A geometrical framework for the study of non-holonomic Lagrangian systems

W. Sarlet, F. Cantrijn*
Theoretical Mechanics Division
University of Gent
Krijgslaan 281
B–9000 Gent (Belgium)

D.J. Saunders
BBC Production Centre
The Open University
Walton Hall
Milton Keynes, MK7 6BH (UK)

Abstract. A geometrical framework is presented for the treatment of a class of dynamical systems, which are modeled by a system of second-order differential equations, coupled with first-order equations linear in the derivatives. Such problems in particular make their appearance in the study of Lagrangian systems with non-holonomic constraints. Among other things, we discuss the concepts of symmetry and adjoint symmetry for such systems and identify for that purpose an appropriate notion of 'dynamical covariant derivative' and 'Jacobi endomorphism'. The intrinsic tools which are being developed further allow a direct geometrical construction of the dynamics of non-holonomic systems. The vertically rolling disc is chosen as an illustrative example for the newly proposed formalism and results.

*Senior Research Associate at the National Fund for Scientific Research (Belgium)

1 Introduction

Consider a class of mechanical systems with non-holonomic constraints. For a system with n degrees of freedom (coordinates q^A), subject to m (linear) non-holonomic constraints, the general description of these constraints is of the form:

$$A_{aA}(t,q)\dot{q}^A + b_a(t,q) = 0, \qquad a = 1,\dots,m,$$

whereby the regularity condition which is normally assumed is that the matrix (A_{aA}) has maximal rank. As a result, one can in principle solve the constraint relations for m of the velocity coordinates in terms of the $k = n - m$ remaining ones, yielding say

$$\dot{q}^a = B^a_\alpha(t,q)\dot{q}^\alpha + B^a(t,q), \qquad a = 1,\dots,m,$$

where the summation over α runs from 1 to k. We henceforth assume that such an operation has been carried out prior to setting up the dynamical equations. If the unconstrained physical system is derivable from a Lagrangian L, the classical procedure for arriving at the equations of motion is to introduce Lagrange-multipliers λ_a and to consider the system of $n+m$ differential equations for the dynamical variables q^A and the m multipliers λ_a, which consists of the m constraint equations, together with:

$$\begin{aligned} \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^\alpha}\right) - \frac{\partial L}{\partial q^\alpha} &= -\lambda_a\, B^a_\alpha, \qquad \alpha = 1,\dots,k, \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^a}\right) - \frac{\partial L}{\partial q^a} &= \lambda_a, \qquad a = 1,\dots,m. \end{aligned}$$

Due to the special way in which the constraints are written, it is obvious here that the unknown multipliers can be eliminated from the picture. Solution curves $q^A(t)$ will have to satisfy the equations

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^\alpha}\right) - \frac{\partial L}{\partial q^\alpha} + \left[\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^a}\right) - \frac{\partial L}{\partial q^a}\right] B^a_\alpha = 0,$$

together with the constraint equations. Substitution of the constraints and their derivatives in the above displayed equations will produce second-order differential equations for the q^α. To see what they are, it is convenient to introduce the function

$$\overline{L}(t, q^A, \dot{q}^\alpha) \equiv L(t, q^A, \dot{q}^\alpha, B^a_\beta \dot{q}^\beta + B^a).$$

We then have the identities,

$$\begin{aligned} \frac{\partial \overline{L}}{\partial q^\alpha} &= \frac{\partial L}{\partial q^\alpha} + \frac{\partial L}{\partial \dot{q}^a}\left(\frac{\partial B^a_\beta}{\partial q^\alpha}\dot{q}^\beta + \frac{\partial B^a}{\partial q^\alpha}\right), \\ \frac{\partial \overline{L}}{\partial \dot{q}^\alpha} &= \frac{\partial L}{\partial \dot{q}^\alpha} + \frac{\partial L}{\partial \dot{q}^a} B^a_\alpha, \end{aligned}$$

from which it follows, taking the preceding result into account, that along solution curves $q^A(t)$, we will have

$$\frac{d}{dt}\left(\frac{\partial \overline{L}}{\partial \dot{q}^\alpha}\right) - \frac{\partial \overline{L}}{\partial q^\alpha} = B^a_\alpha \frac{\partial L}{\partial q^a} + \frac{\partial L}{\partial \dot{q}^a}\left(\dot{B}^a_\alpha - \frac{\partial}{\partial q^\alpha}(B^a_\beta \dot{q}^\beta + B^a)\right).$$

With the understanding that the derivatives of L appearing in the right-hand side are expressed, via the constraints, as functions of $t, q^A, \dot{q}^\alpha$, these are effectively second-order equations for the q^α, which are generally coupled with the first-order constraint equations. Using the identity,

$$\frac{\partial \overline{L}}{\partial q^a} = \frac{\partial L}{\partial q^a} + \frac{\partial L}{\partial \dot{q}^b}\left(\frac{\partial B^b_\beta}{\partial q^a}\dot{q}^\beta + \frac{\partial B^b}{\partial q^a}\right),$$

they can finally be recast into the form:

$$\frac{d}{dt}\left(\frac{\partial \overline{L}}{\partial \dot{q}^\alpha}\right) = X_\alpha(\overline{L}) + C^a_\alpha \frac{\partial L}{\partial \dot{q}^a},$$

where we have introduced, for shorthand, the notations

$$\begin{aligned} X_\alpha &= \frac{\partial}{\partial q^\alpha} + B^a_\alpha \frac{\partial}{\partial q^a}, \\ C^a_\alpha &= \dot{B}^a_\alpha - X_\alpha(B^a_\beta \dot{q}^\beta + B^a). \end{aligned}$$

Assuming that the reduced Lagrangian $\overline{L}$ is regular, the equations for the q^α can be put in normal form, arriving this way at a coupled system of second and first order differential equations of the form

$$\begin{aligned} \ddot{q}^\alpha &= f^\alpha(t, q^A, \dot{q}^\beta), \qquad \alpha = 1, \ldots, k \ (= n - m), \\ \dot{q}^a &= B^a_\alpha(t, q^A)\, \dot{q}^\alpha + B^a(t, q^A), \qquad a = 1, \ldots, m. \end{aligned}$$

Recently, the study of non-holonomic mechanics has again become a field of intensive research. Among the many interesting contributions which focus on the Lagrangian description, we mention papers by Vershik [18], Giachetta [5], Massa and Pagani [11], Koiller [8], Yang [20], Yang et al [21], Cariñena and Rañada [3] and refer in particular to [8, 11] for a long list of references to relevant classical and modern treatments of the subject.

Motivated by this renewed interest, the purpose of the present paper is to describe some geometrical aspects of general systems of mixed differential equations of the above type. Irrespective of their origin, the first order equations in such a system can somehow be regarded as imposing constraints on the accompanying second-order equations. One of the main features we thereby wish to emphasize is that these constraints have a natural interpretation in terms of an Ehresmann connection on an appropriate bundle. This idea is to a certain extent also present in a few of the above cited papers. In fact, while finalising this manuscript, we have become aware of a preprint by Bloch et al [1], where this connection is also at the heart of the matter, but the discussion is restricted to time-independent, non-holonomic Lagrangian dynamics and the emphasis is on topics which are rather different from the ones under investigation here.

In Section 2, we introduce the geometrical framework for the kind of mixed systems of differential equations presented above. It is shown that there are two connections which are naturally associated with such systems: one is needed to define the kind of constraint manifold where the dynamics is supposed to live, the other one comes for free whenever a second-order system on this manifold is being considered. In Section 3, some features of the curvature of both connections are discussed and we introduce the dynamical covariant derivative and Jacobi endomorphism. These two concepts play a fundamental role in the characterisation of symmetries and adjoint symmetries of the given dynamical system in Section 4. In Section 5, we go back to the special case of non-holonomic Lagrangian systems and succeed in defining intrinsically the dynamics associated to such a system, directly on the constraint manifold and without recourse to the celebrated principle of d'Alembert. An illustrative example in the next section is finally followed by some concluding remarks and an outlook for future studies in preparation.

2 A geometrical framework

The structure of the above described types of differential equations strongly suggests that for an intrinsic description, we ought to look at a space where there are two lots of coordinates, q^α and q^a, and that admissible coordinate transformations should respect this distinction, meaning that new Q^a coordinates can possibly depend on all q^A, but the Q^α can be functions of the q^β only. This means that we are thinking of a fibration or, more precisely, of a bundle (i.e. a locally trivial fibred manifold).

Disregarding for a while the special role which is accorded to the variable t, consider a bundle $\pi : E \to M$; let x^α denote coordinates on M and (x^α, y^a) corresponding coordinates on E. A connection on π is a section of $J^1\pi$ over E (see e.g. [17, 10, 19])

$$\tilde{\sigma} : E \to J^1\pi, \quad (x^\alpha, y^a) \mapsto (x^\alpha, y^a, y^a_\alpha = B^a_\alpha(x, y)).$$

Alternatively, such a connection can be thought of as a splitting of the sequence (see e.g. [6])

$$0 \to VE \to TE \to \pi^*TM \to 0,$$

i.e. as a map

$$\sigma : \pi^*TM \to TE, \quad (x^\alpha, y^a, v^\alpha) \mapsto (x^\alpha, y^a, v^\alpha, w^a = B^a_\alpha v^\alpha).$$

The image of σ defines a vector subbundle of TE, which is isomorphic to π^*TM. Starting from a curve $t \mapsto (x^\alpha(t), y^a(t))$ in E, there is an associated curve $t \mapsto (x^\alpha(t), y^a(t), \dot{x}^\alpha(t))$ in π^*TM and a lifted curve $t \mapsto (x^\alpha(t), y^a(t), \dot{x}^\alpha(t), \dot{y}^a(t))$ in TE. It makes perfectly sense to consider dynamical systems, which are constrained by the condition that solution curves must be curves in E, whose lift to TE coincides with the image under σ of the corresponding curve in π^*TM. Such a system will thus be defined by some vector field on $\sigma(\pi^*TM)$ and its solutions will satisfy, apart from some dynamical equations, the differential equations $\dot{y}^a = B^a_\alpha(x, y)\dot{x}^\alpha$.

Assume now, in addition, that both E and M are fibred over $\mathbb{R}$, with projections, say, $\tau_1 : E \to \mathbb{R}$ and $\tau_0 : M \to \mathbb{R}$, satisfying $\tau_1 = \tau_0 \circ \pi$. $J^1\tau_1$ and $J^1\tau_0$ can be naturally identified with submanifolds of TE and TM, respectively, and the restriction of the tangent map $T\pi$ provides (the essential part of) a projection of $J^1\tau_1$ onto $\pi^*J^1\tau_0$. Starting from a connection $\tilde{\sigma}$ as before, the restriction to $\pi^*J^1\tau_0$ of what was called precedingly the corresponding map σ, defines a section of $J^1\tau_1$ over $\pi^*J^1\tau_0$. The image of σ is what we call the constraint manifold J^1_σ; it is an affine subbundle of $J^1\tau_1$, which is isomorphic to $\pi^*J^1\tau_0$:

$$J^1_\sigma = \left\{ j^1_t\phi \in J^1\tau_1 \mid j^1_t\phi = \sigma\left(\phi(t), T\pi(\dot{\phi}(t))\right)\right\}.$$

For a schematic view of the situation, see the diagram in Section 3.

In this special situation, we write the x^α coordinates as (t, q^α) and, to make the link with the previous section, q^a instead of y^a. With regard to the above discussion of special curves in E, it is obvious in the present case that they will have to be sections of τ_1. The different fibrations involved in this picture impose certain 'naturality conditions' on atlases which are appropriate to define the differentiable stucture on each of the manifolds under consideration. In other words, admissible coordinate transformations on E will have to be of the form:

$$\begin{aligned} T &= t, \\ Q^\alpha &= Q^\alpha(t, q^\beta), \\ Q^a &= Q^a(t, q^\alpha, q^b). \end{aligned}$$

For the induced transformations on $\pi^*J^1\tau_0$ and $J^1\tau_1$, they are supplemented by, respectively

$$\dot{Q}^\alpha = \frac{\partial Q^\alpha}{\partial q^\beta}\,\dot{q}^\beta + \frac{\partial Q^\alpha}{\partial t},$$

and

$$\dot{Q}^a = \frac{\partial Q^a}{\partial q^b}\,\dot{q}^b + \frac{\partial Q^a}{\partial q^\alpha}\,\dot{q}^\alpha + \frac{\partial Q^a}{\partial t}.$$

The connection σ gives rise to the following local basis of vector fields, which at each point span the horizontal subspace of TE,

$$X_t = \frac{\partial}{\partial t} + B^a \frac{\partial}{\partial q^a}, \quad X_\alpha = \frac{\partial}{\partial q^\alpha} + B^a_\alpha \frac{\partial}{\partial q^a}.$$

Together with $\partial/\partial q^a$, they form a local basis for $\mathcal{X}(E)$, the dual basis of which is given by

$$dt, \quad dq^\alpha, \quad \eta^a = dq^a - B^a_\alpha \, dq^\alpha - B^a \, dt.$$

The transformation rule for the connection coefficients (under admissible coordinate transformations) reads,

$$\mathcal{B}^a_\alpha = \frac{\partial q^\beta}{\partial Q^\alpha} X_\beta(Q^a), \quad \mathcal{B}^a = X_t(Q^a) + \frac{\partial q^\alpha}{\partial t} X_\alpha(Q^a).$$

These, together with the natural coordinate transformations on $\pi^* J^1\tau_0$, are the relations one needs when it comes to check whether certain objects living on J^1_σ are tensorial.

Let i denote the injection of J^1_σ into $J^1\tau_1$ and ρ the projection of J^1_σ onto E. A local basis of 1-forms on $J^1\tau_1$ consists of course of dt, the contact forms $\theta^\alpha = dq^\alpha - \dot q^\alpha dt$, $\theta^a = dq^a - \dot q^a dt$ and $d\dot q^\alpha$, $d\dot q^a$. Unlike the restricted contact forms $i^*\theta^\alpha$, the 1-forms $i^*\theta^a = dq^a - (B^a_\alpha \dot q^\alpha + B^a) dt$ do not properly behave with respect to the natural coordinate transformations on J^1_σ. A more suitable local basis of 1-forms on J^1_σ therefore is given by,

$$dt, \quad i^*\theta^\alpha, \quad \rho^*\eta^a, \quad \text{and} \quad d\dot q^\alpha.$$

Note that the $\rho^*\eta^a$, which do have proper tensorial behaviour, are related to the restricted contact forms by: $\rho^*\eta^a = i^*\theta^a - B^a_\alpha(i^*\theta^\alpha)$. For simplicity of notation, we will also write θ^α and η^a for the corresponding forms on J^1_σ.

A second-order differential equation field (SODE) on J^1_σ, which by the nature of the present formalism will in fact partly correspond to first-order equations (the constraints), is defined here as follows.

Definition. *A* SODE *on J^1_σ is a vector field $\Gamma \in \mathcal{X}(J^1_\sigma)$, satisfying the requirements*

$$\langle \Gamma, dt \rangle = 1, \quad \langle \Gamma, \theta^\alpha \rangle = 0, \quad , \langle \Gamma, \eta^a \rangle = 0.$$

In coordinates, Γ is of the form

$$\Gamma = \frac{\partial}{\partial t} + \dot q^\alpha \frac{\partial}{\partial q^\alpha} + (B^a_\beta \dot q^\beta + B^a) \frac{\partial}{\partial q^a} + f^\alpha(t, q^A, \dot q^\beta) \frac{\partial}{\partial \dot q^\alpha}.$$

Note in passing that a more general system, with constraint equations of the form $\dot q^a = g^a(t, q^A, \dot q^\alpha)$, would be obtained if we would replace the map $\sigma : \pi^* J^1\tau_0 \to J^1\tau_1$ by an arbitrary section of this fibration (not coming from a connection on π).

As soon as a SODE Γ is given, it looks advantageous to let it be part of an adapted local basis, so for the time being we choose

$$dt, \quad i^*\theta^\alpha, \quad \rho^*\eta^a, \quad \omega^\alpha = d\dot{q}^\alpha - f^\alpha dt$$

as basis of 1-forms on J^1_σ, with dual basis

$$\Gamma, \quad X_\alpha, \quad \frac{\partial}{\partial q^a}, \quad \frac{\partial}{\partial \dot{q}^\alpha}.$$

The canonical vertical endomorphism

$$S = \frac{\partial}{\partial \dot{q}^\alpha} \otimes \theta^\alpha$$

of $J^1\tau_0$ is inherited by $\pi^*J^1\tau_0$ and also carries over to J^1_σ through the isomorphism σ. Computing $\mathcal{L}_\Gamma S$, we find

$$\mathcal{L}_\Gamma S = -X_\alpha \otimes \theta^\alpha - \frac{\partial f^\beta}{\partial \dot{q}^\alpha}\frac{\partial}{\partial \dot{q}^\beta} \otimes \theta^\alpha + \frac{\partial}{\partial \dot{q}^\alpha} \otimes \omega^\alpha,$$

form which it follows that

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

where I is the identity tensor on J^1_σ and

$$N = \frac{\partial}{\partial q^a} \otimes \eta^a$$

is another canonically defined tensor field on J^1_σ. As a matter of fact, N is essentially the vertical projector on E, determined by the connection σ, but appears to behave tensorially also on J^1_σ. We will elaborate on this point in the next section. Introducing the tensor fields

$$\begin{aligned} P_H &= \frac{1}{2}(I - \mathcal{L}_\Gamma S + \Gamma \otimes dt + N), \\ P_V &= \frac{1}{2}(I + \mathcal{L}_\Gamma S - \Gamma \otimes dt - N), \end{aligned}$$

it is easy to verify that $P_H^2 = P_H$, $P_V^2 = P_V$ and $P_H \circ P_V = P_V \circ P_H = 0$. This means, of course, that we have identified for each given SODE Γ a corresponding connection on the bundle $\rho : J^1_\sigma \to E$. The connection coefficients are easily identified by looking at P_H in the coordinate basis; we find

$$P_H = dt \otimes \left(\frac{\partial}{\partial t} + (f^\alpha + \dot{q}^\beta \Gamma^\alpha_\beta)\frac{\partial}{\partial \dot{q}^\alpha}\right) + dq^\alpha \otimes \left(\frac{\partial}{\partial q^\alpha} - \Gamma^\beta_\alpha \frac{\partial}{\partial \dot{q}^\beta}\right) + dq^a \otimes \frac{\partial}{\partial q^a},$$

where, as is the convention in the standard theory of second-order systems, we have put $\Gamma^\alpha_\beta = -(1/2)(\partial f^\alpha/\partial\dot q^\beta)$.

Since we have two tensors S and N at our disposal, it is of interest to investigate some of their further properties. First of all, we know that $S^2 = 0,\ N^2 = N$, and is is easy to see that

$$S\circ N = N\circ S = 0, \quad \text{and} \quad \mathcal{L}_\Gamma S\circ N = N\circ \mathcal{L}_\Gamma S = 0,$$

from which it further follows that also $S\circ\mathcal{L}_\Gamma N = \mathcal{L}_\Gamma N\circ S = 0$. We have,

$$\mathcal{L}_\Gamma N = -C^a_\alpha \frac{\partial}{\partial q^a}\otimes\theta^\alpha - \frac{\partial f^\alpha}{\partial q^a}\frac{\partial}{\partial \dot q^\alpha}\otimes\eta^a,$$

where, in principle, the first term of C^a_α, as defined in the introduction, should be read here as $\Gamma(B^a_\alpha)$, but does clearly not depend on the choice of the Sode Γ. One can further verify that

$$\mathcal{L}_\Gamma S\circ\mathcal{L}_\Gamma N = \mathcal{L}_\Gamma N\circ N, \quad \text{and} \quad \mathcal{L}_\Gamma N\circ\mathcal{L}_\Gamma S = -N\circ\mathcal{L}_\Gamma N.$$

Remark. Consider the tensor

$$U = \mathcal{L}_\Gamma S + \mathcal{L}_\Gamma N\circ N = \mathcal{L}_\Gamma S\circ(I+\mathcal{L}_\Gamma N).$$

Using the above properties, one easily obtains the identity

$$U^2 = I - \Gamma\otimes dt - N + \mathcal{L}_\Gamma N\circ N.$$

It follows that we can define two other, mutually orthogonal projectors, namely

$$P_{H_U} = P_H - \frac{1}{2}\mathcal{L}_\Gamma N\circ N, \quad P_{V_U} = P_V + \frac{1}{2}\mathcal{L}_\Gamma N\circ N.$$

In the coordinate basis we have,

$$\begin{aligned} P_{H_U} \;&=\; dt\otimes\left(\frac{\partial}{\partial t} + \left(f^\alpha + \dot q^\beta\Gamma^\alpha_\beta - \frac{1}{2}\frac{\partial f^\alpha}{\partial q^a}B^a\right)\frac{\partial}{\partial\dot q^\alpha}\right) \\ &\quad + dq^\alpha\otimes\left(\frac{\partial}{\partial q^\alpha} - \left(\Gamma^\beta_\alpha + \frac{1}{2}\frac{\partial f^\beta}{\partial q^a}B^a_\alpha\right)\frac{\partial}{\partial\dot q^\beta}\right) + dq^a\otimes\left(\frac{\partial}{\partial q^a} + \frac{1}{2}\frac{\partial f^\alpha}{\partial q^a}\frac{\partial}{\partial\dot q^\alpha}\right), \end{aligned}$$

which exhibits the alternative connection coefficients. We will not make use of this alternative Sode connection in what follows.

3 Curvature, the dynamical covariant derivative and the Jacobi endomorphism

Two essentially different Ehresmann connections have entered our discussion by now: the first one, $\tilde{\sigma}$ or σ, is needed to define the constraint manifold J^1_σ, the second one comes along with any second-order dynamics living on J^1_σ. As usual the curvature of these connections will play a decisive role somewhere and one of the features we want to highlight in this section is the way these two different curvatures jointly appear in the 'calculus along the projection $\rho : J^1_\sigma \to E$'. It is not our intention here to develop such a calculus in its full extent, the way it has been done for ordinary second-order dynamics in [12, 13, 15]. Experience with interesting applications of such a calculus (see e.g. [14, 4, 2]) has shown that the most important operations related to second-order dynamics are a dynamical covariant derivative and a type (1,1) tensor field, called the Jacobi endomorphism. We will limit ourselves in the present context to precisely these two aspects.

For a start, there is a canonical vector field $\mathbf{T}$ along ρ, defined in coordinates by

$$\mathbf{T} = \frac{\partial}{\partial t} + \dot{q}^\alpha \frac{\partial}{\partial q^\alpha} + (B^a_\alpha \dot{q}^\alpha + B^a)\frac{\partial}{\partial q^a}.$$

As usual, the horizontal lift operation corresponding to some connection on ρ extends to $\mathcal{X}(\rho)$, the set of vector fields along ρ, and we have here

$$\mathbf{T}^H = \Gamma, \quad {X_\alpha}^H = X_\alpha - \Gamma^\beta_\alpha \frac{\partial}{\partial \dot{q}^\beta}, \quad {\frac{\partial}{\partial q^a}}^H = \frac{\partial}{\partial q^a}.$$

In addition, there is a vertical lift operation from $\mathcal{X}(\rho)$ to $\mathcal{X}(J^1_\sigma)$, which is roughly determined by the tensor field S and can be characterized by the properties:

$$\mathbf{T}^V = 0, \quad {X_\alpha}^V = \partial/\partial \dot{q}^\alpha, \quad (\partial/\partial q^a)^V = 0.$$

It will be appropriate to introduce the following submodules of $\mathcal{X}(\rho)$:

$$\begin{aligned} \overline{\mathcal{X}}(\rho) &= \{\overline{X} \in \mathcal{X}(\rho) |\ \langle \overline{X}, dt\rangle = 0,\ N(\overline{X}) = 0\} \\ \widetilde{\mathcal{X}}(\rho) &= \{\tilde{X} \in \mathcal{X}(\rho) |\ N(\tilde{X}) = \tilde{X}\}. \end{aligned}$$

Observe hereby that, as can be done for every tensor field living on E, the vertical projector N is here regarded as a tensor field along the projection ρ.

Every vector field $Z \in \mathcal{X}(J^1_\sigma)$ has a unique decomposition in the form

$$Z = X^H + \overline{Y}^V.$$

Indeed, starting from a given Z, X is defined by $X = T\rho \circ Z$, while $\overline{Y}$ subsequently is the unique element of $\overline{\mathcal{X}}(\rho)$ such that $Z - X^{H} = \overline{Y}^{V}$. The element $X \in \mathcal{X}(\rho)$ itself can be further decomposed as

$$X = \langle X, dt\rangle \mathbf{T} + \overline{X} + \tilde{X}.$$

Dually, we can write every 1-form along ρ as a sum of three terms:

$$\alpha = \langle \mathbf{T}, \alpha\rangle dt + \overline{\alpha} + \tilde{\alpha},$$

with $\overline{\alpha} = \langle X_\beta, \alpha\rangle \theta^\beta$, $\tilde{\alpha} = \langle \partial/\partial q^a, \alpha\rangle \eta^a$. There are dual horizontal and vertical lift operations from $\bigwedge(\rho)$ to $\bigwedge(J^1_\sigma)$.

Turning now to aspects of curvature, let us first look at the connection $\tilde{\sigma}$ (see the diagram below) and recall that one way of defining its curvature is via the Nijenhuis tensor of the horizontal projector on E [9]. One easily verifies, however, that this vector-valued 2-form on E is in fact identical to the Nijenhuis tensor of the vertical projector N. Since N will be shown to be well-defined on J^1_σ as well, we will base our considerations on its Nijenhuis tensor and write N_0, respectively N_1, if we want to make a notational distinction between the N living on E and the N living on J^1_σ.

For the Nijenhuis tensor of N_0, we have computed its various components with respect to the local basis of vector fields $\{X_t, X_\alpha, \partial/\partial q^a\}$, adapted to the connection. The result reads:

$$\mathcal{N}_{N_0} = \left[\frac{1}{2}\Big(X_\alpha(B^a_\beta) - X_\beta(B^a_\alpha)\Big)\, dq^\alpha \wedge dq^\beta + \Big(X_t(B^a_\beta) - X_\beta(B^a)\Big)\, dt \wedge dq^\beta\right] \otimes \frac{\partial}{\partial q^a}.$$

Now, as said before, every tensor field on E can be regarded also as a tensor field along the projection ρ and, in this interpretation, its coordinate expression can more appropriately be written with contact forms θ^α replacing the coordinate 1-forms dq^α in the formula above. Computing then the contraction of the Nijenhuis tensor with the canonical vector field along ρ, we define the type (1,1) tensor field Ψ along ρ:

$$\Psi = i_{\mathbf{T}}\mathcal{N}_{N_0} = \Big[\mathbf{T}(B^a_\beta) - X_\beta(B^a_\alpha \dot{q}^\alpha + B^a)\Big]\, \theta^\beta \otimes \frac{\partial}{\partial q^a} \; = C^a_\beta \theta^\beta \otimes \frac{\partial}{\partial q^a}.$$

This way, we effectively have an intrinsic characterization of the tensor components C^a_α, depending solely on the connection σ on $\pi : E \to M$.

We could proceed in the same way for the SODE connection on $\rho : J^1_\sigma \to E$, leading to a curvature tensor which is a vertical-vector-valued 2-form on J^1_σ. Alternatively (cf. [15]) the curvature here can be regarded as a vector-valued 2-form along ρ, which is determined by the vertical part of the Lie bracket of two horizontal lifts.

To be more precise, for any $X, Y \in \mathcal{X}(\rho)$, computing the decomposition of the vector field $[X^H, Y^H]$, one can verify that its vertical part depends tensorially on X and Y and thus can be written in the form $R(X,Y)^V$, where R is a tensor field along ρ. We define the tensor field Φ along ρ by:

$$\Phi = i_{\mathbf{T}}R = -\Big[X_\alpha(f^\beta) + \Gamma^\gamma_\alpha\Gamma^\beta_\gamma + \Gamma(\Gamma^\beta_\alpha)\Big]X_\beta \otimes \theta^\alpha - \frac{\partial f^\beta}{\partial q^a} X_\beta \otimes \eta^a.$$

For reasons which will become clear in the next section, we introduce the following concept.

Definition. *The tensor field $\Phi + \Psi$ is called the Jacobi endomorphism associated to the given vector field Γ.*

Before closing the discussion on curvature components, it is of interest here to make a digression on the reason why N_1 is well-defined as a tensor field on J^1_σ. To this end, observe that there is a natural process of lifting the connection $\tilde{\sigma}$ on $\pi : E \to M$ to a connection $\tilde{\sigma}_1$ on $\pi_1 : \pi^* J^1\tau_0 \to J^1\tau_0$, which is obtained via the following argumentation. First note that given an arbitrary section $h : M \to E$ of π one can define a lifted section $h_1 : J^1\tau_0 \to \pi^*J^1\tau_0 = E \times_M J^1\tau_0$ by

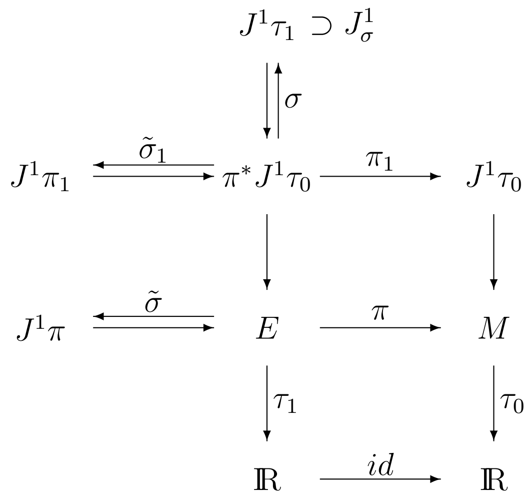
$$h_1(j^1_{t_0}\gamma) = (h(\gamma(t_0)), j^1_{t_0}\gamma),$$

where $\gamma : \mathbb{R} \to M$ is a section of τ_0. It is clear that the definition of h_1 is independent of the choice of γ within the equivalence class of sections determining the point $j^1_{t_0}\gamma \in J^1\tau_0$, and that h_1 takes its values in $\pi^*J^1\tau_0$. Now, given $\tilde{\sigma}$, for any $(m_0, j^1_{t_0}\gamma) \in \pi^*J^1\tau_0$, we can choose a section h of π such that $j^1_{\gamma(t_0)}h = \tilde{\sigma}(m_0)$. In particular, we then have $h(\gamma(t_0)) = m_0$ and hence $h_1(j^1_{t_0}\gamma) = (m_0, j^1_{t_0}\gamma)$. Define then:

$$\tilde{\sigma}_1(m_0, j^1_{t_0}\gamma) = j^1_{j^1_{t_0}\gamma}h_1.$$

Using a coordinate representation one easily verifies that this definition is independent of the choice of γ and h.

The full picture of all spaces of interest is schematically visualized in the diagram below.



In coordinates, the section $\tilde\sigma_1$ is given by

$$q^a_t = B^a(t, q^A), \quad q^a_\alpha = B^a_\alpha(t, q^A), \quad q^a_{\dot\alpha} = 0,$$

from which one can see that the corresponding vertical projector on $\pi^*J^1\tau_0$ reads $N_1 = (\partial/\partial q^a) \otimes \eta^a$ again. Via the isomorphism with J^1_σ, N_1 is also well defined on the constraint manifold. Its Nijenhuis tensor $\mathcal{N}_{N_1}$ formally coincides with the expression of $\mathcal{N}_{N_0}$, regarded as tensor field along ρ, i.e.

$$\mathcal{N}_{N_1} = \left[\frac{1}{2}\Big(X_\alpha(B^a_\beta) - X_\beta(B^a_\alpha)\Big)\,\theta^\alpha \wedge \theta^\beta + C^a_\alpha\, dt \wedge \theta^\beta\right] \otimes \frac{\partial}{\partial q^a}.$$

Hence, the type (1,1) tensor field Ψ, related to the curvature of σ can also be regarded as a tensor field on J^1_σ, namely as

$$\Psi = i_\Gamma \mathcal{N}_{N_1} = C^a_\alpha\, \theta^\alpha \otimes \frac{\partial}{\partial q^a},$$

with the additional remark that this construction turns out to be independent of the choice of a SODE Γ.

There are certainly different paths along which one could arrive at the construction of a suitable dynamical covariant derivative ∇ for the present context. Inspired by previous work, we expect to be able to gather what it should be by calculating the Lie derivative with respect to Γ of horizontal and vertical lifts. For

$$X = \tau\mathbf{T} + \xi^\alpha X_\alpha + \xi^a \frac{\partial}{\partial q^a},$$

we find:

$$\mathcal{L}_\Gamma X^{V} = \Big(\Gamma(\xi^\alpha) + \Gamma^\alpha_\beta \xi^\beta\Big)\frac{\partial}{\partial \dot q^\alpha} - \xi^\alpha {X_\alpha}^{H}.$$

This leads us to define the derivation ∇ of degree 0 in part by

$$\nabla F = \Gamma(F) \;\text{ for } F \in C^\infty(J^1_\sigma), \quad \nabla X_\alpha = \Gamma^{\,\beta}_{\,\alpha} X_\beta.$$

Furthermore, we will require as in [15] that $i_{\nabla X} dt = \Gamma(i_X dt)$ and expect this to tell us that $\nabla \mathbf{T} = 0$. The assumption for this to be true is that $\nabla(\partial/\partial q^a)$ will not have a $\mathbf{T}$-component. In fact, it is natural to expect that it will have no X_α-components either, so that we will recover the same property as in [15]:

$$\mathcal{L}_\Gamma X^{V} = -\overline{X}^{H} + (\nabla X)^{V}.$$

The completion of the definition of ∇ now should follow from computing the decomposition of $\mathcal{L}_\Gamma X^H$. One easily verifies that

$$\begin{aligned}\mathcal{L}_\Gamma X^{H} \;&=\; \Gamma(\tau)\Gamma + \Big(\Gamma(\xi^\alpha) + \Gamma^\beta_\alpha \xi^\alpha\Big){X_\alpha}^{H} + \Big(\Gamma(\xi^a) - \xi^b \frac{\partial}{\partial q^b}(B^a_\beta \dot q^\beta + B^a)\Big)\frac{\partial}{\partial q^a}^{H} \\ &\quad + \Psi(X)^{H} + \Phi(X)^{V}.\end{aligned}$$

This imposes the final defining relation

$$\nabla \frac{\partial}{\partial q^a} = -\frac{\partial}{\partial q^a}(B^b_\beta \dot q^\beta + B^b)\,\frac{\partial}{\partial q^b},$$

so that,

$$\mathcal{L}_\Gamma X^{H} = (\nabla X)^{H} + \Psi(X)^{H} + \Phi(X)^{V}.$$

For completeness, one can check that all defining relations of ∇ thus obtained behave consistently under admissible coordinate transformations on E and on J^1_σ.

Definition. *The derivation ∇ is called the dynamical covariant derivative associated to Γ.*

The dual action of ∇ on 1-forms along ρ is given by,

$$\nabla(dt) = 0, \quad \nabla\theta^\alpha = -\Gamma^{\alpha}_{\,\beta}\,\theta^\beta, \quad \nabla\eta^a = \frac{\partial}{\partial q^b}(B^a_\beta \dot q^\beta + B^a)\,\eta^b.$$

Observe finally that we have $\nabla N = 0$.

4 Symmetries and adjoint symmetries

A (dynamical) symmetry of Γ, as usual, is a vector field $Z \in \mathcal{X}(J^1_\sigma)$ for which $[Z,\Gamma] = h\,\Gamma$ for some function h. As always, we can regard symmetries as being equivalent if they differ by a multiple of Γ and therefore, without loss of generality, restrict our attention to vector fields Z which have no Γ-component and thus satisfy $[Z,\Gamma] = 0$. Writing such a vector field in coordinates as

$$Z = \mu^\alpha X_\alpha + \mu^a \frac{\partial}{\partial q^a} + \nu^\alpha \frac{\partial}{\partial \dot{q}^\alpha},$$

the symmetry requirement tells us that $\nu^\alpha = \Gamma(\mu^\alpha)$ is a first condition, whereas μ^α and μ^a must further satisfy the following coupled system of partial differential equations:

$$\begin{aligned}
\Gamma^2(\mu^\alpha) &= -2\Gamma^\alpha_\beta\,\Gamma(\mu^\beta) + \mu^\beta X_\beta(f^\alpha) + \mu^a \frac{\partial f^\alpha}{\partial q^a}, \\
\Gamma(\mu^a) &= -C^a_\alpha \mu^\alpha + \mu^b \frac{\partial}{\partial q^b}(B^a_\alpha \dot{q}^\alpha + B^a).
\end{aligned}$$

A symmetry Z is clearly completely determined by a vector field along ρ, say of the form $X = \overline{X} + \tilde{X}$, whose horizontal lift produces the horizontal part of Z. It is straightforward to verify, using the results of the preceding section, that the above equations have the following coordinate free formulation:

$$\begin{aligned}
\nabla^2 \overline{X} + \Phi(X) = 0, \\
\nabla \tilde{X} + \Psi(X) = 0.
\end{aligned}$$

For reasons which will become clear in a moment, we can define a symmetry of Γ from now on to be a vector field $X \in \mathcal{X}(\rho)$, which satisfies:

$$\nabla^2 \overline{X} + \nabla \tilde{X} + (\Phi + \Psi)(X) = 0.$$

This single Jacobi-type equation indeed is equivalent to the preceding system of two equations.

Adjoint symmetries of Γ, as known from earlier work [16], can essentially be regarded as so called Γ-basic 1-forms, i.e. forms $\phi \in \bigwedge^1(J^1_\sigma)$ which satisfy $\langle \Gamma, \phi \rangle = 0$ and $\mathcal{L}_\Gamma \phi = 0$. If we write such a form as

$$\phi = a_\alpha \omega^\alpha + b_\alpha \theta^\alpha + c_a \eta^a,$$

it will be invariant under Γ, provided that $b_\alpha = -\Gamma(a_\alpha) + 2a_\beta \Gamma^\beta_\alpha$ and the coefficients a_α and c_a satisfy the equations:

$$\begin{aligned}
\Gamma^2(a_\alpha) &= 2\Gamma(a_\beta \Gamma^\beta_\alpha) + a_\beta X_\alpha(f^\beta) - c_a C^a_\alpha, \\
\Gamma(c_a) &= -a_\alpha \frac{\partial f^\alpha}{\partial q^a} - c_b \frac{\partial}{\partial q^a}(B^b_\alpha \dot{q}^\alpha + B^b).
\end{aligned}$$

This time, ϕ is clearly determined by some element $\alpha = a_\alpha \theta^\alpha + c_a \eta^a \in \bigwedge^1(\rho)$, and all pieces now fit together if we observe that the above partial differential equations in fact express that α satisfies:

$$\nabla^2 \overline{\alpha} - \nabla \tilde{\alpha} + (\Phi + \Psi)(\alpha) = 0.$$

Needless to say, we refer to this as the adjoint equation of the equation for symmetries. Contrary to the latter, it does not split into equations involving only one of the tensors Φ or Ψ.

5 Non-holonomic Lagrangian systems

The idea now is to arrive at a direct geometrical construction of the kind of SODE's Γ, which describe non-holonomic Lagrangian systems, as discussed in the introduction. The point to observe first is that there exists a kind of vertical lift procedure which turns the tensor field Ψ along ρ into a vertical-vector-valued tensor field on $J^1\tau_1$. To see this in a pedestrian way, note first that according to the preceding considerations, the transformation rule for the functions C^a_α, under admissible coordinate transformations, is as follows:

$$C'^b_\beta = C^a_\alpha \frac{\partial q^\alpha}{\partial Q^\beta} \frac{\partial Q^b}{\partial q^a}.$$

On $J^1\tau_1$ on the other hand, under the induced coordinate transformations, contact forms such as θ^α pick up a Jacobian $\partial q^\alpha / \partial Q^\beta$ and $\partial / \partial \dot{q}^a$ transforms into $\partial / \partial \dot{Q}^b$, multiplied by the Jacobian $\partial Q^b / \partial q^a$. It follows that

$$\dot{\Psi} = C^a_\alpha \, \theta^\alpha \otimes \frac{\partial}{\partial \dot{q}^a}$$

is well defined as tensor field on $J^1\tau_1$.

A non-holonomic Lagrangian system, in our approach, is determined by a couple (L, σ), with $L \in C^\infty(J^1\tau_1)$ and σ a connection on π. Putting $\overline{L} = i^* L$ and assuming the Hessian of $\overline{L}$ with respect to the $\dot{q}^\alpha$ to be a regular matrix, we construct the following 1-forms on J^1_σ,

$$\begin{aligned} \theta_{\overline{L}} &= \overline{L} dt + S(d\overline{L}), \\ \psi_{(L,\sigma)} &= i^*\Big(\dot{\Psi}(dL)\Big) - N(d\overline{L}). \end{aligned}$$

Definition. *The fundamental 2-form of a non-holonomic Lagrangian system is the 2-form Ω on J^1_σ, defined by*

$$\Omega = d\theta_{\overline{L}} + \psi_{(L,\sigma)} \wedge dt.$$

Definition. *The dynamics of a non-holonomic Lagrangian system is governed by the* SODE Γ *on* J^1_σ*, which is uniquely determined by the requirement:* $i_\Gamma\Omega = 0$.

In coordinates, we have

$$\psi_{(L,\sigma)} = -\frac{\partial \overline{L}}{\partial q^a}\,\eta^a + \left(i^*\frac{\partial L}{\partial \dot q^a}\right) C^a_\alpha\,\theta^\alpha,$$

and with respect to the frame adapted to Γ, where Γ is to be understood here in the sense that the functions f^α are as yet to be determined, the fundamental 2-form becomes:

$$\begin{aligned}\Omega \;=\; & \left[\Gamma\left(\frac{\partial \overline{L}}{\partial \dot q^\alpha}\right) - X_\alpha(\overline{L}) - C^a_\alpha\left(i^*\frac{\partial L}{\partial \dot q^a}\right)\right] dt\wedge\theta^\alpha \\ & + \left[X_\beta\left(\frac{\partial \overline{L}}{\partial \dot q^\alpha}\right)\theta^\beta + \frac{\partial^2\overline{L}}{\partial q^a\partial \dot q^\alpha}\,\eta^a + \frac{\partial^2\overline{L}}{\partial \dot q^\beta\partial \dot q^\alpha}\,\omega^\beta\right]\wedge\theta^\alpha.\end{aligned}$$

It is clear from this expression that the SODE Γ in the kernel of Ω indeed gives rise to the differential equations for non-holonomic Lagrangian systems, as described in the introduction.

In the classical study of adjoint symmetries of general second-order equations (see [16]), there is a very simple theorem, which encompasses things like Noether's theorem for Lagrangian systems, its generalisation for systems with non-conservative forces and all known statements in the literature about certain symmetries (or 'pseudo-symmetries') which produce so-called alternative Lagrangians. In the most economical picture, where an adjoint symmetry is regarded as a 1-form along the tangent bundle projection (or 1st jet bundle for the time-dependent case) (see [13, 15]), this theorem arises from the case that such a 1-form happens to be the 'vertical exterior derivative' of some function F. In the present situation, where an adjoint symmetry is a 1-form $\alpha \in \bigwedge^1(\rho)$ of the form $\alpha = a_\alpha\theta^\alpha + c_a\eta^a$, we may expect that something special will occur if the coefficients of α happen to be of the form

$$a_\alpha = \frac{\partial F}{\partial \dot q^\alpha},\quad c_a = \frac{\partial F}{\partial q^a}\quad \text{for some function } F.$$

Definition. *A* SODE Γ *on* J^1_σ *is said to be Lagrangian, if the second-order equations for* q^α *can be recast into the form of genuine Euler-Lagrange equations with a Lagrangian not depending on the coordinates* q^a.

Theorem. *Let* $\alpha = a_\alpha\theta^\alpha + c_a\eta^a \;\in \bigwedge^1(\rho)$ *be an adjoint symmetry of* Γ*, with coefficients satisfying the above displayed conditions. Then, the function* $L^* = \Gamma(F)$*, provided it is regular, produces a Lagrangian for the given system.*

Proof. A straightforward calculation shows that, under the given assumptions, the relation for $\Gamma^2(a_\alpha)$ tells us that we have

$$\Gamma\left(\frac{\partial L^*}{\partial \dot q^\alpha}\right) = X_\alpha(L^*),$$

while the relation for $\Gamma(c_a)$ reduces to $\partial L^*/\partial q^a = 0$. The result readily follows. □

Needless to say, under certain further restrictions, we will actually have $L^* = \Gamma(F) = 0$ and this then will cover a generalised Noether theorem for non-holonomic systems.

6 Illustrative example

The paradigm of non-holonomic mechanics is the problem of a vertically rolling disc. Appropriate generalised coordinates are: the coordinates (x, y) of the centre of mass of the disc, the azimuthal angle ψ which determines the position of the plane of the disc and the angle ϕ which describes its internal rotation. If R is the radius of the disc, the condition of rolling without slipping gives rise to non-holonomic constraints of the form

$$\begin{aligned} \dot{x} &= (R\cos\psi)\dot{\phi}, \\ \dot{y} &= (R\sin\psi)\dot{\phi}. \end{aligned}$$

Clearly, we are in a situation here where the dimension of the manifolds M and E is, respectively, 2 and 4 and we can make the following notational identifications: $(q^{\alpha_1}, q^{\alpha_2}, q^{a_1}, q^{a_2}) = (\phi, \psi, x, y)$.

Let us now first follow the procedure for setting up directly the dynamical equations on the constraint manifold. Putting for simplicity the mass of the disc equal to 1, the Lagrangian L on $J^1\tau_1$ is given by (with moments of inertia I_1 and I_2 which need not be specified)

$$L = \frac{1}{2}(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I_1\dot{\phi}^2 + \frac{1}{2}I_2\dot{\psi}^2,$$

and its pull-back to J^1_σ reads,

$$\overline{L} = \frac{1}{2}(R^2 + I_1)\dot{\phi}^2 + \frac{1}{2}I_2\dot{\psi}^2.$$

The terms $X_\alpha(\overline{L})$ in the equations of motion here give 0, whereas the components of the tensor Ψ are given by

$$\begin{aligned} C^{a_1}_{\alpha_1} &= -R\sin\psi\,\dot{\psi}, \qquad & C^{a_1}_{\alpha_2} &= R\sin\psi\,\dot{\phi}, \\ C^{a_2}_{\alpha_1} &= R\cos\psi\,\dot{\psi}, \qquad & C^{a_2}_{\alpha_2} &= -R\cos\psi\,\dot{\phi}. \end{aligned}$$

It is then straightforward to compute that the second-order equations on J^1_σ simply read

$$(R^2 + I_1)\ddot{\phi} = 0, \qquad I_2\ddot{\psi} = 0,$$

i.e. the functions f^α and the connection coefficients Γ^β_α are all zero.

We next try to find an illustration of the theorem on adjoint symmetries. It is easy to verify that the partial differential equations for adjoint symmetries here reduce to

$$\Gamma^2(a_\alpha) = -c_a\, C^a_\alpha, \qquad \Gamma(c_a) = 0.$$

As usual with determining equations for symmetries or adjoint symmetries, one can in principle proceed in a systematic way to construct particular solutions, by gradually allowing polynomials of higher degree in the derivatives. It will be sufficient for our purposes here to discuss just a couple of obvious particular solutions.

Two simple first integrals of the system under consideration are $\dot\phi$ and $\dot\psi$. This leads to four different combinations of particular solutions for the coefficients c_a, but only one of these gives rise to a situation where one subsequently can find by inspection a particular solution for the other coefficients a_α. The favourable case is $c_{a_1} = c_{a_2} = \dot\psi$, with corresponding solution of the second-order conditions: $a_{\alpha_1} = R(\cos\psi - \sin\psi)$, $a_{\alpha_2} = x + y$. The resulting adjoint symmetry happens to satisfy the requirements of the theorem at the end of the preceding section, with

$$F = R(\cos\psi - \sin\psi)\,\dot\phi + (x+y)\,\dot\psi.$$

It turns out that this F is a first integral. In fact, one can see that it is actually the sum of two first integrals which would be obtained along the same lines if we would choose only one of the c_a to be $\dot\psi$ and the other one zero.

The simplest way to satisfy the requirements for the c_a is of course to take them both equal to zero. The other requirements will then be satisfied, for example, if we take the a_α to be first integrals. More interesting particular solutions, however, will be such that $\Gamma^2(a_\alpha) = 0$, but $\Gamma(a_\alpha) \neq 0$. Such a solution is given by $a_{\alpha_1} = \frac{1}{2}\phi$, $a_{\alpha_2} = \frac{1}{2}\psi$. Again, it gives rise to an adjoint symmetry which satisfies the assumptions of the theorem, with

$$F = \frac{1}{2}(\phi\dot\phi + \psi\dot\psi).$$

This time, we find $L^* = \Gamma(F) = \frac{1}{2}(\dot\phi^2 + \dot\psi^2)$ and since L^* satisfies the regularity condition, the theorem tells us (not surprisingly here) that L^* is effectively a Lagrangian for the reduced dynamics on J^1_σ.

7 Concluding remarks

The direct construction of the dynamics of a non-holonomic Lagrangian system, which we described in Section 5, is afterall quite remarkable. All other treatments, to the best of our knowledge, construct the dynamics one way or another via a vector field on the full space $J^1\tau_1$, which is subsequently shown to be tangent to the constraint submanifold. This way, these modern treatments stay close to the classical analysis, which starts from the often misunderstood principle of d'Alembert and involves things like 'virtual velocities'. It is true, of course, that the only justification for our formalism is that it produces differential equations which are known to be the right ones from the other treatments. Nevertheless, turning the arguments around, it looks like an interesting topic for future study to explore more deeply the interrelationship between these different approaches, with the possibility that our present construction may shed new light on the nature of d'Alembert's principle. In the same context, we have to be aware of the fact that our construction depends on the selection of a fibration in the configuration space E. For some physical problems, there is no preferential or 'most natural' choice of such a fibration. It would therefore be of interest to understand whether there is a mechanism by which one can make the transition, at the level of $J^1\tau_1$, from one construction to another.

A second point to be emphasized here is that a large part of our analysis is about a certain class of coupled first and second-order equations and need not have anything to do with non-holonomic Lagrangian mechanics. From this point of view, there is an obvious question for a generalisation to the case where the first-order equations are not restricted in their dependence on the variables $\dot{q}^\alpha$. That such mixed systems of equations may originate from an entirely different context can be seen e.g. in [7]. We have already given a hint about the way such a generalisation can be captured in the present framework in Section 2. Work is in progress along these lines. Since the main motivation for studying such systems need not come from Lagrangian mechanics, we will in this forthcoming contribution pay more attention to various other aspects of the construction of the dynamical covariant derivative and Jacobi endomorphism. Nevertheless, it is worthwhile to add that also in this more general set-up there will be an impact on Lagrangian mechanics: we have indeed good indications that it will again be possible to develop a direct construction of the dynamics of Lagrangian systems with non-linear non-holonomic constraints.

Acknowledgements. This research was partially supported by NATO Collaborative Research Grant (no. CRG 940195). W.S. and F.C. further thank the Belgian National Fund for Scientific Research for continuing support.

References

[1] A.M. Bloch, P.S. Krishnaprasad, J.E. Marsden and R.M. Murray, Nonholonomic mechanical systems with symmetry, preprint (1994).

[2] F. Cantrijn, W. Sarlet, A. Vandecasteele and E. Martínez, Complete separability of time-dependent second-order equations, preprint 1994.

[3] J.F. Cariñena and M.F. Rañada, Lagrangian systems with constraints: a geometric approach to the method of Lagrange multipliers, *J. Phys. A: Math. Gen.* **26** (1993) 1335–1351.

[4] M. Crampin, W. Sarlet, E. Martínez, G.B. Byrnes and G.E. Prince, Towards a geometrical understanding of Douglas's solution of the inverse problem of the calculus of variations, *Inverse Problems* (1994), to appear.

[5] G. Giachetta, Jet methods in nonholonomic mechanics, *J. Math. Phys.* **33** (1992) 1652–1665.

[6] R. Hermann, *Gauge Fields and Cartan-Ehresmann Connections* — Part A, Interdisciplinary Mathematics Vol. X, (Math. Sci. Press, Brookline, 1977).

[7] L.A. Ibort and J. Marín Solano, On the inverse problem of the calculus of variations for a class of coupled dynamical systems, *Inverse Problems* **7** (1991) 713–725.

[8] J. Koiller, Reduction of some classical non-holonomic systems with symmetry, *Arch. Rational Mech. Anal.* **118** (1992) 113–148.

[9] I. Kolář, P.W. Michor and J. Slovák, *Natural Operations in Differential Geometry* (Springer-Verlag, Berlin, 1993).

[10] L. Mangiarotti and M. Modugno, Fibred spaces, jet spaces and connections for field theories, in: Proc. of the International Meeting "*Geometry and Physics*", Florence, 1982 (Pitagora Editrice, Bologna, 1983) 135–165.

[11] E. Massa and E. Pagani, Classical dynamics of non-holonomic systems: a geometric approach, *Ann. Inst. H. Poincaré* **55** (1991) 511–544.

[12] E. Martínez, J.F. Cariñena and W. Sarlet, Derivations of differential forms along the tangent bundle projection, *Diff. Geometry and its Applications* **2** (1992) 17–43.

[13] E. Martínez, J.F. Cariñena and W. Sarlet, Derivations of differential forms along the tangent bundle projection II, *Diff. Geometry and its Applications* **3** (1993) 1–29.

[14] E. Martínez, J.F. Cariñena and W. Sarlet, Geometric characterization of separable second-order equations, *Math. Proc. Camb. Phil. Soc.* **113** (1993) 205–224.

[15] W. Sarlet, A. Vandecasteele, F. Cantrijn and E. Martínez, Derivations of forms along a map: the framework for time-dependent second-order equations, *Diff. Geometry and its Applications*, (1994) to appear.

[16] W. Sarlet, G.E. Prince and M. Crampin, Adjoint symmetries for time-dependent second-order equations, *J. Phys. A: Math. Gen.* **23** (1990) 1335–1347.

[17] D.J. Saunders, *The Geometry of Jet Bundles*, London Math. Soc. Lect. Notes Series **142** (Cambridge University Press, 1989).

[18] A.M. Vershik, Classical and non-classical dynamics with constraints, in: Yu.G. Borisovich and Yu.E. Gliklikh, eds., *Global Analysis — Studies and Applications I*, Lecture Notes in Math. **1108**, (Springer-Verlag, Berlin, 1984) 278–301

[19] A. Vondra, Symmetries and connections on fibered manifolds, *Arch. Math.* (Brno) **30** (**2**) (1994) 97–115.

[20] R. Yang, *Nonholonomic Geometry, Mechanics and Control*, University of Maryland, College Park, Ph.D. Thesis, 1991.

[21] R. Yang, P.S. Krishnaprasad and W.P. Dayawansa, Chaplygin dynamics and Lagrangian reduction, in: W-z Chien, Z-h Guo and Y-z Guo, eds., *Proceedings of the 2nd International Conference on Nonlinear Mechanics*, Beijing, 1993 (Peking University Press, Beijing, 1993) 745–749.