The second variation formula in Lagrange and Finsler geometry

M. Crampin
Department of Applied Mathematics, The Open University,
Walton Hall, Milton Keynes MK7 6AA, UK.

June 13, 2002

1 Introduction

One of the main ideas underpinning Chern's assertion that 'Finsler geometry is just Riemannian geometry without the quadratic restriction' [3] is that the formula for the second variation of arc length familiar from Riemannian geometry has an extension to Finsler geometry which is formally identical. As a consequence, a number of theorems of global Riemannian geometry, such as the theorems of Myers and Synge, hold in the Finsler setting. That this is so had already been observed by Auslander in the 1950s [1]; his results were rederived in a somewhat different form by Bao and Chern in 1993 [2].

The ideas associated with the second variation formula are not restricted to Riemannian and Finsler geometry, of course, but apply equally to any ordinary problem in the calculus of variations for a first-order Lagrangian with finitely many degrees of freedom. As I shall show below, it is possible to derive a second variation formula which is formally similar to the formula of Riemannian geometry in this situation also. I want to emphasise that I am referring here, in contrast to the Finslerian case, to variational problems in which the Lagrangian is regular, so that the parametrization of the extremals is fixed by the Euler-Lagrange equations; such problems are sometimes called, slightly confusingly, nonparametric ([8]). For maximum similarity with Riemannian geometry (which of course is the special case in which the Lagrangian is the kinetic energy associated with the metric) it must be assumed that the Hessian of the Lagrangian with respect to the velocity variables is not just non-singular but is in addition positive definite.

The second variation formula for such variational problems has been widely discussed, for example in [9] and [8], but only in coordinate form; what is distinctive about my version is that it is covariant. My aim in this paper, in addition to presenting a relatively straightforward derivation of this general covariant second variation formula, is to investigate its relationship with the formula obtained by Bao and Chern.

One of the complicating factors of Finsler geometry is the existence of several different linear connections, each of which shares some of the desirable features of the Levi-Civita connection but none of which has all of them. These various possibilities have been admirably surveyed by Szilasi in a recent paper [12]; see also [5]. One major difference between the approach of Auslander, on the one hand, and Bao and Chern, on the other, to the formulation of the second variation formula is in the choice of linear connection used; indeed, a large part of the Bao-Chern paper is devoted to 'explor[ing] the merits of a special connection in Finsler geometry' (as it says in the Abstract), now known as the Chern-Rund connection, and in particular to demonstrating its advantages in the context of the variational theory of geodesics over the Cartan connection, which Auslander employed. However, Chern admits that the so-called flag curvature, which is the quantity which in Finsler geometry plays the role of the sectional curvature in Riemannian geometry, is 'independent of whichever of the standard connections (among several!) we choose', and indeed that it may not require a connection for its definition at all [3].

My derivation of the second variation formula does not involve a linear connection. It makes use instead of a certain differential operator and a certain type $(1,1)$ tensor field intrinsically associated with the Euler-Lagrange field of the system. The differential operator was called by Foulon [7], who first introduced it (though in a slightly different context from that prevailing here) the dynamical derivative; the tensor field he called the Jacobi endomorphism. It is important to note that these objects do not require a linear connection for their definition. They can be defined in fact for any second-order differential equation field, whether or not it is derivable from a Lagrangian, as I shall explain below.

Though the dynamical derivative and the Jacobi endomorphism are logically independent of any connection, they can be related to the various Finsler connections: the dynamical derivative can be expressed as a covariant differentiation operator, and the Jacobi endomorphism as a curvature component. I shall explain these relations in detail, and in doing so I hope to throw new light on the role of connection theory in relation to the second variation equation in Finsler geometry.

2 Second-order differential equation fields

I shall begin by giving a brief resumé of the geometrical theory of systems of second-order ordinary differential equations. Fuller details and proofs may be found in several publications, including [6] and [11]; and with a somewhat different emphasis, [7].

The function of this section in relation to the rest of the paper is to describe some concepts — in particular, the dynamical derivative and the Jacobi endomorphism — which will be used in the discussion of the Euler-Lagrange equations of a calculus of variations problem. But since the relevant ideas can be developed for second-order differential equations in general, not just for those of Euler-Lagrange type, it seems appropriate to explain them in this more general context.

Even so, the end-point of application in the calculus of variations has an immediate influence on the choice of manifold in which to set the theory. The extremals of ordinary problems in the calculus of variations for regular Lagrangians are parametrized curves, whose parametrizations are fixed as part of their definition as extremals. It is therefore better to regard them as sections of a fibre bundle over $\mathbb{R}$ rather than simply as 1-dimensional subsets of a manifold (which could be envisaged as the standard fibre of this bundle). Taking this into account, it is clearly appropriate to develop the theory on an $(n+1)$-dimensional manifold M which is a fibre bundle over $\mathbb{R}$, where t, the coordinate on $\mathbb{R}$, will eventually play the role of the parameter on any extremal. One could indeed simply take M to be a product $\mathbb{R} \times N$ for some n-dimensional manifold N; however, it seems preferable not to choose one trivialization of M, but rather to treat all on an equal footing; in other words, to insist that the geometric objects which are defined in the study of calculus of variations problems should be tensorial with respect to t-dependent transformations of the fibre coordinates x^i.

So let $\pi: M \to \mathbb{R}$ be a fibre bundle. I shall denote the tangent bundle to M by $\tau: TM \to M$; and for any point $p \in M$ I shall denote by $V_p(\pi)$ the vector subspace of T_pM consisting of those vectors vertical with respect to π, that is, tangent to the fibre of M, and by $V(\pi)$ the corresponding vector sub-bundle of TM.

Let $\pi_1^0: J^1\pi \to M$ be the first-order jet bundle of $\pi: M \to \mathbb{R}$. A point of $J^1\pi$ can be regarded as a tangent vector to M of the form

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

The fibre of $J^1\pi$ over $p \in M$ is an affine space modelled on $V_p(\pi)$, and we can take the v^i as fibre coordinates on $J^1\pi$. The fibration $\pi_1^0: J^1\pi \to M$ also determines a vertical sub-bundle, this time of $T(J^1\pi)$; I shall denote it by $V(\pi_1^0)$. The quotient of each fibre of $T(J^1\pi)$ by its vertical subspace can be identified with a tangent space to M, so we have the exact sequence of vector bundles over $J^1\pi$

$$0 \longrightarrow V(\pi_1^0) \longrightarrow T(J^1\pi) \longrightarrow \pi_1^{0*}(\tau) \longrightarrow 0,$$

where $\pi_1^{0*}(\tau)$ is the pull-back of the tangent bundle $\tau: TM \to M$ to $J^1\pi$ by $\pi^0 + 1$. Corresponding to this exact sequence of vector bundles there is an exact sequence of

$C^\infty(J^1\pi)$-modules of sections

$$0 \longrightarrow \mathcal{V}(\pi_1^0) \longrightarrow \mathcal{X}(J^1\pi) \longrightarrow \mathcal{X}(\pi_1^0) \longrightarrow 0,$$

where

$\mathcal{X}(J^1\pi)$ denotes the module of vector fields on $J^1\pi$

$\mathcal{V}(\pi_1^0)$ denotes the sub-module of vector fields vertical with respect to π_1^0

$\mathcal{X}(\pi_1^0)$ denotes the module of sections of $\pi_1^{0*}(\tau) \to J^1\pi$, that is, of vector fields along the projection π_1^0.

There is a naturally defined vector field along π_1^0 called the total derivative, which is denoted by $\mathbf{T}$, and is given in coordinates by

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

The sub-bundle $\pi_1^{0*}(V\pi)$ of $\pi_1^{0*}(\tau)$ has a naturally defined complement, which is spanned at each point by $\mathbf{T}$; so $\pi_1^{0*}(\tau)$ is a direct sum of vector bundles. The corresponding direct sum decomposition of its module of sections is written

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

The elements of $\overline{\mathcal{X}}(\pi_1^0)$ are annihilated by dt, while the annihilators of $\mathbf{T}$ are spanned by $\{dx^i - v^i dt\}$, the contact 1-forms. In this section and the next, symbols such as X and Y will generally denote elements of $\overline{\mathcal{X}}(\pi_1^0)$ (more properly, local sections of $\tau^*(TM) \to TM$, though the distinction between local and global sections is not very important here).

The module $\mathcal{V}(\pi_1^0)$ can be identified in a natural way with $\overline{\mathcal{X}}(\pi_1^0)$. I shall denote by X^V the vertical vector field on $J^1\pi$ corresponding to $X \in \overline{\mathcal{X}}(\pi_1^0)$, and call it the vertical lift of X: thus

$$X^V = X^i \frac{\partial}{\partial v^i}, \quad \text{where } X = X^i \frac{\partial}{\partial x^i}.$$

The mapping $X \to X^V$ may be extended to $\mathcal{X}(\pi_1^0)$ by setting $\mathbf{T}^V = 0$. Associated with the vertical lift construction there is a type $(1,1)$ tensor field on $J^1\pi$ called the vertical endomorphism, denoted by S, which is defined as follows: for $\xi \in \mathcal{X}(J^1\pi)$, $S(\xi)$ is the vertical lift to $J^1\pi$ of the projection of ξ into $\mathcal{X}(\pi_1^0)$. The coordinate representation of S is

$$S = \frac{\partial}{\partial v^i} \otimes (dx^i - v^i dt).$$

A second-order differential equation field is a vector field Γ on $J^1\pi$ which projects onto $\mathbf{T}$. The coordinate representation of a second-order differential equation field Γ takes the form

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

where $f^i = f^i(t, x^j, v^j)$. The integral curves of a second-order differential equation field are jets of sections of π, corresponding to sections of π which are solution curves of the equations $\ddot{x}^i = f^i(t, x^k, \dot{x}^k)$.

Any second-order differential equation field determines a horizontal distribution on $J^1\pi$, that is to say, a distribution complementary to $\mathcal{V}(\pi_1^0)$; or equivalently a splitting of the vector bundle exact sequence. The Lie derivative $\mathcal{L}_\Gamma S$ of the vertical endomorphism S by a second-order differential equation field Γ has the following properties:

$$(\mathcal{L}_\Gamma S)^2 = I - \Gamma \otimes dt;$$

$$(\mathcal{L}_\Gamma S)(\xi) = \xi \text{ if and only if } \xi \in \mathcal{V}(\pi_1^0);$$

$$(\mathcal{L}_\Gamma S)(\Gamma) = 0;$$

$$\langle (\mathcal{L}_\Gamma S)(\xi), dt \rangle = 0 \text{ for all } \xi \in \mathcal{X}(J^1\pi).$$

It is a simple consequence of these properties that the type $(1, 1)$ tensor field P_H defined by

$$P_H = \tfrac{1}{2} (I - \mathcal{L}_\Gamma S + \Gamma \otimes dt)$$

is a projection operator whose kernel is precisely $\mathcal{V}(\pi_1^0)$. The horizontal distribution determined by Γ consists of those $\xi \in \mathcal{X}(J^1\pi)$ which satisfy $P_H(\xi) = \xi$; P_H is the corresponding horizontal projector, and $P_V = I - P_H$ the vertical one. The second-order differential equation field which defines the horizontal distribution is itself horizontal, and for a local basis $\{H_a\}$, $a = 0, 1, 2, \ldots, n$, of horizontal vector fields we can take

$$H_0 = \Gamma, \quad H_i = \frac{\partial}{\partial x^i} - \Gamma_i^j \frac{\partial}{\partial v^j} \quad \text{where } \Gamma_i^j = -\tfrac{1}{2} \frac{\partial f^j}{\partial v^i}.$$

The $C^\infty(J^1\pi)$-module $\mathcal{X}(\pi_1^0)$ may be identified with the module of horizontal vector fields on $J^1\pi$ as follows:

$$\mathbf{T} \leftrightarrow H_0 = \Gamma = \mathbf{T}^H$$
$$X \in \overline{\mathcal{X}}(\pi_1^0) \leftrightarrow X^i H_i = X^H, \quad \text{where } X = X^i \frac{\partial}{\partial x^i}.$$

Any vectorfield ξ on $J^1\pi$ determines two elements of $\mathcal{X}(\pi_1^0)$, ξ_H and ξ_V, by

$$P_H(\xi) = (\xi_H)^H, \quad P_V(\xi) = (\xi_V)^V;$$

in fact $\xi_V \in \overline{\mathcal{X}}(\pi_1^0)$. Note that

$$(X^H)_H = X = (X^V)_V; \quad (\xi_H)^H + (\xi_V)^V = \xi.$$

I can now define the dynamical derivative ∇ and the Jacobi endomorphism Φ associated with a second-order differential equation field Γ. They are both operators $\overline{\mathcal{X}}(\pi_1^0) \to \overline{\mathcal{X}}(\pi_1^0)$. For any $X \in \overline{\mathcal{X}}(\pi_1^0)$,

$$\nabla X = [\Gamma, X^H]_H, \quad \Phi(X) = [\Gamma, X^H]_V.$$

It follows from these definitions that both ∇ and Φ are $\mathbb{R}$-linear, while for $f \in C^\infty(J^1\pi)$,

$$\nabla(fX) = f(\nabla X) + (\Gamma f)X, \quad \Phi(fX) = f\Phi(X).$$

Thus ∇ behaves like a covariant derivative in the direction of Γ (though no linear connection is required to define it); while Φ can be thought of as a type $(1,1)$ tensor field on $\overline{\mathcal{X}}(\pi_1^0)$ (that is to say, a $C^\infty(J^1\pi)$-linear map $\overline{\mathcal{X}}(\pi_1^0) \to \overline{\mathcal{X}}(\pi_1^0)$). The dynamical derivative extends in the usual way to an operator on $\overline{\mathcal{X}}(\pi_1^0)$-tensor fields.

The horizontal distribution and the Jacobi endomorphism form part of the data required to solve the equivalence problem for second-order differential equation fields.

It can be shown that, for any second-order differential equation field, $[\Gamma, X^V]_H = -X$, while $[\Gamma, X^V]_V = [\Gamma, X^H]_H$. Thus

$$[\Gamma, X^V] = -X^H + (\nabla X)^V$$
$$[\Gamma, X^H] = (\nabla X)^H + \Phi(X)^V.$$

From these results the reason for calling Φ the Jacobi endomorphism becomes clear. Suppose that ξ is a vector field defined along an integral curve of Γ by Lie transport, so that $\mathcal{L}_\Gamma \xi = 0$; we can assume, with no essential loss of generality, that $\langle \xi, dt \rangle = 0$. Set

$$\xi = X^H + Y^V, \quad X, Y \in \overline{\mathcal{X}}(\pi_1^0).$$

Then $Y = \nabla X$, and

$$\nabla^2 X + \Phi(X) = 0.$$

This equation is a generalization of the Jacobi equation for geodesics (the equation of geodesic deviation).

Let X be a π-vertical vector field on M. There is a unique vector field $\tilde{X}$ on $J^1\pi$ such that $\pi^0_{1*}\tilde{X} = X$ and that the flow generated by $\tilde{X}$ maps the contact module to itself. This lifted vector field is the prolongation of X to $J^1\pi$. It will be useful to have expressions for the vertical and horizontal components of a prolongation; they are given by the following simple formula, not unrelated to the results of the previous paragraph:

$$\tilde{X} = X^H + (\nabla X)^V.$$

In particular, $[\Gamma, \tilde{X}]$ is vertical.

3 Ordinary problems of the calculus of variations

Suppose given a Lagrangian, in other words a smooth function L on $J^1\pi$. For any local section σ of π defined on $[a,b] \subset \mathbb{R}$, let $\dot{\sigma}$ be its 1-jet, the section of $J^1\pi \to \mathbb{R}$ given by

$$\dot{\sigma}(t) = (t, \sigma^i(t), \dot{\sigma}^i(t)), \quad \text{where } \dot{\sigma}^i = \frac{d\sigma^i}{dt}.$$

Set

$$\mathcal{I}(\sigma) = \int_a^b \dot{\sigma}^*(L dt).$$

The calculus of variations problem for the Lagrangian L is the problem of finding those sections σ of π for which $\mathcal{I}(\sigma)$ is stationary with respect to variations of σ, in an appropriate sense. (Some of the material in this section is developed in more detail in [4].)

I shall consider only variations of the somewhat restricted kind described below; they will nevertheless be enough for my purposes.

For a given section σ of π defined on $[a, b]$, let X be a π-vertical vector field on M defined in a neighbourhood of the image of σ, such that $X_{\sigma(a)} = X_{\sigma(b)} = 0$; and let ϕ be the flow generated by X. There is some $\varepsilon > 0$ such that $\phi_u(\sigma(t))$ is defined for all $u \in (-\varepsilon, \varepsilon)$ and $t \in [a, b]$; define the variation $v: [a, b] \times (-\varepsilon, \varepsilon) \to M$ by $v(t, u) = \phi_u(\sigma(t))$. The variation field of v is just the restriction of X to the image of σ, and the varied section σ_u is given by the operation of ϕ_u on σ.

The prolongation $\tilde{X}$ of X to $J^1\pi$ has the property that the action of its flow on the 1-jet of a section σ is the 1-jet of $\phi_u \circ \sigma$. It follows that if, for any 1-form α on $J^1\pi$,

$$\mathcal{I}(u) = \int_a^b \dot{\sigma}_u{}^* \alpha$$

then

$$\frac{d\mathcal{I}}{du}(u) = \int_a^b \dot{\sigma}_u{}^* (\mathcal{L}_{\tilde{X}} \alpha).$$

To recover the calculus of variations problem for the Lagrangian L one may put $\alpha = Ldt$; then setting $d\mathcal{I}/du(0) = 0$ leads to the Euler-Lagrange equations for L. However, it is expedient to take advantage of the fact that $\dot{\sigma}_u$ is a 1-jet. The characteristic property of a 1-jet of a section is that it pulls back the contact forms to zero. Thus the value of $\mathcal{I}(u)$ is unchanged by the addition of any contact 1-form to Ldt. We exploit this freedom by replacing Ldt by the 1-form Θ_L of L given by

$$\Theta_L = Ldt + \frac{\partial L}{\partial v^i}(dx^i - v^i dt) = Ldt + dL \circ S.$$

This 1-form is the Cartan 1-form of L, and $d\Theta_L$ the Cartan 2-form. Note that the Cartan 1-form is semi-basic, that is to say, it vanishes on vertical vector fields.

A Lagrangian is said to be regular if its Cartan 2-form has maximal rank, which means (since the dimension of $J^1\pi$ is odd) that its characteristic distribution is 1-dimensional. The solutions of the variation problem for a regular Lagrangian — its extremals — can be described as follows. There is a unique vector field Γ_L on $J^1\pi$, the Euler-Lagrange field of L, which is characteristic for $d\Theta_L$ and satisfies $\Gamma_L(t) = 1$. This vector field is a second-order differential equation field, so its integral curves are 1-jets of sections of π. The solutions of the variation problem are the segments of those sections of π whose 1-jets are integral curves of Γ_L.

There is a close relation between the horizontal distribution corresponding to the Euler-Lagrange field Γ_L and the algebraic structure of $d\Theta_L$:

$$d\Theta_L(X^V, Y^V) = d\Theta_L(X^H, Y^H) = 0, \quad d\Theta_L(X^V, Y^H) = d\Theta_L(Y^V, X^H).$$

Thus $(X, Y) \mapsto d\Theta_L(X^V, Y^H)$ defines a symmetric type $(0,2)$ tensor field on $\overline{\mathcal{X}}(\pi_1^0)$, and the regularity of L implies that it is non-singular. I shall denote it by g_L. The Euler-Lagrange field is characteristic for the Cartan 2-form, which is (evidently) closed, so $\mathcal{L}_{\Gamma_L} d\Theta_L = 0$; it follows from this, together with the formulas for the brackets of a second-order differential equation field with vertical and horizontal lifts given in the previous section, that

$$\nabla g_L = 0 \quad \text{and} \quad g_L(\Phi(X), Y) = g_L(X, \Phi(Y)).$$

4 The Lagrangian second variation formula

With the notation of the previous section we have

$$\frac{d^2\mathcal{I}}{du^2}(0) = \int_a^b \dot{\gamma}^*({\mathcal{L}_{\tilde{X}}}^2\Theta_L),$$

where γ is an extremal, so that $\dot{\gamma}$ is a segment of an integral curve of Γ_L. To evaluate the integrand it is therefore enough to evaluate ${\mathcal{L}_{\tilde{X}}}^2\Theta_L(\Gamma_L)$. Now

$$\begin{aligned}
{\mathcal{L}_{\tilde{X}}}^2\Theta_L(\Gamma_L) &= \tilde{X}\left(\mathcal{L}_{\tilde{X}}\Theta_L(\Gamma_L)\right) - \mathcal{L}_{\tilde{X}}\Theta_L([\tilde{X}, \Gamma_L]) \\
&= \tilde{X}\left(d\Theta_L(\tilde{X}, \Gamma_L) + \Gamma_L(\Theta_L(\tilde{X}))\right) - d\Theta_L(\tilde{X}, [\tilde{X}, \Gamma_L]) - [\tilde{X}, \Gamma_L]\left(\Theta_L(\tilde{X})\right) \\
&= \Gamma_L\left(\tilde{X}(\Theta_L(\tilde{X}))\right) - d\Theta_L([\Gamma_L, \tilde{X}], \tilde{X}),
\end{aligned}$$

using the fact that Γ_L is a characteristic vector field of $d\Theta_L$. But $[\Gamma_L, \tilde{X}]$ is vertical, and $d\Theta_L$ gives zero when both of its arguments are vertical; so in the term $d\Theta_L([\Gamma_L, \tilde{X}], \tilde{X})$ the second $\tilde{X}$ may be replaced by its horizontal component, which is X^H. Since $\mathcal{L}_{\Gamma_L} d\Theta_L = 0$,

$$d\Theta_L([\Gamma_L, \tilde{X}], X^H) = \Gamma_L\left(d\Theta_L(\tilde{X}, X^H)\right) - d\Theta_L(\tilde{X}, [\Gamma_L, X^H]).$$

Thus

$${\mathcal{L}_{\tilde{X}}}^2\Theta_L(\Gamma_L) = \Gamma_L\left(\tilde{X}(\Theta_L(\tilde{X})) - d\Theta_L(\tilde{X}, X^H)\right) + d\Theta_L(\tilde{X}, [\Gamma_L, X^H]).$$

Consider the first term on the right-hand side, the time-derivative term. Since Θ_L is semi-basic, $\Theta_L(\tilde{X}) = \Theta_L(X^H)$, and so

$$\tilde{X}\left(\Theta_L(\tilde{X})\right) = \tilde{X}\left(\Theta_L(X^H)\right) = \mathcal{L}_{X^H}\Theta_L(\tilde{X}) - d\Theta_L(X^H, \tilde{X});$$

thus

$$\tilde{X}\left(\Theta_L(\tilde{X})\right) - d\Theta_L(\tilde{X}, X^H) = \mathcal{L}_{X^H}\Theta_L(\tilde{X}) = X^H\left(\Theta_L(\tilde{X})\right) - \Theta_L([X^H, \tilde{X}]).$$

Now X^H and $\tilde{X}$ are both π_1^0-projectable, and both project onto X; it follows that $[X^H, \tilde{X}]$ is vertical, and so $\Theta_L([X^H, \tilde{X}]) = 0$. Thus

$$\tilde{X}\left(\Theta_L(\tilde{X})\right) - d\Theta_L(\tilde{X}, X^H) = X^H\left(\Theta_L(\tilde{X})\right).$$

But $X^H\left(\Theta_L(\tilde{X})\right)$ vanishes with X, and so the time-derivative term makes no contribution to the integral since the variation leaves the end-points of the segment fixed. All that remains of the integrand is $d\Theta_L(\tilde{X}, [\Gamma_L, X^H])$. Now

$$\begin{aligned}
&\tilde{X} = X^H + (\nabla X)^V,\\
&[\Gamma_L, X^H] = (\nabla X)^H + \Phi(X)^V,
\end{aligned}$$

and so

$$\begin{aligned}
d\Theta_L(\tilde{X}, [\Gamma_L, X^H]) &= d\Theta_L(X^H + (\nabla X)^V, (\nabla X)^H + \Phi(X)^V)\\
&= g(\nabla X, \nabla X) - g(X, \Phi(X)).
\end{aligned}$$

So finally

$$\frac{d^2\mathcal{I}}{du^2}(0) = \int_a^b \{g_L(\nabla X, \nabla X)_{\dot{\gamma}} - g_L(X, \Phi(X))_{\dot{\gamma}}\}\, dt,$$

which is the second variation formula for the Lagrangian L.

Notice that although in the derivation the variation field X is assumed to be defined in a neighbourhood of the image of the section γ, the integrand in the second variation formula depends only on the values of X along the image of γ.

More generally, for a problem in which the end-points of the extremals are not required to be fixed but may vary over some submanifolds Q_a of $\pi^{-1}(a)$ and Q_b of $\pi^{-1}(b)$, we have

$$\begin{aligned}\tfrac{d^2\mathcal{I}}{du^2}(0) = X^H_{\dot\gamma(b)}\left(\Theta_L(\tilde X)\right) - X^H_{\dot\gamma(a)}\left(\Theta_L(\tilde X)\right)\\ + \textstyle\int_a^b \{g_L(\nabla X, \nabla X)_{\dot\gamma} - g_L(X, \Phi(X))_{\dot\gamma}\}\, dt.\end{aligned}$$

In addition to being a segment of an integral curve of Γ_L the extremal will satisfy the transversality conditions

$$\Theta_L(\tilde X)_{\dot\gamma(a)} = \Theta_L(\tilde X)_{\dot\gamma(b)} = 0$$

for any X such that $X_{\gamma(a)} \in T_{\gamma(a)}Q_a$ and $X_{\gamma(b)} \in T_{\gamma(b)}Q_b$ (in fact since Θ_L is semi-basic the transversality conditions involve only the values of $X_{\gamma(a)}$ and $X_{\gamma(b)}$).

Let Q be any submanifold of $\pi^{-1}(t)$ and let $v \in J^1\pi$ with $\pi^1_0(v) = x \in Q$ satisfy the transversality condition $\Theta_L(\zeta) = 0$ for all $\zeta \in T_vJ^1\pi$ such that $\pi^1_{0*}\zeta \in T_xQ$. For $\xi, \eta \in T_xQ$ set

$$\lambda_v(\xi, \eta) = \xi^H\left(\Theta_L(\tilde Y)\right)$$

where Y is any π-vertical vector field on M such that $Y_x = \eta$. It can be shown that λ_v is well-defined (it does not depend on the choice of Y to extend η, thanks to the transversality condition), and that it defines a symmetric bilinear form on T_xQ. Moreover, in the case in which L is the kinetic energy of a Riemannian manifold, v satisfies the transversality condition if and only if it is normal to Q at x, and then λ_v is the second fundamental form of Q at x corresponding to the normal v.

Thus the second variation formula for a regular Lagrangian L generalises the Riemannian second variation formula for variable end-point as well as for fixed end-point problems.

The formula for the second variation given in terms of coordinates in [9] must of course be essentially equivalent to the one given above when both the integral and the boundary terms are taken into account. However, neither the integrand nor the boundary terms in the formula in [9] is covariant, whereas in my formulation both the integrand and the boundary terms are. The integrand in [9] consists of the second-order terms in the Taylor expansion of L relative to $\dot\gamma$. Hidden in my method is a partial integration which transfers some of this integrand to the boundary terms, with the result that both components become covariant.

5 From Lagrange to Finsler geometry

In apparent contrast to Lagrangian geometry, in the first instance Finsler geometry takes place on the tangent bundle TM of a manifold M, with its zero section removed. A Finsler function F is a function on TM which is positively homogeneous of degree 1 in the fibre coordinates; I shall always use this condition in the form $\Delta(F) = F$, where Δ is the Liouville vector field on TM, the vertical vector field which is the infinitesimal generator of dilations in the fibres. I take M to be $(n+1)$-dimensional, with coordinates x^a, $a = 0, 1, 2, \ldots, n$, and I denote the corresponding fibre coordinates on TM by y^a. Then $\Delta = y^a\partial/\partial y^a$.

The tangent bundle TM is equipped with a vertical endomorphism $\tilde{S}$, which encodes its basic geometrical structure in a way analoguous to the way the canonical 1-form does for the cotangent bundle. A vector field Γ on TM is a second-order differential equation field if $\tilde{S}(\Gamma) = \Delta$.

The Hilbert 1-form Θ_F of the Finsler function F is the 1-form

$$\Theta_F = dF \circ \tilde{S} = \frac{\partial F}{\partial y^a} dx^a.$$

Because of the homogeneity assumption on F, the 2-form $d\Theta_F$ is singular (that is, its rank is at most n). In fact it is easy to see that $\mathcal{L}_\Delta \Theta_F = 0$, whence (noting that Θ_F is semi-basic) $\Delta \lrcorner\, d\Theta_F = 0$.

The energy E of a Finsler function F is the function $E = \frac{1}{2}F^2$. The Finsler function F is regular if the 2-form $d\Theta_E$ is non-singular, where $\Theta_E = dE \circ \tilde{S}$. There is then a unique vector field Γ_E such that

$$\Gamma_E \lrcorner\, d\Theta_E = -dE;$$

it is a second-order differential equation field, and is homogeneous of degree 1 in the fibre coordinates, which is to say that $[\Delta, \Gamma_E] = \Gamma_E$. Moreover, $\Gamma_E(E) = 0$, and therefore $\Gamma_E(F) = 0$, assuming F to be nowhere zero. Now $\Theta_E = F\Theta_F$, so that $d\Theta_E = F\, d\Theta_F + dF \wedge \Theta_F$; so

$$\Gamma_E \lrcorner\, d\Theta_E = F \Gamma_E \lrcorner\, d\Theta_F - \Theta_F(\Gamma_E) dF.$$

But $\Theta_F(\Gamma_E) = \Delta(F) = F$, whence $\Gamma_E \lrcorner\, d\Theta_F = 0$.

Thus the characteristic distribution of $d\Theta_F$ contains both Δ and Γ_E. When F is regular, the characteristic distribution of $d\Theta_F$ is precisely the integrable 2-dimensional distribution spanned by Δ and Γ_E. Each vector field in the characteristic distribution of the form $\Gamma_E + \alpha\Delta$ is a second-order differential equation field, whose base integral curves are the same as those of Γ_E as point sets, but differ from them in parametrization. These base integral curves, as sets, are the unparametrized geodesics of the Finsler function.

I shall denote by $\langle \Gamma_F \rangle$ the family of second-order differential equation fields in the characteristic distribution of $d\Theta_F$. There are various ways of handling this slight embarassment of riches. One is to pass to the quotient under the action of the one-parameter group of fibre dilations on TM (having first excised the zero section), so as to obtain the projectivized tangent bundle, and to note that the characteristic distribution of $d\Theta_F$ passes to the quotient as a line-element field. Another is to fix a parametrization. For example, the integral curves of Γ_E are parametrized in such a way that F is constant along them; they are therefore tangent to the codimension 1 submanifolds $F =$ constant, and interpreting F as a measure of length, we can regard the integral curves of the restriction of Γ_E to the submanifold $F = 1$ as being parametrized by arc length. More generally, any function φ on TM for which $\Delta\varphi \neq 0$ allows us to pick out a second-order differential equation field $\Gamma \in \langle \Gamma_F \rangle$ satisfying $\Gamma(\varphi) = 0$, by taking $\alpha = -\Gamma_E(\varphi)/\Delta(\varphi)$; if φ is homogeneous of degree 1 in the fibre coordinates, so will Γ be. We may further restrict Γ to the submanifold $\varphi = 1$, say.

Let us assume, by restricting to an open submanifold if necessary, that M is fibred over $\mathbb{R}$; for example, on any coordinate patch in M we can consider the fibration given by $\pi\colon (x^0, x^1, x^2, \ldots, x^n) \mapsto x^0$. Then $J^1\pi$ is just the submanifold $y^0 = 1$ of TM. The second-order differential equation field defined on $J^1\pi$ by the above construction with $\varphi = y^0$ is just the Euler-Lagrange field Γ_L of the Lagrangian L given by

$$L(t, x^i, v^i) = F(t, x^i, 1, v^i).$$

To show that this is so, it is helpful to denote the injection $J^1\pi \to TM$ by ι: thus in coordinates, $\iota(t, x^i, v^i) = (x^a, y^a)$ where $x^0 = t$, $y^0 = 1$, $y^i = v^i$. Note that

$$\iota^* \frac{\partial F}{\partial y^i} = \frac{\partial L}{\partial v^i}.$$

From the homogeneity of F we have

$$y^0 \frac{\partial F}{\partial y^0} = F - y^i \frac{\partial F}{\partial y^i},$$

from which it follows that

$$\iota^* \frac{\partial F}{\partial y^0} = L - v^i \frac{\partial L}{\partial v^i}.$$

Thus

$$\iota^*\Theta_F = \left(L - v^i \frac{\partial L}{\partial v^i}\right) dt + \frac{\partial L}{\partial v^i} dx^i = \Theta_L,$$

and so the restriction to $J^1\pi$ of the characteristic second-order differential equation field Γ of $d\Theta_F$ which satisfies $\Gamma(y^0) = 0$ is a characteristic vector field of $d\Theta_L$. But since Γ is a second-order differential equation field (on TM), its restriction to $y^0 = 1$ will satisfy $\Gamma(x^0) = \Gamma(t) = 1$. Thus $\Gamma = \Gamma_L$, as claimed.

Let φ be a function on TM such that $\Delta\varphi = \varphi$. Consider the codimension 1 submanifold $\mathcal{M}_\varphi$ of TM defined by $\varphi = 1$. As we have seen, the jet bundle $J^1\pi$ is an example of such a submanifold: it is defined by $y^0 = 1$. The unit sphere bundle of a Finsler manifold is another: it is defined by $F = 1$, where F is the Finsler function. Let Γ be a second-order differential equation field on TM, which is homogeneous of degree 1 in the fibre coordinates, and is tangent to $\mathcal{M}_\varphi$. I shall now discuss the relation between the horizontal distribution construction associated with Γ considered as a vector field on TM, and the similar construction in which Γ is considered as a vector field on $\mathcal{M}_\varphi$.

I shall denote the $C^\infty(\mathcal{M}_\varphi)$-module of vector fields on $\mathcal{M}_\varphi$ by $\mathcal{X}(\mathcal{M}_\varphi)$. By a type $(1,1)$ tensor field on $\mathcal{M}_\varphi$ I shall mean a module map $\mathcal{X}(\mathcal{M}_\varphi) \to \mathcal{X}(\mathcal{M}_\varphi)$.

On $\mathcal{M}_\varphi$, $\Delta\varphi = 1$, so Δ is transverse to $\mathcal{M}_\varphi$. It follows that $\mathcal{M}_\varphi$ projects onto (an open subset of) M under τ. In fact, for any $\xi \in \mathcal{X}(TM)$, the vector field $\hat{\xi} = \xi - (\xi(\varphi)/\varphi)\Delta$, which is defined for $\varphi \neq 0$, satisfies $\hat{\xi}(\varphi) = 0$, and in particular its restriction to $\mathcal{M}_\varphi$ belongs to $\mathcal{X}(\mathcal{M}_\varphi)$. Furthermore, $\hat{\xi} = 0$ if and only if ξ is a multiple of the vertical vector field Δ. Thus given any vectorfield on M, there is a τ-projectable element of $\mathcal{X}(\mathcal{M}_\varphi)$ which projects onto it.

For any second-order differential equation field Γ, $\mathcal{L}_\Gamma \tilde{S}$ satisfies $(\mathcal{L}_\Gamma \tilde{S})^2 = I$, and it acts as the identity on vertical vectors. Its eigendistribution with eigenvalue -1 is complementary to the vertical, and so defines a horizontal distribution. If in addition Γ is homogeneous of degree 1 in the fibre coordinates, so that $[\Delta, \Gamma] = \Gamma$, then Γ is horizontal.

Under what circumstances is the horizontal distribution defined by Γ tangent to $\mathcal{M}_\varphi$? To answer this question, I consider the 1-form $d\varphi \circ \tilde{P}_H$, where $\tilde{P}_H = \frac{1}{2}(I - \mathcal{L}_\Gamma \tilde{S})$ is the horizontal projector associated with the distribution. Now since by assumption $\Gamma(\varphi) = 0$ and $\Delta(\varphi) = \varphi$,

$$d\varphi \circ \mathcal{L}_\Gamma \tilde{S} = \mathcal{L}_\Gamma (d\varphi \circ \tilde{S}) = \Gamma \lrcorner\, d\Theta_\varphi + d\varphi,$$

where $\Theta_\varphi = d\varphi \circ \tilde{S}$ is the Hilbert 1-form of φ. It follows that

$$d\varphi \circ \tilde{P}_H = -\tfrac{1}{2}\Gamma \lrcorner\, d\Theta_\varphi.$$

Thus the necessary and sufficient condition for the horizontal distribution defined by Γ to be tangent to $\mathcal{M}_\varphi$ is that $\Gamma \lrcorner\, d\Theta_\varphi = 0$.

This condition is satisfied when $\varphi = F$ and $\Gamma \in \langle \Gamma_F \rangle$. When $\varphi = y^0$, on the other hand, $\Theta_\varphi = dx^0$, so $d\Theta_\varphi = 0$ and the condition is certainly satisfied. Thus the condition is satisfied in the cases of interest, and I shall assume that it is satisfied for the remainder of this section.

Consider the restriction of $\mathcal{L}_\Gamma \tilde{S}$ to $\mathcal{M}_\varphi$. For $\xi \in \mathcal{X}(\mathcal{M}_\varphi)$, we may write $\xi = V + H$ where $V \in \mathcal{X}(\mathcal{M}_\varphi)$ is vertical and $H \in \mathcal{X}(\mathcal{M}_\varphi)$ is horizontal. Then

$$(d\varphi \circ \mathcal{L}_\Gamma \tilde{S})(\xi) = (d\varphi \circ \mathcal{L}_\Gamma \tilde{S})(V + H) = V(\varphi) - H(\varphi) = 0.$$

Thus $\mathcal{L}_\Gamma \tilde{S}$ induces by restriction a type $(1,1)$ tensor field on $\mathcal{M}_\varphi$, and the horizontal distribution on $\mathcal{M}_\varphi$ is just the eigendistribution of this tensor field with eigenvalue -1.

This same horizontal distribution on $\mathcal{M}_\varphi$ may be constructed by more intrinsic means. The first step in this construction is to modify $\tilde{S}$ so as to obtain a type $(1,1)$ tensor field on $\mathcal{M}_\varphi$. This may be done by composing $\tilde{S}$ with the projection $\mathcal{X}(TM) \to \mathcal{X}(\mathcal{M}_\varphi)$ by $\xi \mapsto \xi - (\xi(\varphi)/\varphi)\Delta$ defined above. The result is the type $(1,1)$ tensor field S on $\mathcal{M}_\varphi$ (where $\varphi = 1$) given by

$$S = (I - \Delta \otimes d\varphi) \circ \tilde{S} = \tilde{S} - \Delta \otimes \Theta_\varphi.$$

Making use of the homogeneity of Γ and the fact that $\Gamma(\varphi) = 0$ we find that

$$\mathcal{L}_\Gamma S = \mathcal{L}_\Gamma \tilde{S} + \Gamma \otimes \Theta_\varphi - \Delta \otimes (d\varphi \circ \mathcal{L}_\Gamma \tilde{S}).$$

Since we are also restricting the arguments of $\mathcal{L}_\Gamma S$ to elements of $\mathcal{X}(\mathcal{M}_\varphi)$ we can dispense with the last term, because $\mathcal{L}_\Gamma \tilde{S}$ maps $\mathcal{X}(\mathcal{M}_\varphi)$ to itself, as was noted above. Thus

$$\mathcal{L}_\Gamma \tilde{S} = \mathcal{L}_\Gamma S - \Gamma \otimes \Theta_\varphi$$

(as an equality between type $(1,1)$ tensor fields on $\mathcal{M}_\varphi$). It follows that the tensor on the right may also be used to define the horizontal distribution on $\mathcal{M}_\varphi$.

When $\varphi = y^0$, $\Theta_\varphi = dx^0 = dt$, so on $y^0 = 1$

$$S = \frac{\partial}{\partial y^a} \otimes dx^a - \left(\frac{\partial}{\partial y^0} + y^i \frac{\partial}{\partial y^i} \right) \otimes dx^0 = \frac{\partial}{\partial v^i} \otimes (dx^i - v^i dt),$$

and

$$\mathcal{L}_\Gamma \tilde{S} = \mathcal{L}_\Gamma S - \Gamma \otimes dt.$$

So the construction described in Section 2 is a particular case of the one given here.

We have seen how the restriction of a Finsler geometry to $\mathcal{M}_{y^0}$ produces a Lagrange geometry. This may suggest that Finsler geometry is more general than Lagrange geometry. However, it is possible to go the other way as well, and at the local level the two geometries are essentially equivalent. Given a Lagrangian L on $\mathcal{M}_{y^0}$, it is possible to extend it to a Finsler function F, by the following 'homogenization trick':

$$F(x^a, y^a) = y^0 L(x^0, x^i, y^i/y^0).$$

It is clear that F is well-defined for $y^0 \neq 0$, and is homogeneous of degree 1 in the fibre coordinates. This local equivalence of Lagrangian and Finslerian structures will be assumed in the following section.

6 The Finslerian second variation formula

As we have seen, given a Finsler function F on TM the representative of $\langle \Gamma_F \rangle$ which is tangent to the submanifold $y^0 = 1$ is the Euler-Lagrange field Γ_L of the Lagrangian L defined by $L(t, x^i, v^i) = F(t, x^i, 1, v^i)$, with its integral curves parametrized by the 'time' $t = x^0$; while the representative of $\langle \Gamma_F \rangle$ which is tangent to the submanifold $F(x^a, y^a) = 1$ is the geodesic field Γ_E of the energy E associated with the Finsler function F, with its integral curves parametrized by arc length. I want now to discuss the relationship between these two parametrizations in detail, because it is an important step towards deriving a second variation formula in Finsler geometry from the Lagarangian formula obtained earlier.

I shall denote the submanifold $y^0 = 1$ by $\mathcal{M}_t$ (for time parametrization), and the submanifold $F(x^\alpha, y^\alpha) = 1$ by $\mathcal{M}_l$ (for length parametrization).

Reparametrization is represented in TM by the map $\rho\colon \mathcal{M}_t \to \mathcal{M}_l$ defined as follows: for any point $v \in \mathcal{M}_t$, $\rho(v)$ is the point in $T_{\tau(v)}M$ at which the ray through v meets

$\mathcal{M}_l$. Thus $\rho(t, x^i, v^i) = (x^a, y^a)$ where $x^0 = t$, $y^i = y^0 v^i$ and $F(x^a, y^a) = 1$. By the homogeneity of F, $y^0 F(x^a, 1, v^i) = 1$, or $y^0 = L(t, x^i, v^i)^{-1}$. So ρ is given explicitly by

$$\rho(t, x^i, v^i) = (t, x^i, L^{-1}, v^i L^{-1}).$$

I next examine the effect of pullback by ρ on the Hilbert 1-form Θ_F, or more precisely its restriction to $\mathcal{M}_l$ (I shall not distinguish the two notationally, so this must be understood in what follows). (This is of course not the same as the restriction of Θ_F to $\mathcal{M}_t$ which was derived in the previous section, though the result will turn out to be the same.) Now

$$\Theta_F = \frac{\partial F}{\partial y^a} dx^a,$$

and so

$$\rho^* \Theta_F = \rho^* \left(\frac{\partial F}{\partial y^a} \right) dx^a.$$

But F and L are related as follows: $F(x^a, y^a) = y^0 L(x^0, x^i, y^i/y^0)$. Thus

$$\begin{aligned} \frac{\partial F}{\partial y^i}(x^a, y^a) &= \frac{\partial L}{\partial v^i}(x^0, x^i, y^i/y^0) \\ \frac{\partial F}{\partial y^0}(x^a, y^a) &= L(x^0, x^i, y^i/y^0) - \frac{y^i}{y^0} \frac{\partial L}{\partial v^i}(x^0, x^i, y^i/y^0), \end{aligned}$$

whence

$$\rho^* \left(\frac{\partial F}{\partial y^i} \right) = \frac{\partial L}{\partial v^i}, \quad \rho^* \left(\frac{\partial F}{\partial y^0} \right) = L - v^i \frac{\partial L}{\partial v^i}.$$

So

$$\rho^* \Theta_F = \Theta_L.$$

Consider next $\rho_* \Gamma_L$. We have

$$\rho^* \left(\rho_* \Gamma_L \lrcorner\, d\Theta_F \right) = \Gamma_L \lrcorner\, \rho^* d\Theta_F = \Gamma_L \lrcorner\, d\Theta_L = 0.$$

It follows that $\rho_*\Gamma_L \lrcorner\, \Theta_F$ is a multiple of dF. But $\Delta \lrcorner\, (\rho_*\Gamma_L \lrcorner\, d\Theta_F) = 0$, while $\Delta \lrcorner\, dF = \Delta(F) = F = 1$. Thus $\rho_*\Gamma_L \lrcorner\, \Theta_F = 0$, and so $\rho_*\Gamma_L \in \langle \Gamma_F \rangle$; since it is tangent to $\mathcal{M}_l$ it must be a multiple of Γ_E. But $\rho_*\Gamma_L(x^0) = 1$, so

$$\rho_*\Gamma_L = (y^0)^{-1}\Gamma_E.$$

Let γ be a segment of a base integral curve of Γ_L, parametrized by time. Reparametrize it by arc length s: that is to say, find an increasing function $s(t)$ on $[a, b]$ such that the curve $\hat{\gamma} = \gamma \circ s^{-1}$ has unit speed relative to F. Thus (with prime denoting differentiation with respect to s)

$$\dot{s} = F(\hat{\gamma} \circ s, (\hat{\gamma}' \circ s)\dot{s}) = L(\gamma, \dot{\gamma}) = ((\hat{\gamma}')^0 \circ s)^{-1}.$$

It follows that

$$\rho \circ \dot{\gamma} = \hat{\gamma}' \circ s,$$

and that

$$\rho_*\dot{\gamma}_*\frac{\partial}{\partial t} = \dot{s}\hat{\gamma}'_*\frac{\partial}{\partial s} = ((\hat{\gamma}')^0 \circ s)^{-1}\hat{\gamma}'_*\frac{\partial}{\partial s}.$$

Comparing this with $\rho_*\Gamma_L = (y^0)^{-1}\Gamma_E$, we see that $\hat{\gamma}'$ is an integral curve of Γ_E.

I want now consider the effect of reparametrization on the Lagrangian second variation formula

$$\frac{d^2\mathcal{I}}{du^2}(0) = \int_a^b \dot{\gamma}^*(\mathcal{L}_{\tilde{X}}{}^2\Theta_L).$$

I must first consider what happens to the variation field $\tilde{X}$. Recall that X is a π-vertical vector field on M, and $\tilde{X}$ its prolongation to $J^1\pi = \mathcal{M}_t$. Then $\rho_*\tilde{X}$ is tangent to $\mathcal{M}_l$ and is τ-projectable, with $\tau_*\rho_*\tilde{X} = X$. Set $\rho_*\tilde{X} = \hat{X}$. Then since $[\Gamma_L, \tilde{X}]$ is π_1^0-vertical, and $\tau \circ \rho = \pi_1^0$,

$$\tau_*\rho_*[\Gamma_L, \tilde{X}] = 0 = \tau_*[(y^0)^{-1}\Gamma_E, \hat{X}] = \tau_*\left((y^0)^{-1}[\Gamma_E, \hat{X}] - \hat{X}((y^0)^{-1})\Gamma_E\right).$$

That is to say, $[\Gamma_E, \hat{X}]$ differs from a τ-vertical vector field on $\mathcal{M}_l$ by a multiple of Γ_E.

Moreover,

$$\dot{\gamma}^*(\mathcal{L}_{\tilde{X}}{}^2\Theta_L) = \dot{\gamma}^*\rho^*(\mathcal{L}_{\hat{X}}{}^2\Theta_F) = s^*(\hat{\gamma}')^*(\mathcal{L}_{\hat{X}}{}^2\Theta_F),$$

and so

$$\frac{d^2\mathcal{I}}{du^2}(0) = \int_a^b \dot{\gamma}^*(\mathcal{L}_{\tilde{X}}{}^2\Theta_L) = \int_0^l (\hat{\gamma}')^*(\mathcal{L}_{\hat{X}}{}^2\Theta_F),$$

where (without loss of generality) $s(a) = 0$, and $s(b) = l$ (so that l is the length of the segment).

Since $\hat{\gamma}'$ is an integral curve of Γ_E, to compute the integrand it is enough to compute $\mathcal{L}_{\hat{X}}{}^2\Theta_F(\Gamma_E)$, as before. The calculation works in the same way as for the Lagrangian case, with one exception: $[\Gamma_E, \hat{X}]$ is not vertical. However, it differs from a vertical vector field by a multiple of Γ_E, and Γ_E is a characteristic vector field of $d\Theta_F$; thus in the term $d\Theta_F([\Gamma_E, \hat{X}], \hat{X})$ it is still legitimate to replace the second $\hat{X}$ by its horizontal component (calculated now with respect to Γ_E, of course).

Thus

$$\frac{d^2\mathcal{I}}{du^2}(0) = \int_0^l \{g_F(\nabla X, \nabla X)_{\hat{\gamma}'} - g_F(X, \Phi(X))_{\hat{\gamma}'}\}\, ds,$$

where ∇ and Φ here represent the operators associated with Γ_E, not Γ_L; and g_F is derived from $d\Theta_F$ in the same way that g_L is derived from $d\Theta_L$.

This reformulation of the second variation formula is less satisfactory than it may appear: in the first place, the bilinear form g_F is singular — in fact $g_F(\mathbf{T}, \cdot) = 0$; in the second, the variation field X corresponds to variations with constant t, which is not natural in the Finslerian context. Both of these difficulties may be resolved by re-expressing the integrand in terms of the energy metric g_E.

The relationship between g_E and g_F is easily found from consideration of the forms $d\Theta_E$ and $d\Theta_F$. On $\mathcal{M}_l$, $d\Theta_E = d\Theta_F + dF \wedge \Theta_F$. Thus

$$\begin{aligned} g_E(X, Y) = d\Theta_E(X^V, Y^H) &= d\Theta_F(X^V, Y^H) + X^V(F)\Theta_F(Y^H) \\ &= g_F(X, Y) + X^V(F)Y^V(F). \end{aligned}$$

Now $g_F(\mathbf{T}, \cdot) = 0$, and $\mathbf{T}^V(F) = \Delta(F) = F = 1$, so $X^V(F) = g_E(X, \mathbf{T})$. Furthermore, $g_E(\mathbf{T}, \mathbf{T}) = 1$. It follows that

$$g_F(X, Y) = g_E(X^\perp, Y^\perp),$$

where $X^\perp = X - g_E(X, \mathbf{T})\mathbf{T}$ is the component of X perpendicular to the unit vector $\mathbf{T}$ with respect to g_E. Thus

$$g_F(\nabla X, \nabla X) - g_F(X, \Phi(X)) = g_E((\nabla X)^\perp, (\nabla X)^\perp) - g_E(X^\perp, \Phi(X)^\perp).$$

From the formula $[\Gamma_E, X^H] = (\nabla X)^H + \Phi(X)^V$ with $X = \mathbf{T}$ (so $X^H = \Gamma_E$) we see that $\nabla \mathbf{T} = 0$ and $\Phi(\mathbf{T}) = 0$. From the first of these, together with $\nabla g_E = 0$, it follows that

$$(\nabla X)^\perp = \nabla X - g_E(\nabla X, \mathbf{T})\mathbf{T} = \nabla(X^\perp);$$

and from the second, together with the symmetry of Φ, that

$$\Phi(X)^\perp = \Phi(X^\perp).$$

Thus ∇ and Φ both preserve perpendicularity to $\mathbf{T}$, and the second variation formula can be written

$$\frac{d^2\mathcal{I}}{du^2}(0) = \int_0^l \left\{ g_E(\nabla X^\perp, \nabla X^\perp)_{\hat{\gamma}'} - g_E(X^\perp, \Phi(X^\perp))_{\hat{\gamma}'} \right\} ds;$$

or alternatively

$$\frac{d^2 I}{du^2}(0) = \int_0^l \{ g_E(\nabla X, \nabla X)_{\hat{\gamma}'} - g_E(X, \Phi(X))_{\hat{\gamma}'} \} \, ds$$

where X is any vector field along $\hat{\gamma}$ which is perpendicular to $\hat{\gamma}'$ with respect to g_E.

So long as $\hat{\gamma}$ is transverse to the fibres of π, the space of vector fields along it which are perpendicular to it and vanish at the end-points is isomorphic to the space of vector fields along it which are π-vertical and vanish at the end-points.

7 Connection theory

In this section I shall deal only with the length-parametrized geodesic vector field Γ_E, so I shall drop the distinguishing subscript E.

A Finsler connection D has the following torsions:

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

For any Finsler connection, therefore, the following relations hold between ∇ and Φ on the one hand, and D_Γ and the torsions on the other:

$$\begin{aligned}
\nabla X &= \mathrm{D}_\Gamma X - \mathrm{D}_{X^H}\mathbf{T} - \mathcal{A}(\mathbf{T}, X) \\
\Phi(X) &= -\mathcal{R}(\mathbf{T}, X) \\
\nabla X &= \mathrm{D}_\Gamma X - \mathcal{P}(\mathbf{T}, X);
\end{aligned}$$

and in addition

$$\mathrm{D}_{X^V}\mathbf{T} + \mathcal{B}(\mathbf{T}, X) = X.$$

Let D be any Finsler connection whose torsions have the following properties:

$$\mathcal{A} = 0, \qquad \mathcal{B}(\mathbf{T}, \cdot) = 0, \qquad \mathcal{P}(\mathbf{T}, \cdot) = 0.$$

For such a connection,

$$\nabla = \mathrm{D}_\Gamma.$$

Moreover, $\mathrm{D}_{X^H}\mathbf{T} = 0$ (and in particular $\mathrm{D}_\Gamma \mathbf{T} = 0$); and $\mathrm{D}_{X^V}\mathbf{T} = X$. If *curv* stands for the curvature of the connection, then

$$\begin{aligned}
curv(X^H, \Gamma)\mathbf{T} &= \mathrm{D}_{X^H}\mathrm{D}_\Gamma \mathbf{T} - \mathrm{D}_\Gamma \mathrm{D}_{X^H}\mathbf{T} - \mathrm{D}_{[X^H, \Gamma]}\mathbf{T} \\
&= \mathrm{D}_{\Phi(X)^V}\mathbf{T} + \mathrm{D}_{(\nabla X)^H}\mathbf{T} = \Phi(X).
\end{aligned}$$

We may therefore express the second variation formula in terms of such a connection as

$$\frac{d^2\mathcal{I}}{du^2}(0) = \int_0^l \left\{ g(\mathrm{D}_\Gamma X, \mathrm{D}_\Gamma X)_{\hat{\gamma}'} - g(X, curv(X^H, \Gamma)\mathbf{T})_{\hat{\gamma}'} \right\} ds,$$

for any vector field X along $\hat{\gamma}$ perpendicular to it with respect to g (where g is the energy metric), and vanishing at the end-points. In such an expression D and *curv* refer to whichever connection is under consideration — though in fact D_Γ and $curv(X^H,\Gamma)\mathbf{T}$ are always the same, namely ∇ and $\Phi(X)$.

The properties of Finsler connections quoted in the following paragraph may be found in Szilasi [12] (and in [5]).

All of the torsions of the Berwald connection associated with Γ vanish, with the exception of $\mathcal{R}$ ([12] Theorem 1 (O1), (O2) and (O5); note that I always assume that the horizontal projection is the one generated by Γ, so that, in Szilasi's terminology, h is the Barthel endomorphism). For each of the other three significant Finsler connections, namely the Cartan, the Chern-Rund and the Hashigushi connections, $\mathcal{A}=0$ ([12] Theorem 2, (M3); Corollary 4, (ChR3); and Corollary 5, (Hsg5)). In the case of the Chern-Rund connection, $\mathcal{B}=0$ ([12] Corollary 4, (R1)), but $\mathcal{P}$ is not identically zero; in the case of the Hashigushi connection, $\mathcal{P}=0$ ([12] Corollary 5, (Hsg1)) but $\mathcal{B}$ is not identically zero; while for the Cartan connection neither $\mathcal{B}$ nor $\mathcal{P}$ is identically zero. However, for all three of these connections $\mathcal{B}(\mathbf{T},\cdot)=0$ and $\mathcal{P}(\mathbf{T},\cdot)=0$: this follows from the occurrence of the metric-related quantities which Szilasi calls $\mathcal{C}$ and $\mathcal{C}'$ (the Cartan tensors) in the definitions of the covariant differentiation operators of these connections, and the properties of $\mathcal{C}$ and $\mathcal{C}'$ represented as $\mathcal{C}^0=0$ ([12] Lemma 5) and $\mathcal{C}'^0=0$ ([12] Lemma 7). These results are also proved, more directly, in [5].

Thus any of the four standard Finsler connections may be used for the covariant derivative and curvature in the expression above for the second variation formula.

References

[1] L. Auslander. On curvature in Finsler geometry. *Trans. Amer. Math. Soc.* **79** (1955), 378–388.

[2] D. Bao and S.-S. Chern. On a notable connection in Finsler geometry. *Houston J. Math.* **19** (1993), 135–180.

[3] S.-S. Chern. Finsler geometry is just Riemannian geometry without the quadratic restriction. *Notices Amer. Math. Soc.* **43** (1996), 959–963.

[4] M. Crampin. *Jet Bundle Techniques in Analytical Mechanics* (Quaderno no. 47, Consiglio Nazionale dell Ricerche, Italy, 1995).

[5] M. Crampin. Connections of Berwald type. Preprint: Dept. of Applied Mathematics, The Open University (1999).

[6] 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.

[7] P. Foulon. Géométrie des equations différentielles du second ordre. *Ann. Inst. H. Poincaré* **45** (1986), 1–28.

[8] M. Giaquinta and S. Hildebrandt. *Calculus of Variations, Vols I and II* (Springer, 1996).

[9] M. Morse. *Variational Analysis* (Wiley, 1973).

[10] 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.

[11] W. Sarlet, A. Vandecasteele, F. Cantrijn and E. Martínez. Derivations of forms along a map: the framework for time-dependent second-order equations. *Diff. Geom. Appl.* **5** (1994), 329–344.

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