Publ. Math. Debrecen
57 / 3-4 (2000), 455–473

Connections of Berwald type

By M. CRAMPIN (Milton Keynes)

Abstract. Associated with any horizontal distribution on the tangent bundle of a differentiable manifold is a certain linear connection which can be regarded as a linearization of the corresponding non-linear connection. I call this linear connection a connection of Berwald type, and show how its covariant derivative operator can be specified in terms of the projections of the horizontal distribution. I explain how the Berwald connection of Finsler geometry can be regarded as a special case of this general construction, and describe the relation between the Berwald connection and the other standard Finsler connections from this point of view.

1. Introduction

A number of papers have been published recently (see [1], [5] and references therein) in which axiom systems are given for the Berwald connection and the other important connections of Finsler geometry. The discussion is generally carried out in relation to the metric derived from the energy associated with a Finsler function. This may be an appropriate approach in context, but it has the side effect, unfortunate in my view, of obscuring the fact that the Finslerian Berwald connection is a particular case of a general and natural construction which has nothing in particular to do with metrics. Thus unlike the Cartan, Chern–Rund and Hashiguchi connections, the Berwald connection can be regarded as associated primarily with the geodesic spray of the energy metric, and its metrical properties as consequences of those of the geodesic spray. The construction of the Berwald connection from the geodesic spray is in turn a particular case of a more general construction, which associates in a unique way a certain

Mathematics Subject Classification: 53B40, 53B05.
Key words and phrases: Finsler geometry, theory of connections.

linear connection with an arbitrary non-linear connection on the tangent bundle of a differentiable manifold. Following Szilasi, I have called these more general connections, connections of Berwald type.

These facts suggest an alternative to the axiomatic approach to connection theory in Finsler geometry, which is to regard the Berwald-type connection as the logically prior concept, and to treat the others as deriving from it. This paper is a description of the theory of Berwald-type and Finslerian connections using this approach.

The observation that Berwald-type connections can be defined in quite general situations dates back at least to Vilms's paper of 1968, [6], and is discussed in a number of places, for example Bejancu's book [2]. What is new and distinctive about the present paper is that it develops a synthesis of results whose interrelations have not been spelled out so explicitly before, by presenting them in the distinctive framework described in the opening paragraphs. It also attempts to give coherence to a collection of results which have previously appeared in a number of different guises: sometimes in coordinate or tensorial versions, and sometimes in more modern formalisms, of which there are at least three in current use. By the latter remark I mean that we are concerned with a family of connections in some vector bundle over the tangent bundle $\tau : TM \to M$ of a manifold M, and there are in current use three different vector bundles in which the theory is formulated. These are $T(TM)$, the tangent bundle over TM; the vertical sub-bundle $V(TM)$ of $T(TM)$; and the pull-back bundle $\tau^*(TM)$. I shall impose coherence by developing the theory in terms of the latter bundle; this is not an arbitrary choice, but the one which seems to me to be geometrically the most natural, as I shall explain below. A reader who has also read Szilasi's excellent study of Finsler connections in [5] will realise my debt to him; in fact I started work on this subject by translating his results, which are expressed in the $T(TM)$ formalism, into the $\tau^*(TM)$ formalism with which I have become familiar through joint researches with Eduardo Martínez, Willy Sarlet and others into the geometry of second-order differential equation fields and the inverse problem of the calculus of variations (see [3], [4]).

2. Linearization of non-linear connections

I shall motivate my account of Berwald-type connections by first considering the theory of linear connections on the tangent bundle.

Given a linear connection on a manifold M one can construct in a well-known way a horizontal distribution on its tangent bundle $\tau : TM \to M$. This horizontal distribution provides a comprehensive way of describing the notion of parallel translation of vectors along curves in M as an operation carried out in TM: in fact it provides two such descriptions which are, in the case of a linear connection, equivalent.

The first description is given directly in terms of the horizontal lift of a curve. Suppose given a curve $\sigma : [a,b] \to M$, with $\sigma(a) = x$ and $\sigma(b) = y$; and suppose given some $v \in T_xM$. Then the parallel translate of v along σ is obtained as follows. Let σ_v^H be the horizontal lift of σ to TM which starts at $v \in T_xM$; then $\sigma_v^H(b) \in T_yM$ is the parallel translate of v to y along σ.

Parallel translation can of course be thought of as a map between tangent spaces, say $\Pi_\sigma : T_xM \to T_yM$; in terms of horizontal lifts this map is given by $\Pi_\sigma(v) = \sigma_v^H(b)$. Since Π_σ is a linear map, it may be replaced by its tangent map $\Pi_{\sigma*}$, if the tangent spaces to the linear spaces T_xM and T_yM are identified with those spaces themselves. This observation leads to a second way of describing parallel translation. Consider the restriction of TM to the image of σ, or more exactly the pull-back bundle $\sigma^*(TM)$, which can be considered as a bundle over some open interval of $\mathbb{R}$ containing $[a,b]$. The horizontal distribution on TM pulls back to a one-dimensional distribution on $\sigma^*(TM)$, which contains a unique vector field which projects onto the coordinate vector field on $\mathbb{R}$. Since the value of this vector field at each point in $\sigma^*(TM)$ is just the tangent vector to the horizontal lift of σ at that point, or equally the horizontal lift to that point of the tangent vector to σ, it is appropriate to denote the vector field by $\dot\sigma^H$. Then the parallel translate along σ of a vector $v \in T_xM$ to y is its Lie translate with respect to the flow generated by $\dot\sigma^H$. To be exact, take any point $w \in T_xM$, and regard $v \in T_xM$, the vector which is to be parallel-translated, as a vector tangent to T_xM at w; that is to say, identify v with v_w^V, its vertical lift to w. Then the Lie translate of v_w^V along the flow of $\dot\sigma^H$ to $\sigma_w^H(b) \in T_yM$ is a vertical vector which is the vertical lift of the parallel translate of v to y along σ. The linearity of the connection ensures that the construction is well-defined in the sense that the result does not depend on the choice of w. The operation of covariant differentiation ∇ (with respect to the given linear connection on M) may be recovered as follows: for any vector fields X and Y on M,

$$[X^H, Y^V] = (\nabla_X Y)^V ,$$

where X^H is the horizontal lift of X and Y^V the vertical lift of Y, both vector fields on TM. This formula fits in naturally with this way of thinking about parallel translation.

One can think of the second description of parallel translation as being in an obvious sense a linearization of the first; though in the case of a linear connection this is a pretty pointless observation, since linearizing something which is already linear merely takes one back to where one started. But suppose now that one is given a non-linear connection on M, that is to say, a horizontal distribution on TM for which the maps Π_σ defined by horizontal lifting of curves σ in M cannot be assumed to be linear. It will still be possible to introduce a notion of linear parallel transport associated with the non-linear connection by using $\Pi_{\sigma *}$, as before; this will indeed be linear, and it will be a linearization of the given non-linear connection. When it is fully developed, this idea will prove to give rise to the Berwald-type linear connection associated with the horizontal distribution which defines the non-linear connection.

Consider again the pull-back bundle $\sigma^*(TM)$. The horizontal vector field $\dot\sigma^H$ is defined as it was before. However, it will no longer be the case that the Lie translate of v_w^V along the flow of $\dot\sigma^H$ from $w \in T_xM$ to $\sigma_w^H(b) \in T_yM$ is independent of the choice of w. So linear parallel translation will not be an operation defined on vectors at $x \in M$ along curves in M; rather it will be defined on vectors at $x \in M$ along horizontal curves in TM. That is to say, given a point $w \in T_xM$, let σ_w^H be a horizontal curve starting at w, which is the horizontal lift of some curve σ in M; and let v be an element of T_xM, considered as a vector tangent to T_xM at the point w, and therefore identifiable with a vertical vector in $T_w(TM)$ itself. Then the linear parallel translate of v to $\sigma_w^H(b) \in T_yM$ along σ_w^H is the Lie translate of $v_w^V \in T_w(TM)$ to $\sigma_w^H(b)$ by the flow of the vector field $\dot\sigma^H$ on $\sigma^*(TM)$. Since $\dot\sigma^H$ is projectable with respect to the projection $\sigma^*(TM) \to \mathbb{R}$, the action of its flow will preserve the fibration, and so the Lie translate of v_w^V will be vertical, that is, tangent to T_yM, and therefore identifiable with a vector in T_yM itself. Thus parallel translation may be regarded as a linear map $T_xM \to T_yM$, but dependent on which horizontal lift of σ is chosen, rather than just on σ itself as it is in the linear case.

The construction just described defines parallel translation of elements of each tangent space T_xM along a certain class of curves which start at points of TM over $x \in M$, namely horizontal ones. It is clearly desirable to extend the class of curves in TM along which parallel translation may be defined, if possible to all (smooth) curves. There is certainly a natural way to introduce a notion of parallel translation of elements of T_xM along curves in TM which are vertical above x, that is, along curves in T_xM: since T_xM is a vector space, it is endowed with a canonical flat symmetric affine connection or complete parallelism, in which vectors at points w_1 and w_2 are parallel if and only if they correspond to the same element of T_xM when $T_{w_1}(T_xM)$ and $T_{w_2}(T_xM)$ are identified with T_xM in the canonical way. It turns out that knowing what it is for vectors to be parallelly translated along horizontal and along vertical curves is enough to define a linear connection.

The connection thus defined will therefore be based on a law of parallel translation of vectors tangent to M, along curves in TM. To express it in terms of a covariant derivative operator, as will be done below, it is necessary to choose a vector bundle on whose sections the operator can act. That is, the required covariant derivative operator ∇_ξ, where ξ is either a tangent vector to TM or a local (i.e. locally defined) vector field on TM, must act on local sections of some vector bundle over TM, and give back elements or local sections of the same bundle as the case may be. Among the possible arguments of ∇_ξ must be vector fields on M defined along curves in M to which $\tau_*\xi$ is tangent (where $\tau : TM \to M$ is the projection). The natural choice of vector bundle over TM whose local sections over curves include such vector fields on M is the pull-back of the vector bundle $TM \to M$ to TM by τ, which I write as $\tau^*(TM)$. Indeed, a point in the fibre of $\tau^*(TM)$ over $w \in TM$ is just an element of $T_{\tau(w)}M$, by the definition of $\tau^*(TM)$. This bundle may be identified with the vertical sub-bundle of $T(TM)$, but conceptually the two are distinct: in particular, a fibre of $\tau^*(TM)$ need not, and should not, be thought of as a subspace of a larger vector space.

I denote the $C^\infty(TM)$ module of sections of $\tau^*(TM)$ by $\mathcal{X}(\tau)$; its elements may be (and often are) described as vector fields along the projection τ. Given a horizontal distribution on TM I denote by v^H_w the horizontal lift of $v \in T_xM$ to w, where $x = \tau(w)$; thus v^H_w is the unique horizontal vector at w which projects onto v. The identification of $T_w(TM)$ with $(T_xM)^H \oplus (T_xM)^V$ extends to a direct sum decomposition of $\mathcal{X}(TM)$, the module of vector fields on TM, as

$$\mathcal{X}(TM) = (\mathcal{X}(\tau))^H \oplus (\mathcal{X}(\tau))^V.$$

Every horizontal (vertical) vector field $\xi \in \mathcal{X}(TM)$ may be written uniquely in the form $\xi = X^{\scriptscriptstyle H}$ ($\xi = X^{\scriptscriptstyle V}$) for some $X \in \mathcal{X}(\tau)$; I denote such X by $\xi_{\scriptscriptstyle H}$ ($\xi_{\scriptscriptstyle V}$). Let $P_{\scriptscriptstyle H}$ ($P_{\scriptscriptstyle V}$) be the horizontal (vertical) projector. Then for arbitrary $\xi \in \mathcal{X}(TM)$ I define $\xi_{\scriptscriptstyle H} \in \mathcal{X}(\tau)$ ($\xi_{\scriptscriptstyle V} \in \mathcal{X}(\tau)$) by $(\xi_{\scriptscriptstyle H})^{\scriptscriptstyle H} = P_{\scriptscriptstyle H}(\xi)$ ($(\xi_{\scriptscriptstyle V})^{\scriptscriptstyle V} = P_{\scriptscriptstyle V}(\xi)$). Thus for $\xi \in \mathcal{X}(TM)$,

$$(\xi_{\scriptscriptstyle H})^{\scriptscriptstyle H} + (\xi_{\scriptscriptstyle V})^{\scriptscriptstyle V} = \xi,$$

while for $X \in \mathcal{X}(\tau)$,

$$\left(X^{\scriptscriptstyle H}\right)_{\scriptscriptstyle H} = \left(X^{\scriptscriptstyle V}\right)_{\scriptscriptstyle V} = X, \quad \left(X^{\scriptscriptstyle V}\right)_{\scriptscriptstyle H} = \left(X^{\scriptscriptstyle H}\right)_{\scriptscriptstyle V} = 0.$$

Let (x^i) be local coordinates on M and (x^i, v^i) the corresponding local coordinates on TM. I shall denote by $\{H_i\}$ the local basis of horizontal vector fields on TM which project onto $\partial/\partial x^i$, so that

$$H_i = \frac{\partial}{\partial x^i} - N_i^j \frac{\partial}{\partial v^j}$$

for certain local functions N_i^j on TM; and by $\{V_i\}$ the local basis of vertical vector fields given by $V_i = \partial/\partial v^i$. I shall denote by $\{\partial_i\}$ the local basis of $\mathcal{X}(\tau)$ such that $\partial_i{}^{\scriptscriptstyle V} = V_i$; then $\partial_i{}^{\scriptscriptstyle H} = H_i$.

Theorem 1. *Given any horizontal distribution on TM, there is a unique linear connection on the vector bundle $\tau^*(TM)$ with the properties that*

- *parallel translation along any horizontal curve is given by Lie transport with respect to the corresponding horizontal vector field, as described above;*
- *the restriction of the connection to any fibre T_xM of $\tau : TM \to M$ is the canonical complete parallelism on the vector space T_xM.*

The covariant derivative operator $\nabla : \mathcal{X}(TM) \times \mathcal{X}(\tau) \to \mathcal{X}(\tau)$ of the connection is defined as follows:

$$\nabla_\xi X = \left[P_{\scriptscriptstyle H}(\xi), X^{\scriptscriptstyle V}\right]_{\scriptscriptstyle V} + \left[P_{\scriptscriptstyle V}(\xi), X^{\scriptscriptstyle H}\right]_{\scriptscriptstyle H}.$$

Proof. I shall show first that the given operator is a covariant derivative operator, and then that it has the stated properties. Finally, I shall show that it is uniquely determined by these properties.

To show that the operator is indeed a covariant derivative, it is enough to show that it obeys the correct rules when its arguments are multiplied by functions. For any $f \in C^\infty(TM)$,

$$\begin{aligned}\nabla_{f\xi} X &= [fP_H(\xi), X^V]_V + [fP_V(\xi), X^H]_H \\ &= f\nabla_\xi X - (X^V f)\, P_H(\xi)_V - (X^H f)\, P_V(\xi)_H = f\nabla_\xi X,\end{aligned}$$

since the terms involving derivatives of f also involve $P_H(\xi)_V$ and $P_V(\xi)_H$, both of which are zero. On the other hand,

$$\begin{aligned}\nabla_\xi (fX) &= [P_H(\xi), fX^V]_V + [P_V(\xi), fX^H]_H \\ &= f\nabla_\xi X + (P_H(\xi)f)\,(X^V)_V + (P_V(\xi)f)\,(X^H)_H \\ &= f\nabla_\xi X + (\xi f)X.\end{aligned}$$

The condition for X to be parallel along a horizontal curve σ^H, where σ is a curve in M, is that $\nabla_{\dot\sigma^H} X = 0$. But

$$\nabla_{\dot\sigma^H} X = (\mathcal{L}_{\dot\sigma^H} X^V)_V\,,$$

where the Lie derivative is calculated in $\sigma^*(TM)$. The vector field $\dot\sigma^H$ on $\sigma^*(TM)$ is projectable, so $\mathcal{L}_{\dot\sigma^H} X^V$ is vertical. Thus $\nabla_{\dot\sigma^H} X = 0$ if and only if $\mathcal{L}_{\dot\sigma^H} X^V = 0$.

To say that the restriction of ∇ to any fibre of τ is the canonical complete parallelism is to say that the covariant derivative in any vertical direction of any basic vector field (vector field on M) – which, regarded as a section of $\tau^*(TM)$, is constant on the fibres – is zero. But if Y is basic then Y^H is projectable (and projects onto Y); and therefore $[X^V, Y^H]$ is vertical, and so $[X^V, Y^H]_H = 0$. Thus $\nabla_{X^V} Y = 0$ for any X and any basic Y.

So the given formula defines the covariant derivative operator of a connection which satisfies the given conditions. I shall now show that this connection is unique.

Suppose there is another connection which satisfies the given conditions, whose covariant derivative operator is D. The new connection determines, and is determined by, a map $\delta : \mathcal{X}(TM) \times \mathcal{X}(\tau) \to \mathcal{X}(\tau)$, such that the new covariant derivative is given in terms of the old by

$$\mathrm{D}_\xi X = \nabla_\xi X + \delta(\xi, X);$$

moreover, δ is tensorial (that is to say, $C^\infty(TM)$-linear in both arguments, and consequently pointwise defined). I shall show that if D satisfies the conditions in the statement of the theorem, as ∇ does, then $\delta = 0$ at every point of TM.

Take any point $w \in TM$, with $\tau(w) = x$ (so $w \in T_xM$), let $\xi_w \in T_w(TM)$ be an arbitrary horizontal vector, and v an arbitrary element of T_xM. Take any curve σ in M through x such that $\dot\sigma(t_0) = \tau_*\xi_w$, where $x = \sigma(t_0)$; then $\dot\sigma^H(w) = \xi_w$. Let V be the vector field along σ^H_w defined by Lie translation of v^V_w by $\dot\sigma^H$ in $\sigma^*(TM)$. Then $\mathrm{D}_{\dot\sigma^H}V = \nabla_{\dot\sigma^H}V = 0$, and therefore $\delta_w(\xi_w, v) = 0$. Now let $\eta_w \in T_w(TM)$ be vertical, and let X be any basic vector field such that $X(x) = v$: then $\mathrm{D}_{\eta_w}X = \nabla_{\eta_w}X = 0$, and so $\delta_w(\eta_w, v) = 0$ also. But these two results together show that $\delta_w(\zeta_w, v) = 0$ for any $w \in TM$, any $\zeta_w \in T_w(TM)$, and any $v \in T_{\tau(w)}M$, and so $\mathrm{D} = \nabla$. $\square$

If $H_i = \partial/\partial x^i - N^j_i\partial/\partial v^j$, then

$$\nabla_{H_i}\partial_j = \frac{\partial N^k_i}{\partial v^j}\partial_k.$$

3. Berwald-type connections and others

Not every connection on $\tau^*(TM)$ is of Berwald type. In this section I shall derive the necessary and sufficient conditions for a connection D on $\tau^*(TM)$ which reduces to the canonical complete parallelism on the fibres of TM to be of Berwald type.

It may easily be checked that for any connection on $\tau^*(TM)$, and any horizontal distribution on TM, each of the following expressions is tensorial in both of its arguments (that is, it is $C^\infty(TM)$-linear in $X, Y \in \mathcal{X}(\tau)$), where D is the covariant derivative operator:

$$\begin{aligned}
\mathcal{A}(X,Y) &= \mathrm{D}_{X^H}Y - \mathrm{D}_{Y^H}X - [X^H, Y^H]_H\\
\mathcal{R}(X,Y) &= -[X^H, Y^H]_V\\
\mathcal{B}(X,Y) &= -\mathrm{D}_{Y^V}X - [X^H, Y^V]_H\\
\mathcal{P}(X,Y) &= \mathrm{D}_{X^H}Y - [X^H, Y^V]_V\\
\mathcal{S}(X,Y) &= \mathrm{D}_{X^V}Y - \mathrm{D}_{Y^V}X - [X^V, Y^V]_V\,.
\end{aligned}$$

These type $(1,2)$ tensor fields along τ may be regarded as torsions of D. The reason for this is to be found in the fact that any connection on $\tau^*(TM)$, and any horizontal distribution on TM, between them induce a connection on $T(TM)$, whether or not the two are in any way related. The covariant derivative operator of this connection on $T(TM)$ is defined by

$$\mathrm{D}_\xi\left(X^H\right) = \left(\mathrm{D}_\xi X\right)^H, \quad \mathrm{D}_\xi(Y^V) = \left(\mathrm{D}_\xi Y\right)^V.$$

(That is to say, the connection on $\tau^*(TM)$ may be extended to $T(TM)$ by exploiting the ambiguity in each of the expressions $\mathrm{D}_\xi X^H$ and $\mathrm{D}_\xi Y^V$.) The tensors defined above are the components of the torsion of the induced linear connection on $T(TM)$. (The conditions under which a given linear connection on $T(TM)$ comes from one on $\tau^*(TM)$ in this manner are discussed by SZILASI [5], and my notation for the components of the torsion is adapted from his.)

Note that $\mathcal{R}$ depends only on the horizontal distribution (it measures the lack of integrability, in the sense of Frobenius, of the distribution). Any connection which reduces to the canonical complete parallelism on the fibres of $T(TM)$ must have $\mathcal{B} = \mathcal{S} = 0$ for any horizontal distribution. For the connection of Berwald type associated with the horizontal distribution $X \mapsto X^H$ the torsion $\mathcal{P}$ vanishes in addition.

I shall now derive the necessary and sufficient conditions for a connection on $\tau^*(TM)$ which induces the canonical complete parallelism on the fibres to be of Berwald type – that is, for there to be some horizontal distribution for which it is the associated Berwald-type connection. The conditions may be formulated in two ways, one of which involves the torsion $\mathcal{P}$, the other the curvature *curv*, of the given connection.

Proposition 1. *Let* D *be the covariant derivative operator of a connection on $\tau^*(TM)$ which reduces to the canonical complete parallelism on the fibres. Let $X \mapsto X^H$ be any horizontal distribution. The following two conditions are equivalent: for all $X, Y, Z \in \mathcal{X}(\tau)$*

$$\left(\mathrm{D}_{Y^V}\mathcal{P}\right)(X,Z) = \left(\mathrm{D}_{Z^V}\mathcal{P}\right)(X,Y)$$

$$\mathit{curv}\left(X^H, Z^V\right)Y = \mathit{curv}\left(X^H, Y^V\right)Z;$$

and if either holds for one horizontal distribution both hold for all. Either condition is necessary and sufficient for there to be some horizontal distribution for which D *is the associated Berwald-type connection.*

Proof. I first demonstrate the equivalence of the two conditions. This comes from the first Bianchi identity for the connection, which in turn follows from the Jacobi identity

$$\big[[X^H, Y^V], Z^V\big] + \big[[Y^V, Z^V], X^H\big] + \big[[Z^V, X^H], Y^V\big] = 0.$$

When the brackets are expressed in terms of covariant derivatives the vertical component gives

$$curv\,(X^H, Z^V)\,Y - curv\,(X^H, Y^V)\,Z = (\mathrm{D}_{Y^V}\mathcal{P})\,(X, Z) - (\mathrm{D}_{Z^V}\mathcal{P})\,(X, Y),$$

taking account of the facts that $\mathcal{B} = \mathcal{S} = 0$ for such a connection. The restriction of the connection to vertical vector fields is flat, that is, it satisfies $curv(X^V, Y^V)Z = 0$; thus if the condition

$$curv(X^H, Z^V)Y - curv(X^H, Y^V)Z = 0$$

holds for some horizontal distribution it holds for all. On the other hand, $\mathcal{P} = 0$ for the Berwald-type connection associated with $X \mapsto X^H$, where $\mathcal{P}$ is calculated with respect to this horizontal distribution, and therefore both conditions hold for Berwald-type connections for all horizontal distributions.

The converse may be proved using either condition. Suppose that

$$\mathrm{D}_{H_i}\partial_j = A^k_{ij}\partial_k.$$

Then the relevant curvature component is given by

$$curv(H_i, V_j)\partial_k = -\frac{\partial A^l_{ki}}{\partial v^j}\partial_l,$$

and so the curvature condition is

$$\frac{\partial A^l_{ki}}{\partial v^j} = \frac{\partial A^l_{kj}}{\partial v^i};$$

it follows that there are functions N^l_k such that

$$A^l_{kj} = \frac{\partial N^l_k}{\partial v^j},$$

and the connection is the Berwald-type connection associated with the horizontal distribution spanned by $\{H_i\}$ where

$$H_i = \frac{\partial}{\partial x^i} - N_i^j \frac{\partial}{\partial v^j}.$$

Alternatively, consider the effect of a change of horizontal distribution on $\mathcal{P}$: if the new horizontal lift of X is $X^H + \delta(X)^V$, then the new torsion is $\mathcal{P}(X,Y) + (\mathrm{D}_{Y^V}\delta)(X)$; the condition $(\mathrm{D}_{Y^V}\mathcal{P})(X,Z) - (\mathrm{D}_{Z^V}\mathcal{P})(X,Y) = 0$ is the integrability condition for the differential equation $(\mathrm{D}_{Y^V}\delta)(X) = -\mathcal{P}(X,Y)$ for δ. □

Given any connection on $\tau^*(TM)$ for which $\mathcal{S} = 0$, and any function F on TM, we can associate with F a symmetric type $(0,2)$ tensor field along τ, which may be called the Hessian of F with respect to the fibre coordinates in TM, as follows. (By a type $(0,m)$ tensor field along τ I mean an object with the indicated tensorial properties, which takes its arguments from $\mathcal{X}(\tau)$.)

Proposition 2. *Let* D *be the covariant derivative operator of a connection on $\tau^*(TM)$ for which $\mathcal{S} = 0$, and let F be any function on TM. Set*

$$g_F(X,Y) = X^V\,(Y^V F) - (\mathrm{D}_{X^V} Y)^V F,$$

for any $X, Y \in \mathcal{X}(\tau)$. Then g is a symmetric type $(0,2)$ tensor field along τ. Moreover, if the connection reduces to the canonical complete parallelism on the fibres of TM then

$$(\mathrm{D}_{X^V} g_F)(Y,Z) = (\mathrm{D}_{Y^V} g_F)(X,Z)$$

for any $X, Y, Z \in \mathcal{X}(\tau)$.

Proof. Clearly

$$g_F(X,Y) - g_F(Y,X) = [X^V, Y^V]\,F - (\mathrm{D}_{X^V} Y)^V\,F + (\mathrm{D}_{Y^V} X)^V\,F = 0.$$

Moreover, for any function f on TM, $g_F(fX,Y) = f g_F(X,Y)$. These two facts together establish that g_F is a symmetric type $(0,2)$ tensor field along τ. It is a straightforward matter to show that

$$\begin{aligned}(\mathrm{D}_{X^V} g_F)\,(Y,Z) - (\mathrm{D}_{Y^V} g_F)\,(X,Z) &= [X^V, Y^V]\,(Z^V F) - (\mathrm{D}_{X^V} Y)^V\,(Z^V F) \\ &\quad + (\mathrm{D}_{Y^V} X)^V\,(Z^V F) - (\mathrm{curv}\,(X^V, Y^V)\,Z)^V\,F,\end{aligned}$$

whence if $\mathcal{S} = 0$ and $curv(X^V, Y^V) = 0$, as is the case for a connection which reduces to the canonical complete parallelism on the fibres of TM, then

$$(D_{X^V} g_F)(Y, Z) - (D_{Y^V} g_F)(X, Z) = 0. \qquad \square$$

4. Berwald-type connections associated with semi-sprays and sprays

Having reduced the concept of a Berwald connection to its barest essentials, I now commence the task of reconstruction, that is, of progressively specializing Berwald-type connections so as to obtain the Finslerian Berwald connection. There are three stages to this process, corresponding to three salient facts about the Finslerian Berwald connection: firstly, the horizontal distribution with which it is associated is derived from a semi-spray, or second-order differential equation field; secondly, the semi-spray is actually a spray, so that both it and the horizontal distribution it defines satisfy homogeneity conditions with respect to dilations of the fibres; and thirdly, the spray is actually the Euler–Lagrange field of the energy associated with a Finsler function, or in other words its geodesic spray.

Most of this story is well known, so I shall deal with it briefly and omit some proofs. I shall also ignore complications relating to the behaviour near the zero section of the geometric objects I shall define.

I denote by S the vertical endomorphism on TM, and Δ the Liouville field (the generator of dilations of the fibres): $\Delta = \mathbf{T}^V$, where $\mathbf{T} \in \mathcal{X}(\tau)$ is the total derivative operator $C^\infty(M) \to C^\infty(TM)$ defined by $(\mathbf{T}f)(v) = vf$ for $v \in TM$, $f \in C^\infty(TM)$. A vector field Γ on TM is a semi-spray if $S(\Gamma) = \Delta$.

A semi-spray Γ determines a horizontal distribution on TM, whose horizontal projector is given by

$$P_H = \frac{1}{2}\left(I - \mathcal{L}_\Gamma S\right).$$

Not every horizontal distribution can be derived from a semi-spray in this way. The conditions for a horizontal distribution to be derived from a semi-spray may be expressed in terms of the associated Berwald-type connection ∇, as follows.

Proposition 3. *The necessary and sufficient condition for a horizontal distribution to be derived from a semi-spray is that the torsion $\mathcal{A}$ of the associated Berwald-type connection vanishes.*

Proof. Of the several ways of stating the necessary and sufficient condition for a horizontal distribution to be derived from a semi-spray the most convenient for my purposes is the following: a horizontal distribution on TM, with corresponding horizontal lift $X \mapsto X^H$, is derived from a semi-spray if and only if for every pair of vector fields X and Y on M

$$[X^H, Y^V] - [Y^H, X^V] = [X, Y]^V .$$

Now $\mathcal{A}(X, Y) = \nabla_{X^H} Y - \nabla_{Y^H} X - [X^H, Y^H]_H$, and $\nabla_{X^H} Y = [X^H, Y^V]_V$. When $X, Y \in \mathcal{X}(M)$, X^H and Y^H are τ-projectable, and project onto X and Y respectively. Thus $[X^H, Y^H]_H = [X, Y]$ (regarded, as it may be, as an element of $\mathcal{X}(\tau)$). Moreover, $[X^H, Y^V]$ and $[Y^H, X^V]$ are both vertical, and so for $X, Y \in \mathcal{X}(M)$, $[X^H, Y^V] - [Y^H, X^V] = [X, Y]^V$ if and only if $[X^H, Y^V]_V - [Y^H, X^V]_V = [X, Y]$. But since $\mathcal{A}$ is a tensor, it vanishes if and only if it vanishes whenever its arguments are taken from $\mathcal{X}(M)$. □

Thus a Berwald-type connection associated with the horizontal distribution derived from a semi-spray has all of its torsions which can be zero equal to zero.

I turn next to the homogeneity conditions.

A horizontal distribution is homogeneous if $\mathcal{L}_\Delta P_H = 0$, or equivalently if the horizontal lift of any vector field on M is homogeneous of degree zero with respect to dilations of the fibres, that is, $[\Delta, X^H] = 0$ for all $X \in \mathcal{X}(M)$. The condition for a horizontal distribution to be homogeneous can also be expressed in terms of the corresponding Berwald-type connection.

Proposition 4. *The necessary and sufficient condition for the horizontal distribution $X \mapsto X^H$ to be homogeneous is that $\nabla_{X^H} \mathbf{T} = 0$ for all $X \in \mathcal{X}(\tau)$.*

Proof. We have

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

If the horizontal distribution is homogeneous then $\nabla_{X^H} \mathbf{T} = 0$ for $X \in \mathcal{X}(M)$, and therefore for all $X \in \mathcal{X}(\tau)$ by the $C^\infty(TM)$-linearity of $X \mapsto$

$\nabla_{X^H}\mathbf{T}$. If $\nabla_{X^H}\mathbf{T} = 0$ then $[\Delta, X^H]_V = 0$ for all $X \in \mathcal{X}(\tau)$, and in particular for all $X \in \mathcal{X}(M)$; but in the latter case $[\Delta, X^H]$ is vertical, so it follows that $[\Delta, X^H] = 0$. □

Given any connection which reduces to the canonical complete parallelism on the fibres one can construct a horizontal distribution (as pointed out by Abate [1]) as follows. The map $T_w(TM) \to T_{\tau(w)}M$ by $\xi_w \mapsto D_{\xi_w}\mathbf{T}$ is linear; its restriction to $V_w(TM)$ is the identity; by the rank and nullity theorem its kernel must be of dimension n; the kernel is clearly complementary to $V_w(TM)$ in $T_w(TM)$, and therefore defines a horizontal subspace of $T_w(TM)$. If one starts with a horizontal distribution, takes its Berwald-type connection, and then carries out this construction, then the horizontal distribution so defined will in general be different from the original one. Only in the case of a homogeneous horizontal distribution will they be the same.

A semi-spray Γ is a spray if it is homogeneous of degree one with respect to dilations of the fibres, that is, if $[\Delta, \Gamma] = \Gamma$.

Proposition 5. *If Γ is a spray, then*

1. *the horizontal distribution associated with Γ is homogeneous*
2. *Γ is horizontal*
3. $\nabla_\Gamma \mathbf{T} = 0$.

Proof. (1) If $[\Delta, \Gamma] = \Gamma$ then

$$\mathcal{L}_\Delta P_H = -\frac{1}{2}\mathcal{L}_\Delta \mathcal{L}_\Gamma S = -\frac{1}{2}\left(\mathcal{L}_\Gamma \mathcal{L}_\Delta S + \mathcal{L}_\Gamma S\right) = 0,$$

since $\mathcal{L}_\Delta S = -S$.

(2) We have

$$P_V(\Gamma) = \frac{1}{2}\left(\Gamma + \mathcal{L}_\Gamma S(\Gamma)\right) = \frac{1}{2}(\Gamma - [\Delta, \Gamma]),$$

and so $P_V(\Gamma) = 0$ if (and indeed only if) Γ is a spray.

(3) If Γ is a spray it is horizontal, and the corresponding distribution is homogeneous, so $\nabla_\Gamma \mathbf{T} = 0$ by Proposition 3. □

Given any horizontal distribution, it is natural to ask which curves on M are autoparallel. In general this question is ambiguous: does it mean autoparallel with respect to the non-linear or the linear connection?

If the former, then the answer is the projections onto M of the integral curves of the semi-spray $\mathbf{T}^H$. If the latter, the answer is slightly more complicated: there is a unique semi-spray $\widehat{\Gamma}$ such that $\nabla_{\widehat{\Gamma}}\mathbf{T} = 0$, and the required curves are the projections onto M of the integral curves of $\widehat{\Gamma}$. The semi-sprays $\mathbf{T}^H$ and $\widehat{\Gamma}$ are in general distinct. Worse, if the horizontal distribution is derived from a semi-spray Γ, in general all three semi-sprays Γ, $\mathbf{T}^H$ and $\widehat{\Gamma}$ are distinct. When Γ is a spray, however, they are the same, and the autoparallel curves of both non-linear and linear connections are the projections onto M of the integral curves of Γ.

I now restrict my attention to Berwald-type connections associated with sprays.

The Berwald-type covariant derivative is defined in terms of the bracket operation on vector fields on TM. For a Berwald-type connection derived from a spray, the definition can be turned round so as to express the bracket in terms of covariant derivatives: for $X, Y \in \mathcal{X}(\tau)$,

$$[X^V, Y^V] = (\nabla_{X^V} Y)^V - (\nabla_{Y^V} X)^V$$

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

$$[X^H, Y^H] = (\nabla_{X^H} Y)^H - (\nabla_{Y^H} X)^H - \mathcal{R}(X,Y)^V.$$

The particular cases of these formulae involving the spray, $\Gamma = \mathbf{T}^H$, are useful:

$$[\Gamma, Y^V] = (\nabla_\Gamma Y)^V - Y^H$$

$$[\Gamma, Y^H] = (\nabla_\Gamma Y)^H - \mathcal{R}(\mathbf{T}, Y)_V.$$

I shall need the following result in the next section.

Proposition 6. *Let Γ be a spray. Then for any $X \in \mathcal{X}(\tau)$,*

$$\mathit{curv}(\Gamma, X^V) = 0.$$

Proof. For any $Y \in \mathcal{X}(\tau)$,

$$\big[\Gamma, [X^V, Y^H]\big]_H + \big[X^V, [Y^H, \Gamma]\big]_H + \big[Y^H, [\Gamma, X^V]\big]_H = 0,$$

by Jacobi's identity. If one now substitutes for the brackets, using the formulae above, one finds that

$$\nabla_\Gamma \nabla_{X^V} Y - \nabla_{X^V} \nabla_\Gamma Y - \nabla_{(\nabla_\Gamma X)^V} Y + \nabla_{X^H} Y = 0.$$

But $(\nabla_\Gamma X)^V - X^H = [\Gamma, X^V]$, so this says that $\mathit{curv}(\Gamma, X^V)Y = 0$. $\square$

5. The Finslerian Berwald connection

Let Γ be the geodesic spray of a Finsler space whose energy is E, let $X \mapsto X^H$ be the homogeneous horizontal distribution defined by Γ, and ∇ the covariant derivative of the corresponding Berwald-type connection.

The energy metric g of the Finsler space is the Hessian of E, as defined in Proposition 2. The Cartan tensor C of g is the type $(1, 2)$ tensor field along τ defined by

$$g(C(X, Y), Z) = (\nabla_{X^V} g)\,(Y, Z).$$

It follows from Proposition 2 that C is symmetric, and that $g(C(X, Y), Z)$ is symmetric in all three arguments.

It is well-known that when Γ is the geodesic spray, $\nabla_\Gamma g = 0$. This is easily established by using the expression of the Euler–Lagrange equations in terms of the Cartan 2-form $\omega = d(dE \circ S)$. This 2-form is related to g by $g(X, Y) = \omega(X^V, Y^H)$; moreover, $\mathcal{L}_\Gamma \omega = 0$, from which the required result follows by evaluation on X^V, Y^H. The homogeneity of E is not required in this proof, and in fact the result holds for the Euler–Lagrange semi-spray of any Lagrangian L, where $g = g_L$ is the Hessian. Indeed, $\nabla_\Gamma g = 0$ and $(\nabla_{X^V} g)(Y, Z) = (\nabla_{Y^V} g)(X, Z)$ are two of the Helmholtz conditions, which together comprise the usual starting point for investigations of the inverse problem of the calculus of variations, where now Γ is a given semi-spray and the task is to find g satisfying the conditions, if possible; see for example [4].

I use the fact that $\nabla_\Gamma g = 0$, together with Proposition 6, in the following proposition.

Proposition 7. *Let C' be the type $(1, 2)$ tensor field along τ defined by*

$$g(C'(X, Y), Z) = (\nabla_{X^H} g)\,(Y, Z).$$

Then

$$C' = -\nabla_\Gamma C.$$

Proof. The vanishing of the curvature component $curv(\Gamma, X^V)$, together with the fact that $\nabla_\Gamma g = 0$, implies that

$$\nabla_\Gamma \nabla_{X^V} g - \nabla_{(\nabla_\Gamma X)^V} g = -\nabla_{X^H} g.$$

But $\nabla_{X^V} g = g(C(X,\cdot),\cdot)$, so (using $\nabla_\Gamma g = 0$ again)

$$\nabla_\Gamma \nabla_{X^V} g - \nabla_{(\nabla_\Gamma X)^V} g = g((\nabla_\Gamma C)(X,\cdot),\cdot),$$

whence $C' = -\nabla_\Gamma C$. □

It follows that C' is symmetric, and that $g(C'(X,Y),Z)$ is symmetric in all three arguments.

By homogeneity, the energy may be expressed as $E = \frac{1}{2} g(\mathbf{T},\mathbf{T})$; it follows that for any $X \in \mathcal{X}(\tau)$

$$X^H(E) = \frac{1}{2}\left(\nabla_{X^H} g\right)(\mathbf{T},\mathbf{T}) + g\left(\nabla_{X^H}\mathbf{T},\mathbf{T}\right) = \frac{1}{2}\left(\nabla_\Gamma g\right)(X,\mathbf{T}) = 0.$$

Now according to Abate [1], Theorem 2.3 (slightly modified to fit in with my notation etc.), the Finslerian Berwald connection is uniquely determined by the properties that for all $X \in \mathcal{X}(\tau)$

1. $\nabla_{X^H}\mathbf{T} = 0$
2. $X^H(E) = 0$
3. $\mathcal{A} = \mathcal{B} = \mathcal{P} = 0$.

The Berwald-type connection associated with the geodesic spray Γ has these properties, which confirms (if confirmation is required) that it is indeed the Finslerian Berwald connection.

The symmetry of C and C' has the consequences:
$C(\mathbf{T},\cdot) = 0$, by homogeneity; $C'(\mathbf{T},\cdot) = 0$, since $\nabla_\Gamma g = 0$.

6. The Chern–Rund, Cartan and Hashiguchi connections

As has been noted before, a change of connection determines, and is determined by, a tensorial map $\delta : \mathcal{X}(TM) \times \mathcal{X}(\tau) \to \mathcal{X}(\tau)$, such that the new covariant derivative is given in terms of the old (which I take to be the Berwald one) by

$$D_\xi X = \nabla_\xi X + \delta(\xi, X).$$

The torsions of the new connection are given by

$$\mathcal{A}(X,Y) = \delta\left(X^H, Y\right) - \delta\left(Y^H, X\right)$$

$$\mathcal{B}(X,Y) = -\delta\left(Y^V, X\right)$$

$$\mathcal{P}(X,Y) = \delta\left(X^H, Y\right)$$

$$\mathcal{S}(X,Y) = \delta\left(X^V, Y\right) - \delta\left(Y^V, X\right).$$

In particular, $\mathcal{A} = 0$ if and only if δ^H is symmetric (where δ^H is the type $(1,2)$ tensor field along τ given by $\delta^H(X,Y) = \delta(X^H, Y)$); and $\mathcal{S} = 0$ if and only if δ^V is symmetric (where δ^V is defined analogously to δ^H).

Furthermore,

$$\begin{aligned}(D_\xi g)(X,Y) = g\left(C\left(\xi_V, X\right), Y\right) + g\left(C'\left(\xi_H, X\right), Y\right) \\ - g(\delta(\xi, X), Y) - g(X, \delta(\xi, Y)).\end{aligned}$$

Theorem 2. *The necessary and sufficient condition for* $\mathcal{A} = 0$ *and* $\mathrm{D}_g^H = 0$ *is*

$$\delta^H = \frac{1}{2} C';$$

the necessary and sufficient condition for $\mathcal{S} = 0$ *and* $\mathrm{D}_g^V = 0$ *is*

$$\delta^V = \frac{1}{2} C.$$

PROOF. The condition $\mathrm{D}_g^H = 0$ amountsto

$$g(C'(Z,X),Y) = g\left(\delta^H(Z,X),Y\right) + g\left(X, \delta^H(Z,Y)\right).$$

These equations may be solved for δ^H using the Christoffel trick, taking into account the symmetry of δ^H required to satisfy $\mathcal{A} = 0$. The proof of the second assertion is similar. □

Each of these conditions can be imposed independently, and they can be imposed together. The three possibilities produce the other major connections of Finsler geometry, as follows.

The Chern–Rund connection

The Chern–Rund connection is given by

$$\delta^H = \frac{1}{2} C', \quad \delta^V = 0.$$

All of its torsions except $\mathcal{R}$ and $\mathcal{P}$ vanish; we have $\mathcal{P}(X,Y) = \frac{1}{2} C'(X,Y)$, and in particular $\mathcal{P}(\mathbf{T}, \cdot) = 0$. The Chern–Rund connection is "horizontally-metrical", that is, $\mathrm{D}^H g = 0$ (but $\mathrm{D}^V g = \nabla^V g$ is given by $(D_{X^V} g)(Y,Z) = g(C(X,Y),Z)$). The Chern–Rund connection induces the canonical complete parallelism on the fibres of τ.

The Hashiguchi connection

The Hashiguchi connection is given by

$$\delta^V = \frac{1}{2}C, \quad \delta^H = 0.$$

All of its torsions except $\mathcal{R}$ and $\mathcal{B}$ vanish; we have $\mathcal{B}(X,Y) = -\frac{1}{2}C(X,Y)$, and in particular $\mathcal{B}(\mathbf{T},\cdot) = 0$. The Hashiguchi connection is "vertically-metrical", that is, $\mathrm{D}^V_g = 0$ (but$\mathrm{D}^H g = \nabla^H g$ is given by $(D_{X^H} g)(Y,Z) = g(C'(X,Y),Z))$. The Hashiguchi connection does not induce the canonical complete parallelism on the fibres of τ.

The Cartan connection

The Cartan connection is given by

$$\delta^V = \frac{1}{2}C, \quad \delta^H = \frac{1}{2}C'.$$

Its torsions $\mathcal{A}$ and $\mathcal{S}$ vanish; we have $\mathcal{B}(X,Y) = -\frac{1}{2}C(X,Y)$ and $\mathcal{P}(X,Y) = \frac{1}{2}C'(X,Y)$, and in particular $\mathcal{B}(\mathbf{T},\cdot) = \mathcal{P}(\mathbf{T},\cdot) = 0$. The Cartan connection is metrical, that is, $\mathrm{D}^V g = \mathrm{D}^H g = 0$. The Cartan connection does not induce the canonical complete parallelism on the fibres of τ.

References

[1] M. Abate, A characterization of the Chern and Berwald connections, *Houston J. Math.* **22** (1996), 701–717.

[2] A. Bejancu, Finsler geometry and applications, *Ellis Horwood*, 1990.

[3] M. Crampin E. Martínez and W. Sarlet, Linear connections for systems of second-order ordinary differential equations, *Ann. Inst. H. Poincaré* **65** (1996), 223–249.

[4] 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.

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

[6] J. Vilms, Curvature of nonlinear connections, *Proc. Amer, Math. Soc.* **19** (1968), 1125–1129.

M. CRAMPIN
DEPARTMENT OF APPLIED MATHEMATICS
THE OPEN UNIVERSITY
WALTON HALL, MILTON KEYNES MK7 6AA
UK

(Received April 6, 1999; file arrived September 8, 1999)