline{X}^H)_V$), but also $\mathcal{B}_\mathbf{T} = T(\mathbf{T}^H, \overline{X}^V)_H$ which turns out to be minus the identity on $\mathcal{X}(\pi_1^0)$.

4 Applications

The vertical and horizontal covariant derivative operators which come with the Berwald-type connection associated to a SODE, play an important role in the characterization of a variety of qualitative features of that SODE. An important role in the theory of SODEs is played by the so-called Jacobi endomorphism. It can be viewed as a $C^\infty(J^1\pi)$-linear map $\Phi : \overline{\mathcal{X}}(\pi_1^0) \to \overline{\mathcal{X}}(\pi_1^0)$, and it completely determines the curvature of the SODE connection, as well as the torsion of the associated Berwald-type connection. In fact we have $\Phi = -\mathcal{R}_\mathbf{T}$. Very often, it is appropriate to choose a local basis of vector fields along π_1^0, which is adapted to the structure of Φ (for example a basis of eigenvector fields of Φ). The covariant derivative operators then allow to replace analytical computations by intrinsic, geometrical ones. We briefly sketch two of such applications.

A SODE is said to be linearizable if, by an appropriate coordinate change, it can be cast in the form $\dot{x}^i = v^i, \dot{v}^i = A^i_j(t)v^j + B^i_j(t)x^j + a^i(t)$. From the curvature of the Berwald-type connection $(curv(\xi,\eta)X = \mathrm{D}_\xi\mathrm{D}_\eta X - \mathrm{D}_\eta\mathrm{D}_\xi X - \mathrm{D}_{[\xi,\eta]}X)$, another tensor field of interest can be detected, namely a type (1,3) tensor θ, defined by $\theta(\overline{X}, \overline{Y})\overline{Z} = curv(\overline{X}^V, \overline{Y}^H)\overline{Z}$. It turns out (see [3]) that necessary and sufficient conditions for the linearizability of a given SODE are that $\theta = 0$ and $\mathrm{D}_{\overline{X}^H}\Phi(\overline{Y}) = \mathrm{D}_{\overline{X}^V}\Phi(\overline{Y}) = 0$ for all $\overline{X}, \overline{Y} \in \overline{\mathcal{X}}(\pi_1^0)$. Such intrinsic conditions can be tested on the given data in any coordinates.

Another appliction 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 nonsingular symmetric type (0,2) tensor field g exists, such that the equivalent system $g_{ij}(\ddot{x}^j - f^j) = 0$ is a set of Euler-Lagrange equations. The conditions for the existence of such metric tensor, known as the Helmholtz conditions, can be cast in the following coordinate free way: $\mathrm{D}_\Gamma g = 0$, $\mathrm{D}_{\overline{Z}^V} g(\overline{X}, \overline{Y}) = \mathrm{D}_{\overline{Y}^V} g(\overline{X}, \overline{Z})$ and $g(\Phi(\overline{X}), \overline{Y}) = g(\Phi(\overline{Y}), \overline{X})$. It was shown in [6] that the calculus originating from the Berwald-type connection makes it possible to give a full geometrical treatment of the integrability conditions of these equations for g.

References

[1] 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.

[2] M. Crampin, Connections of Berwald type, *Publ. Math. (Debrecen)*, in press.

[3] 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.

[4] 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.

[5] T. Mestdag and W. Sarlet,The Berwald-type connection associated to time-dependent second-order differential equations, preprint (2000).

[6] 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.* (1998) **54** 233–273.

[7] J. Szilasi, Notable Finsler connections on a Finsler manifold, *Lect. Matematicas* **19** (1998) 7–34.