

Symmetry aspects of nonholonomic field theories

Joris Vankerschaver^{1,2} and David Martín de Diego³

¹ Control and Dynamical Systems, California Institute of Technology, MC 107-81, Pasadena, CA 91125, USA

² Department of Mathematical Physics and Astronomy, Ghent University, Krijgslaan 281, B-9000 Ghent, Belgium

³ Instituto de Matemáticas y Física Fundamental, Consejo Superior de Investigaciones Científicas, Serrano 123, 28006 Madrid, Spain

E-mail: jv@caltech.edu and d.martin@imaff.cfmac.csic.es

Received 28 September 2007, in final form 28 November 2007

Published 4 January 2008

Online at stacks.iop.org/JPhysA/41/035401

Abstract

The developments in this paper are concerned with nonholonomic field theories in the presence of symmetries. Having previously treated the case of vertical symmetries, we now deal with the case where the symmetry action can also have a horizontal component. As a first step in this direction, we derive a new and convenient form of the field equations of a nonholonomic field theory. Nonholonomic symmetries are then introduced as symmetry generators whose virtual work is zero along the constraint submanifold, and we show that for every such symmetry, there exists a so-called momentum equation, describing the evolution of the associated component of the momentum map. Keeping up with the underlying geometric philosophy, a small modification of the derivation of the momentum lemma allows us to also treat generalized nonholonomic symmetries, which are vector fields along a projection. Such symmetries arise for example in practical examples of nonholonomic field theories such as the Cosserat rod, for which we recover both energy conservation (a previously known result) and a modified conservation law associated with spatial translations.

PACS numbers: 46.15.Cc, 47.10.-a, 11.10.Ef

Mathematics Subject Classification: 70S05, 37K05, 53D20, 70F25

1. Introduction

Knowledge of symmetry is fundamental to the understanding of mechanical systems and field theories, whether classical or not. Results such as Noether's theorem, the Marsden–Weinstein reduction procedure, etc all attest to this. In the presence of nonholonomic constraints,

however, the concept of symmetry becomes somewhat more subtle: Noether's theorem (see [10] and the references therein) for example no longer holds automatically.

Surprisingly, there is a more sophisticated analogue of Noether's theorem, going by the name of the *nonholonomic momentum equation* [5, 7], which plays a fundamental role in the reduction procedure for general nonholonomic systems. This equation describes the evolution of the components of the momentum map under the nonholonomic flow, and its derivation relies heavily on the use of differential geometry for the formulation of the nonholonomic problem.

In the case of *classical field theories*, a similar result was derived in [22] for a restricted class of symmetries, namely those symmetries whose infinitesimal vector field is vertical, or alternatively, symmetries which act trivially on the space of independent variables of the theory. Nevertheless, many physically interesting symmetries are not vertical, and an extension of this result to a more general class of symmetries is therefore needed: this is the subject of the present paper. Examples of non-vertical symmetries include translation in time, yielding conservation of energy, without doubt the prime example of a conserved quantity.

Outline of this paper

In this paper, we derive a generalized form of the nonholonomic momentum equation for two classes of symmetries. The first consists of symmetries which have a nontrivial horizontal component, while for the second class we restrict our attention to vertical symmetries whose infinitesimal generator is a generalized vector field, i.e. a vector field along the projection $\pi_{1,0}$. More information on the application of generalized symmetries to differential equations can be found in [19].

In section 3 we derive a new form of the field equations for a classical field theory with nonholonomic constraints, which is particularly well suited to the derivation of the momentum equation in section 4. These equations are equivalent to the nonholonomic field equations derived, for example, in [3, 15, 21] but they do not involve the Lagrange multipliers and are therefore closer in spirit to Hölder's equations (see [18]). We derive the nonholonomic momentum equation for two distinct classes of symmetries: firstly, for the case of a symmetry group whose infinitesimal generators are *projectable* vector fields on the total space, and secondly, for a symmetry group acting *vertically* on the total space, but whose associated nonholonomic symmetries are *generalized vector fields*.

This theory is applied in section 5 to a number of examples. In section 5.1 we obtain a new class of nonholonomic symmetries for mechanical systems with nonlinear constraints; Benenti's nonholonomic system (see [2]) is chosen as a straightforward but relevant example and a number of new conservation laws are derived. In section 5.2 we derive nonholonomic conservation laws for the nonholonomic Cosserat rod, an example of a nonholonomic field theory (see [23]). These conservation laws are associated with translations in time and spatial translations, respectively. The former is a horizontal symmetry, while the latter is described by a (vertical) generalized vector field.

Finally, in the appendix, we elaborate somewhat further on the choice of reaction forces for a nonholonomic field theory. As the material in this section is not crucial to the developments in the remainder of the paper, we have relegated it to the appendix.

Notation

Throughout this paper, all geometric objects are assumed to be smooth. We will denote the contraction of a vector field V and a form α both as $i_V\alpha$ or as $V\lrcorner\alpha$, and we will use these

two notations interchangeably. The Lie derivative is denoted by \mathcal{L} . We will frequently use the notation $\bigwedge^k(M)$, where M is a manifold, to denote the k -fold exterior product of TM with itself. Thus, $\bigwedge^k(M)$ is a bundle over M whose sections are k -forms on M .

2. Classical field theories

As is customary in the geometric treatment of classical field theories (see [8, 10, 13, 20] and the references therein), fields are modelled as sections of a fibre bundle $\pi : Y \rightarrow X$ of rank m , where the base space X is an $(n + 1)$ -dimensional oriented manifold (with volume form η), and the total space Y has dimension $n + m + 1$. Local coordinates on X are denoted by (x^μ) , $\mu = 0, \dots, n$ and are supposed to be such that the volume form η can be locally written as

$$\eta = d^{n+1}x := dx_0 \wedge \dots \wedge dx_n.$$

In addition, we assume a local system of bundle coordinates (x^μ, y^a) , $\mu = 0, \dots, n, a = 1, \dots, m$, on Y to be given.

Over Y , there exists a tower of jet bundles

$$\dots \longrightarrow J^2\pi \xrightarrow{\pi_{2,1}} J^1\pi \xrightarrow{\pi_{1,0}} Y.$$

The elements of $J^k\pi$ are equivalence classes of local sections of π , where two sections are said to be equivalent at a point x of X if their k th-order Taylor expansions agree at that point. We denote the equivalence class of a section as $j_x^k\phi$. More information on jet bundles can be found in [20]. We will mostly only need the first-order jet bundle $J^1\pi$; the second-order jet bundle will make a brief appearance in section 5.2 (see for instance [12] for a review of mechanics on higher order jet bundles).

We will need as basic geometric tools the concepts of contact forms and jet prolongation of vector fields. A contact m -form on $J^k\pi$ will be any m -form θ satisfying $(j^k\phi)^*\theta$ for every local section ϕ of π . The set of contact 1-forms on $J^1\pi$ defines a distribution $\mathcal{D}_{\text{Cartan}}$, called the Cartan distribution. Using this distribution, it is easy to define the 1-jet prolongation of a vector field ξ_Y on Y as the unique vector field $j^1\xi_Y$ projectable on ξ_Y and preserving the Cartan distribution, i.e. $\mathcal{L}_{j^1\xi_Y}\mathcal{D}_{\text{Cartan}} \subseteq \mathcal{D}_{\text{Cartan}}$. In coordinates, if $\xi_Y = \xi^\mu(x, y)\frac{\partial}{\partial x^\mu} + \xi^a(x, y)\frac{\partial}{\partial y^a}$, then

$$j^1\xi_Y = \xi^\mu \frac{\partial}{\partial x^\mu} + \xi^a \frac{\partial}{\partial y^a} + \left(\frac{d\xi^a}{dx^\nu} - y_\mu^a \frac{\partial \xi^\mu}{\partial x^\nu} \right) \frac{\partial}{\partial y_\nu^a}, \quad \text{where} \quad \frac{d}{dx^\nu} = \frac{\partial}{\partial x^\nu} + y_\nu^a \frac{\partial}{\partial y^a}.$$

If we adopt the coordinate systems on X and Y as above, then $J^1\pi$ is equipped with an induced coordinate system $(x^\mu, y^a; y_\mu^a)$, $\mu = 0, \dots, n, a = 1, \dots, m$. With respect to this coordinate system, the projection $\pi_{1,0} : J^1\pi \rightarrow Y$ is given by $\pi_{1,0}(x^\mu, y^a; y_\mu^a) = (x^\mu, y^a)$.

For later use, we mention the existence of a distinguished vector-valued $(n + 1)$ -form S_η on $J^1\pi$, which is called the *vertical endomorphism*. In coordinates, S_η reads

$$S_\eta = (dy^a - y_\nu^a dx^\nu) \wedge d^n x_\mu \otimes \frac{\partial}{\partial y_\mu^a}, \tag{1}$$

where

$$d^n x_\mu := \frac{\partial}{\partial x^\mu} \lrcorner d^{n+1}x.$$

For the purpose of this paper, a Lagrangian will be a function L on $J^1\pi$. Given a Lagrangian L , we define the Poincaré–Cartan $(n + 1)$ -form Θ_L as

$$\Theta_L := S_\eta^*(dL) + L\eta = \frac{\partial L}{\partial y_\mu^a} (dy^a - y_\nu^a dx^\nu) \wedge d^n x_\mu + L d^{n+1}x.$$

We also put $\Omega_L := -d\Theta_L$, to which we refer as the Poincaré–Cartan $(n + 2)$ -form.

For more information on the geometry of jet bundles, see [20].

The Euler–Lagrange equations

Let there be given a Lagrangian L . The dynamics of the field theory is described by the Euler–Lagrange equations associated with L ; they express that the field is an extremum of the following action functional:

$$S(\phi) = \int_X L \left(x^\mu, \phi^a(x), \frac{\partial \phi^a}{\partial x^\mu} \right) d^{n+1}x. \tag{2}$$

By varying the action S with respect to a variation V , we obtain after integrating by parts

$$0 = \frac{d}{d\epsilon} S(j^1(\Phi_\epsilon \circ \phi)) \Big|_{\epsilon=0} = \int_U \left(\frac{\partial L}{\partial y^a} - \frac{d}{dx^\mu} \frac{\partial L}{\partial y^a_\mu} \right) V^a d^{n+1}x, \tag{3}$$

where Φ_ϵ is a finite variation associated with V . Here, a *variation* of a field ϕ over an open set $U \subset X$ is a π -vertical vector field V defined on an open neighborhood of $\phi(U)$. The associated finite variation Φ_ϵ is nothing but the flow of V . Note that the composition $\Phi_\epsilon \circ \phi$ is again a local section of π .

In coordinates, a section ϕ of π is an extremum of (2) if and only if it satisfies the familiar Euler–Lagrange equations, given by

$$\frac{\partial}{\partial x^\mu} \left(\frac{\partial L}{\partial y^a_\mu} (j^1\phi) \right) - \frac{\partial L}{\partial y^a} (j^1\phi) = 0.$$

There exist various intrinsic formulations of these equations (see [11] for an overview), of which we mention just one. It can be shown by a straightforward coordinate calculation that the Euler–Lagrange equations are equivalent to the following set of intrinsic equations:

$$(j^1\phi)^*(i_W\Omega_L) = 0 \quad \text{for all} \quad W \in \mathfrak{X}(J^1\pi).$$

In what follows, we will be mostly interested in similar equations for field theories which are subject to nonholonomic constraints. This is the subject of the next section.

3. The nonholonomic field equations

A nonholonomic field theory is given by the specification of three objects (see also [21]):

- (i) a Lagrangian $L : J^1\pi \rightarrow \mathbb{R}$;
- (ii) a constraint submanifold $\mathcal{C} \hookrightarrow J^1\pi$, such that the restriction of the projection $(\pi_{1,0})|_{\mathcal{C}}$ defines a subbundle of $\pi_{1,0} : J^1\pi \rightarrow Y$;
- (iii) a bundle of reaction forces F , where the elements Φ of F are $(n + 1)$ -forms defined along \mathcal{C} , i.e. maps from $\mathcal{C} \subset J^1\pi$ to $\Omega^{n+1}(J^1\pi)$. The elements of F have to satisfy the following requirements:
 - (a) Φ is n -horizontal, i.e. Φ vanishes when contracted with any two π_1 -vertical vector fields;
 - (b) Φ is 1-contact, i.e. $(j^1\phi)^*\Phi = 0$ for any section ϕ of π .

It can be shown that any element Φ of F is of the following form:

$$\Phi = A_a^\mu (dy^a - y^a_\nu dx^\nu) \wedge d^n x_\mu, \tag{4}$$

where A_a^μ are functions on $J^1\pi$.

For the sake of simplicity, we will assume \mathcal{C} to be defined by the vanishing of k functionally independent functions φ^α on $J^1\pi$. Furthermore, we will assume that F is globally generated by l generators Φ^κ of the following form:

$$\Phi^\kappa = A_a^{\kappa\mu} (dy^a - y_v^a dx^v) \wedge d^n x_\mu \quad (\kappa = 1, \dots, l).$$

In practice, the dimension l of F will be equal to the codimension k of \mathcal{C} . There seems to be no *a priori* reason for supposing that $k = l$. In most cases, however, F will be determined by \mathcal{C} through application of the Chetaev principle, described in remark 3.6 below.

In the nonholonomic treatment of constraints, a special role is played by infinitesimal variations of the fields which are compatible with the constraint, as in the following definition.

Definition 3.1. A variation V of a field ϕ (taking values in \mathcal{C} , i.e. such that $j^1\phi \subset \mathcal{C}$) defined over an open subset U with compact closure is admissible if

$$(j^1\phi)^*(j^1V \rfloor \Phi) = 0 \quad \text{for all} \quad \Phi \in F. \tag{5}$$

By varying the action S with respect to an admissible variation V , we obtain after integrating by parts again (3). If the variations V were arbitrary, then this would immediately yield the Euler–Lagrange equations. However, this is not the case as the variations have to be admissible. Additional *reaction forces* will therefore appear in the Euler–Lagrange equations, whose role is to constrain the solution ϕ to the constraint submanifold.

Definition 3.2. A local section ϕ of π , defined on an open subset $U \subset X$ with compact closure, is a solution of the nonholonomic problem determined by L, \mathcal{C} and F if $j^1\phi(U) \subset \mathcal{C}$ and (3) holds for all admissible variations V of ϕ .

It follows from (3) that a local section ϕ is a solution of the nonholonomic problem if it satisfies the *nonholonomic Euler–Lagrange equations*

$$\left[\frac{\partial L}{\partial y^a} - \frac{d}{dx^\mu} \frac{\partial L}{\partial y_\mu^a} \right] (j^2\phi) = \lambda_{\alpha\kappa} A_a^{\alpha\kappa} (j^1\phi) \quad \text{and} \quad \varphi^\alpha (j^1\phi) = 0. \tag{6}$$

Here, $\lambda_{\alpha\kappa}$ are unknown Lagrange multipliers, to be determined from the constraints. An intrinsic form of these equations is derived below in theorem 3.5, but first we need the following technical results.

Lemma 3.3 (lemma 3.2 in [13]). Let W be a vector field on $J^1\pi$. If ϕ is a section of π and if either W is tangent to the image of $j^1\phi$ or if W is $\pi_{1,0}$ -vertical, then $(j^1\phi)^*(i_W \Omega_L) = 0$.

Now, let ϕ be a section such that the image of $j^1\phi$ is a subset of \mathcal{C} and consider a vector field W which is tangent to the image of $j^1\phi$, i.e. there exists a vector field w on X such that $T_x j^1\phi(w(x)) = W(j_x^1\phi)$ for all $x \in X$. One can follow a similar reasoning as in the proof of lemma 3.2 in [13] to show that

$$(j^1\phi)^*(W \rfloor \Phi) = w \rfloor ((j^1\phi)^* \Phi)$$

for any $\Phi \in F$. Since Φ is 1-contact, the right-hand side of this expression vanishes. On the other hand, if W is $\pi_{1,0}$ -vertical, it follows automatically that $(j^1\phi)^*(W \rfloor \Phi) = 0$. We have therefore proved the following lemma.

Lemma 3.4. Let ϕ be a section of π such that $j_x^1\phi \in \mathcal{C}$ for all $x \in U \subset X$. If either W is tangent to the image of $j^1\phi$ or W is $\pi_{1,0}$ -vertical, then $(j^1\phi)^*(W \rfloor \Phi) = 0$ for all $\Phi \in F$.

Henceforth, we shall call any vector field W on $J^1\pi$ *admissible* with respect to a section ϕ of π if $(j^1\phi)^*(W \rfloor \Phi) = 0$ for all $\Phi \in F$.

Theorem 3.5. *Let ϕ be a section of π . If $\text{Im } j^1\phi \subset \mathcal{C}$, then the following assertions are equivalent:*

- (a) ϕ is a stationary point of the action (2) under admissible variations;
- (b) ϕ satisfies the Euler–Lagrange equations (6);
- (c) for all vector fields W on $J^1\pi$ such that $(j^1\phi)^*(W \rfloor \Phi) = 0$ for all $\Phi \in F$,

$$(j^1\phi)^*(W \rfloor \Omega_L) = 0. \tag{7}$$

Proof. Let us first prove the equivalence of (a) and (c). For arbitrary, not necessarily admissible variations, the following result holds (this is equation 3C.5 in [13]):

$$\left. \frac{d}{d\epsilon} S(\phi_\epsilon) \right|_{\epsilon=0} = - \int_U (j^1\phi)^*(j^1V \rfloor \Omega_L).$$

For admissible variations, from hypothesis (a), we have

$$\int_U (j^1\phi)^*(j^1V \rfloor \Omega_L) = 0.$$

Now, we may multiply V by an arbitrary function on X and this result will still hold true. The fundamental lemma of the calculus of variations therefore shows that

$$(j^1\phi)^*(j^1V \rfloor \Omega_L) = 0, \tag{8}$$

for all admissible variations V defined over U . By using a partition of unity as in [13], it can then be shown that (8) holds for all π -vertical vector fields V such that $(j^1V) \rfloor \Phi = 0$ for all $\Phi \in F$. This expression is equivalent to (7): to see this, take an arbitrary vector field W on $J^1\pi$ such that $(j^1\phi)^*(W \rfloor \Phi) = 0$ for all $\Phi \in F$. The vector field W can be decomposed as the following sum (to be considered along the image of $j^1\phi$):

$$W = w_{\parallel} + j^1V + v_{\pi_{1,0}},$$

where w_{\parallel} is tangent to the image of $j^1\phi$, j^1V is the prolongation of a π -vertical vector field V and $v_{\pi_{1,0}}$ is a $\pi_{1,0}$ -vertical vector field. Using lemma 3.4, we have that

$$(j^1\phi)^*(j^1V \rfloor \Phi) = (j^1\phi)^*(W \rfloor \Phi) = 0,$$

and from lemma 3.3, we get $(j^1\phi)^*(W \rfloor \Omega_L) = (j^1\phi)^*(j^1V \rfloor \Omega_L)$. The right-hand side of this equation vanishes since j^1V is admissible, and therefore we conclude that W satisfies (7).

The equivalence of (a) and (c) is just a matter of writing out the definitions. In coordinates, the left-hand side of (7) reads (for a prolongation of a vertical vector field V)

$$(j^1\phi)^*(j^1V \rfloor \Omega_L) = V^a \left(\frac{\partial L}{\partial y^a}(j^1\phi) - \frac{\partial}{\partial x^\mu} \frac{\partial L}{\partial y^\mu_a}(j^1\phi) \right) d^{n+1}x,$$

and this holds for all admissible variations V . Therefore, if ϕ satisfies (7), then there exist functions $\lambda_{\alpha\kappa}$ such that

$$\left[\frac{\partial L}{\partial y^a} - \frac{d}{dx^\mu} \frac{\partial L}{\partial y^\mu_a} \right] (j^2\phi) = \lambda_{\alpha\kappa} A_a^{\alpha\kappa}(j^1\phi).$$

The converse is similar. □

We see from the proof of this theorem that only vertical vector fields yield nontrivial results for (7).

Remark 3.6. The bundle of reaction forces F is commonly derived from the constraint submanifold \mathcal{C} through application of the Chetaev principle (see [18], as well as [21] for an

extension to the case of field theories). If the constraint submanifold is given as the zero-level set of functions φ^α , then according to this principle, F is locally generated by the following forms:

$$\Phi^\alpha := S_\eta^*(d\varphi^\alpha) = \frac{\partial \varphi^\alpha}{\partial y_\mu^a} (dy^a - y_\nu^a dx^\nu) \wedge d^n x_\mu.$$

In the past, there has appeared some criticism over the use of the Chetaev principle (see [16]), and as we shall see in the appendix, for classical field theories the Chetaev principle sometimes has to be modified.

4. The nonholonomic momentum equation

In this section, we derive the nonholonomic momentum equation, the nonholonomic counterpart to the well-known theorem of Noether. More in detail, we prove that for every *nonholonomic symmetry* there exists a certain partial differential equation, which reduces to a conservation law when the constraints are absent. In proving the momentum lemma, many different starting assumptions can be made, and we study two different setups.

- (i) In section 4.2, we assume that the symmetry group G acts by bundle automorphisms on Y . In particular, we allow for the fact that G acts nontrivially on X as well. In this case, nonholonomic symmetries are *projectable* vector fields on Y .
- (ii) In section 4.3, we model nonholonomic symmetries as vector fields along the projection $\pi_{1,0}$. For the sake of simplicity, we assume in this case that the action of G on Y is *vertical*, i.e. G acts trivially on X . This setup is therefore complementary to the one described before, but this case can probably be extended even further.

4.1. The nonholonomic momentum map

Let G be a Lie group acting on Y by bundle automorphisms, i.e. there exist smooth actions $\overline{\Phi} : G \times Y \rightarrow Y$ and $\underline{\Phi} : G \times X \rightarrow X$ such that $\pi(\overline{\Phi}(g, y)) = \underline{\Phi}(g, \pi(y))$ for all $g \in G$ and $y \in Y$. We use the following abbreviations: $\overline{\Phi}_g := \overline{\Phi}(g, \cdot)$ and $\underline{\Phi}_g := \underline{\Phi}(g, \cdot)$. The Lie group G acts on $J^1\pi$ by prolonged bundle automorphisms, i.e.

$$j^1\Phi_g(j_x^1\phi) = j_x^1(\overline{\Phi}_g \circ \phi \circ \underline{\Phi}_g^{-1}).$$

Let us now assume that G leaves invariant L, \mathcal{C} , and F :

$$j^1\Phi_g^*(L\eta) = L\eta, \quad j^1\Phi_g(\mathcal{C}) \subset \mathcal{C} \quad \text{and} \quad (j^1\Phi_g)^*F \subset F$$

for all $g \in G$.

We consider first the vector bundle \mathfrak{g}^F over Y , defined as follows. Denote by $\mathfrak{g}^F(y)$ the linear subspace of \mathfrak{g} consisting of all $\xi \in \mathfrak{g}$ such that

$$j^1\xi_Y(\gamma)\lrcorner F = 0 \quad \text{for all} \quad \gamma \in \mathcal{C} \cap \pi_{1,0}^{-1}(y), \tag{9}$$

where ξ_Y is the infinitesimal generator of the action corresponding to ξ , that is,

$$\xi_Y(y) = \left. \frac{d}{dt} \overline{\Phi}_{\exp(t\xi)}(y) \right|_{t=0}.$$

We assume that the disjoint union of all $\mathfrak{g}^F(y)$, for all $y \in Y$ can be given the structure of a vector bundle \mathfrak{g}^F over Y .

To any section $\tilde{\xi}$ of \mathfrak{g}^F , one can associate a vector field $\tilde{\xi}_Y$ on Y according to the following prescription:

$$\tilde{\xi}_Y(y) = [\tilde{\xi}(y)]_Y(y). \tag{10}$$

Definition 4.1. A nonholonomic symmetry is a section $\bar{\xi}$ of \mathfrak{g}^F such that the associated vector field $\bar{\xi}_Y$ is π -projectable; i.e. there exists a vector field $\bar{\xi}_X$ on X such that $T\pi \circ \bar{\xi}_Y = \bar{\xi}_X \circ \pi$.

Definition 4.2. A horizontal nonholonomic symmetry is a section $\bar{\xi}$ of \mathfrak{g}^F which is constant, i.e. $\bar{\xi}(y) = \bar{\xi}(y')$ for all $y, y' \in Y$.

We may identify horizontal nonholonomic symmetries with elements $\xi \in \mathfrak{g}$ such that $\xi = \mathfrak{g}^F(y)$ for all $y \in Y$, that is,

$$j^1 \bar{\xi}_Y(\gamma) \lrcorner F = 0 \quad \text{for all } \gamma \in \mathcal{C} \cap \pi_{1,0}^{-1}(y) \quad \text{and for all } y \in Y,$$

We now define the *nonholonomic momentum map* as the map $J^{\text{n.h.}} : \mathcal{C} \rightarrow \bigwedge^n (J^1\pi) \otimes \mathfrak{g}^F$, constructed as follows. Let $\bar{\xi}$ be any section of \mathfrak{g}^F (for the construction of the momentum map it does not matter whether the associated vector field is projectable or not) and put

$$J_{\bar{\xi}}^{\text{n.h.}} = i_{j^1 \bar{\xi}_Y} \Theta_L, \tag{11}$$

where $\bar{\xi}$ is the vector field associated with $\bar{\xi}$ according to (10). Given this definition of $J_{\bar{\xi}}^{\text{n.h.}}$, we define $J^{\text{n.h.}}$ by the following rule:

$$\langle J^{\text{n.h.}}, \bar{\xi} \rangle = J_{\bar{\xi}}^{\text{n.h.}}.$$

If $\bar{\xi}$ is a nonholonomic symmetry, then the prolongation $j^1 \bar{\xi}_Y$ of the associated vector field is admissible. This is proved below in corollary 4.4.

Lemma 4.3. Let $\bar{\xi}$ be a section of \mathfrak{g}^F . For a fixed $y \in Y$, put $\xi := \bar{\xi}(y)$ and consider any $\gamma \in \pi_{1,0}^{-1}(y) \cap \mathcal{C}$. Then there exists a $\pi_{1,0}$ -vertical vector $v_\gamma \in T_\gamma J^1\pi$ such that

$$j^1 \bar{\xi}_Y(\gamma) = j^1 \xi_Y(\gamma) + v_\gamma.$$

Proof. Recall that $j^1 \xi_Y$ is the prolongation of the fundamental vector field ξ_Y associated with the (fixed) Lie algebra element ξ . The lemma follows from the fact that $T_\gamma \pi_{1,0}(j^1 \bar{\xi}_Y(\gamma)) = \bar{\xi}_Y(y)$. On the other hand, $T_\gamma \pi_{1,0}(j^1 \xi_Y(\gamma)) = \xi_Y(y)$, but by definition, $\bar{\xi}_Y(y) = \xi_Y(y)$. \square

Corollary 4.4. Let $\bar{\xi}$ be a nonholonomic symmetry. Then the prolonged vector field $j^1 \bar{\xi}_Y$ is admissible with respect to any section ϕ of π .

Proof. Take $\gamma \in \mathcal{C}$, put $y = \pi_{1,0}(\gamma)$, and let $\xi = \bar{\xi}(y)$. From definition 4.1, we gather that for all $\Phi \in F$, $j^1 \xi_Y(\gamma) \lrcorner \Phi(\gamma) = 0$. But $j^1 \bar{\xi}_Y(\gamma)$ is equal to $j^1 \xi_Y(\gamma)$ up to a vertical vector, while $\Phi \in F$ is semi-basic. Therefore,

$$j^1 \bar{\xi}_Y(\gamma) \lrcorner \Phi(\gamma) = 0.$$

As this holds for all $\gamma \in \mathcal{C}$, we conclude that $j^1 \bar{\xi}_Y$ is admissible with respect to any section. \square

Remark 4.5. In the case of mechanical systems with linear constraints, the bundle \mathfrak{g}^F , defined as above, coincides with the bundle \mathcal{S} used in [5].

4.2. The nonholonomic momentum equation

Theorem 4.6. *If ϕ is a solution of the nonholonomic field equations (7), then for any nonholonomic symmetry $\bar{\xi}$ the associated component of the momentum map $J_{\bar{\xi}}^{\text{n.h.}}$ satisfies the following nonholonomic momentum equation:*

$$(j^1\phi)^*(dJ_{\bar{\xi}}^{\text{n.h.}}) = (j^1\phi)^*(\mathcal{L}_{j^1\bar{\xi}_Y}(L\eta)).$$

Proof. Let $\bar{\xi}$ be a nonholonomic symmetry. Recall that the prolongation $j^1\bar{\xi}$ of the associated vector field is admissible with respect to any section of π . Therefore, since

$$dJ_{\bar{\xi}}^{\text{n.h.}} = \mathcal{L}_{j^1\bar{\xi}_Y}\Theta_L + i_{j^1\bar{\xi}_Y}\Omega_L,$$

pulling back along a solution of the nonholonomic Euler–Lagrange equations gives us

$$(j^1\phi)^*(dJ_{\bar{\xi}}^{\text{n.h.}}) = (j^1\phi)^*(\mathcal{L}_{j^1\bar{\xi}_Y}\Theta_L). \tag{12}$$

For the sake of notational convenience, let us say that two forms α and β on $J^1\pi$ are equivalent (denoted by $\alpha \simeq \beta$) if they agree up to a contact form, i.e. $\alpha \simeq \beta$ iff $\alpha = \beta + \theta$, where θ is contact. This is equivalent to saying that $(j^1\phi)^*\alpha = (j^1\phi)^*\beta$ for all sections ϕ of π . We then have that

$$\mathcal{L}_{j^1\bar{\xi}_Y}\Theta_L = \mathcal{L}_{j^1\bar{\xi}_Y}(S_\eta^*(dL)) + \mathcal{L}_{j^1\bar{\xi}_Y}(L\eta).$$

In coordinates, the first term on the right-hand side becomes

$$\begin{aligned} \mathcal{L}_{j^1\bar{\xi}_Y}(S_\eta^*(dL)) &= \mathcal{L}_{j^1\bar{\xi}_Y}\left(\frac{\partial L}{\partial y_\mu^a}(dy^a - y_v^a dx^v) \wedge d^n x_\mu\right) \\ &\simeq \frac{\partial L}{\partial y_\mu^a}\mathcal{L}_{j^1\bar{\xi}_Y}(dy^a - y_v^a dx^v) \wedge d^n x_\mu \\ &= \frac{\partial L}{\partial y_\mu^a}\frac{\partial \xi^a}{\partial y^b}(dy^b - y_v^b dx^v) \wedge d^n x_\mu \simeq 0 \end{aligned}$$

and hence vanishes when pulled back along a prolongation of a section. Here, we have used the fact that the prolongation $j^1\bar{\xi}_Y$ can locally be written as

$$j^1\bar{\xi}_Y = \xi^\mu(x)\frac{\partial}{\partial x^\mu} + \xi^a(x, y)\frac{\partial}{\partial y^a} + \left(\frac{d\xi^a}{dx^\mu} - y_v^a\frac{d\xi^v}{dx^\mu}\right)\frac{\partial}{\partial y_\mu^a},$$

and that the Lie derivative of a contact form with respect to such a vector field is again a contact form. Note especially that the coefficient $\xi^\mu(x)$ does not depend on y . This is a consequence of the projectability condition in definition 4.1. □

Corollary 4.7. *If ϕ is a solution of the nonholonomic field equations (7), then for any horizontal nonholonomic symmetry ξ the momentum map $J_\xi^{\text{n.h.}}$ is a conservation law for the nonholonomic dynamics, that is,*

$$(j^1\phi)^*(dJ_\xi^{\text{n.h.}}) = 0.$$

4.3. Vertical vector fields along the projection

Not all nonholonomic symmetries can be modelled as sections of \mathfrak{g}^F ; as we shall see, an important class consists of sections of the pullback bundle $\pi_{1,0}^* \mathfrak{g}^F$, i.e. maps $\bar{\xi} : \mathcal{C} \rightarrow \mathfrak{g}^F$ such that $\bar{\xi}(\gamma) \in \mathfrak{g}^F(\pi_{1,0}(\gamma))$ for all $\gamma \in \mathcal{C}$. Such a section can be represented as

$$\bar{\xi} = \xi^A(x^\mu, y^a, y_\mu^a) e_A, \tag{13}$$

where $\xi^A(x^\mu, y^a, y_\mu^a)$ are locally defined functions on \mathcal{C} and $\{e_A\}$ is a basis of \mathfrak{g} .

For this treatment, we will make one important simplification: *we assume that the symmetry group G acts vertically.* It should be noted, though, that this is probably not a fundamental restriction: it is likely that this approach can be generalized further.

Our treatment is therefore somewhat restricted compared to the previous paragraph. On the other hand, by allowing as in (13) sections whose coefficients are functions on \mathcal{C} rather than just on Y , we are able to describe symmetries which do not fit into the framework of the previous paragraph. Examples will be given in section 5.

Definition 4.8. *A generalized nonholonomic symmetry is a section $\bar{\xi}$ of the pullback bundle $\pi_{1,0}^* \mathfrak{g}^F$.*

A generalized nonholonomic symmetry $\bar{\xi}$ induces a vertical vector field $\tilde{\xi}$ along $\pi_{1,0}$, defined as follows for all $\gamma \in \mathcal{C}$:

$$\tilde{\xi}(\gamma) = [\bar{\xi}(\gamma)]_Y(y) \in T_y Y, \quad \text{where} \quad y = \pi_{1,0}(\gamma).$$

There exists a generalized notion of prolongation for vector fields along $\pi_{1,0}$ (see [20]). In coordinates, if $\tilde{\xi} = \xi^a(x^\mu, y^b, y_\mu^b) \frac{\partial}{\partial y^a}$, then

$$j^1 \tilde{\xi}_Y = \xi^a \frac{\partial}{\partial y^a} + \frac{d\xi^a}{dx^\mu} \frac{\partial}{\partial y_\mu^a}, \quad \text{where} \quad \frac{d\xi^a}{dx^\mu} = \frac{\partial \xi^a}{\partial x^\mu} + \frac{\partial \xi^a}{\partial y^b} y_\mu^b + \frac{\partial \xi^a}{\partial y_\nu^b} y_{\mu\nu}^b.$$

Note that $j^1 \tilde{\xi}_Y$ is a vector field along $\pi_{2,1}$, as its coefficients depend on the second-order derivatives $y_{\mu\nu}^b$.

Having defined the prolongation $j^1 \tilde{\xi}_Y$ as above, we now define the associated component of the momentum map as

$$J_{\bar{\xi}}^{\text{n.h.}} = i_{j^1 \tilde{\xi}_Y} \Theta_L.$$

This definition is formally identical to (11); but strictly speaking, the momentum map is now an n -form along the projection $\pi_{2,1}$. By pullback, such a form induces an n -form on $J^2 \pi$. Similarly, the concept of Lie derivation can be extended in a natural fashion to the case of vector fields along the projection (see [20]).

Theorem 4.9. *If ϕ is a solution of the nonholonomic field equations (7), then for any generalized nonholonomic symmetry $\bar{\xi}$ the associated component of the momentum map $J_{\bar{\xi}}^{\text{n.h.}}$ satisfies the following nonholonomic momentum equation:*

$$(j^2 \phi)^*(dJ_{\bar{\xi}}^{\text{n.h.}}) = (j^2 \phi)^*(\mathcal{L}_{j^1 \tilde{\xi}_Y}(L\eta)). \tag{14}$$

Proof. The proof is similar, in spirit, to that of theorem 4.6; but there are some additional technical difficulties that need to be taken into account. For any solution ϕ of the nonholonomic field equations, we have

$$(j^2 \phi)^*(dJ_{\bar{\xi}}^{\text{n.h.}}) = (j^2 \phi)^*(\mathcal{L}_{j^1 \tilde{\xi}_Y} \Theta_L) + (j^2 \phi)^*(i_{j^1 \tilde{\xi}_Y} \Omega_L). \tag{15}$$

For the first term on the right-hand side, one can check that

$$\mathcal{L}_{j^1\tilde{\xi}_Y}(dy^a - y^a_\mu dx^\mu) = \frac{\partial \xi^a}{\partial y^b}(dy^b - y^b_\mu dx^\mu) + \frac{\partial \xi^a}{\partial y^b_\mu}(dy^b_\mu - y^b_{\mu\nu} dx^\nu),$$

a contact form on $J^2\pi$. From this, it follows that the Lie derivative of Θ_L is given by

$$(j^2\phi)^*(\mathcal{L}_{j^1\tilde{\xi}_Y}\Theta_L) = (j^2\phi)^*(\mathcal{L}_{j^1\tilde{\xi}_Y}(L\eta)).$$

The second term on the right-hand side of (15) is zero. This is essentially a consequence of the fact that ϕ is a solution of the field equations and can be proved using the following observation. Take $x \in X$ and consider a vector field W on $J^1\pi$ such that $W(j^1_x\phi) = j^1_{\xi_Y}(j^2_x\phi)$. Then $W(j^1_x\phi)\lrcorner\Phi = 0$ for all $\Phi \in F$, and without loss of generality, we may choose W to be such that this equality holds in a neighbourhood of $j^1_x\phi$. We now have that

$$(j^2\phi)^*(i_{j^1\tilde{\xi}_Y}\Omega_L)(x) = (j^1\phi)^*(i_W\Omega_L)(x).$$

As W is admissible with respect to ϕ , we conclude that the right-hand side is zero, and this in turn implies the vanishing of the second term in (15). \square

5. Examples

5.1. Mechanical systems with nonlinear constraints

As an illustrative example, we take a variation of Benenti’s system (see [9, 14] and the references therein), which describes two point masses moving on a horizontal plane whose velocities are constrained to be parallel. The configuration space of this system is $Q := \mathbb{R}^2 \times \mathbb{R}^2$, and so we put $X = \mathbb{R}$, while $Y = \mathbb{R} \times Q$. The projection π is the projection onto the first factor. The Lagrangian for this system is

$$L = \frac{m}{2}(\dot{x}_1^2 + \dot{y}_1^2) + \frac{m}{2}(\dot{x}_2^2 + \dot{y}_2^2).$$

Strictly speaking, L is a function on TQ , but since $J^1\pi$ is isomorphic to $\mathbb{R} \times TQ$, we view L as a (time-independent) Lagrangian on $J^1\pi$. The constraint is

$$\varphi \equiv \dot{x}_1\dot{y}_2 - \dot{x}_2\dot{y}_1 = 0.$$

This is a nonlinear constraint and determines a submanifold \mathcal{C} of $J^1\pi$. If we assume Chetaev’s principle to hold, the bundle of reaction forces F along \mathcal{C} is generated by the following 1-form:

$$\Phi = \dot{y}_2\theta_{x_1} + \dot{x}_1\theta_{y_2} - \dot{y}_1\theta_{x_2} - \dot{x}_2\theta_{y_1},$$

where the contact 1-forms θ_{q^a} are defined as $\theta_{q^a} = dq^a - \dot{q}^a dt$.

Consider now the obvious action of the Abelian group \mathbb{R}^4 on Y by translations: $(a_1, b_1, a_2, b_2) \cdot (t; x_1, y_1, x_2, y_2) = (t; x_1 + a_1, y_1 + b_1, x_2 + a_2, y_2 + b_2)$. This action fulfils all the necessary conditions needed for the momentum lemma: it is vertical and leaves invariant the Lagrangian L , the constraint submanifold \mathcal{C} and the bundle of reaction forces F .

It is easy to check that the contraction of a vector field $\tilde{\xi}$ of the following form:

$$\tilde{\xi} = \alpha\dot{x}_1\frac{\partial}{\partial x_1} + \beta\dot{y}_1\frac{\partial}{\partial y_1} + \gamma\dot{x}_2\frac{\partial}{\partial x_2} + \delta\dot{y}_2\frac{\partial}{\partial y_2} \quad (\alpha, \beta, \gamma, \delta \in \mathbb{R})$$

with Φ will vanish along \mathcal{C} if $\alpha + \delta = \beta + \gamma$. In this case, $\tilde{\xi}$ is a generalized nonholonomic symmetry of the kind defined in definition 4.8. (Strictly speaking, the term ‘generalized nonholonomic symmetry’ refers to the section $\tilde{\xi} := (\alpha\dot{x}_1, \beta\dot{y}_1, \gamma\dot{x}_2, \delta\dot{y}_2)$ of $\pi_{1,0}^*\mathfrak{g}^F$.)

Table 1. Conservation laws for the Benenti system.

$(\alpha, \beta, \gamma, \delta)$	
$(1, 1, 0, 0)$	$\dot{x}_1\ddot{x}_1 + \dot{y}_1\ddot{y}_1 = 0$
$(0, 0, 1, 1)$	$\dot{x}_2\ddot{x}_2 + \dot{y}_2\ddot{y}_2 = 0$
$(1, 0, 0, -1)$	$\dot{x}_1\ddot{x}_1 - \dot{y}_2\ddot{y}_2 = 0$

For the sake of convenience, let us take $\alpha = \beta = 1$ and $\gamma = \delta = 0$. The other cases are similar and the results are summarized in table 1. In this case, the prolongation $j^1\tilde{\xi}$ of $\tilde{\xi}$ is the vector field along $\pi_{2,1}$ given in coordinates by

$$j^1\tilde{\xi} = \dot{x}_1 \frac{\partial}{\partial x_1} + \dot{y}_1 \frac{\partial}{\partial y_1} + \ddot{x}_1 \frac{\partial}{\partial \dot{x}_1} + \ddot{y}_1 \frac{\partial}{\partial \dot{y}_1}.$$

The component $J_{\tilde{\xi}}^{\text{n.h.}}$ of the nonholonomic momentum map then becomes

$$J_{\tilde{\xi}}^{\text{n.h.}} = j^1\tilde{\xi} \lrcorner \Theta_L = m(\dot{x}_1^2 + \dot{y}_1^2),$$

while the right-hand side of the momentum equation (14) is $\mathcal{L}_{j^1\tilde{\xi}}(L dt) = m(\dot{x}_1\ddot{x}_1 + \dot{y}_1\ddot{y}_1) dt$. The momentum equation hence reduces to $\dot{x}_1\ddot{x}_1 + \dot{y}_1\ddot{y}_1 = 0$, which can also be verified using the equations of motion.

5.2. The nonholonomic Cosserat rod

The nonholonomic Cosserat rod is an example of a nonholonomic field theory studied in [23]. It describes the motion of a rod which is constrained to roll without sliding on a horizontal surface. This theory can be studied using the bundle $\pi : Y \rightarrow X$, where $X = [0, \ell] \times \mathbb{R}$ (space and time) and $Y = X \times \mathbb{R}^2 \times \mathbb{S}^1$, with bundle coordinates $(s, t; x, y, \theta)$.

Its Lagrangian is given by

$$L = \frac{\rho}{2}(\dot{x}^2 + \dot{y}^2) + \frac{\alpha}{2}\dot{\theta}^2 - \frac{1}{2}(\beta(\theta')^2 + K\kappa^2)$$

with $\kappa^2 = (x'')^2 + (y'')^2$, while the constraints are given by

$$\dot{x} + R\dot{\theta}y' = 0 \quad \text{and} \quad \dot{y} - R\dot{\theta}x' = 0. \tag{16}$$

Here, ρ, α, β, K and R are real parameters. The field equations associated with this Lagrangian are given by

$$\begin{cases} \rho\ddot{x} + Kx'''' = \lambda \\ \rho\ddot{y} + Ky'''' = \mu \\ \alpha\ddot{\theta} - \beta\theta'' = R(\lambda y' - \mu x'), \end{cases} \tag{17}$$

where λ and μ are Lagrange multipliers associated with the nonholonomic constraints. These equations are to be supplemented by the constraint equations (16).

For future reference, we note that the bundle of reaction forces is generated in this case by the following forms:

$$\begin{cases} \Phi^1 = (dx - \dot{x} dt) \wedge ds + Ry'(d\theta - \dot{\theta} dt) \wedge ds \\ \Phi^2 = (dy - \dot{y} dt) \wedge ds - Rx'(d\theta - \dot{\theta} dt) \wedge ds. \end{cases} \tag{18}$$

In the absence of nonholonomic constraints, this model is subject to the usual symmetry actions such as translations in time, global translations and global rotations (see [23]). As we shall now show, some of these persist in the nonholonomic case.

Observe that the Lagrangian is of second order. At the cost of sacrificing physical relevance, one may also put $K = 0$ to obtain a purely first-order theory. We will not do so here, as the derivation of a nonholonomic momentum lemma for a second-order field theory proceeds exactly as above, up to a few minor modifications. The nonholonomic momentum map $J^{\text{n.h.}}$ is now defined on $J^3\pi$, and the nonholonomic momentum equation hence becomes

$$(j^3\phi)^*(dJ_{\tilde{\xi}}^{\text{n.h.}}) = (j^2\phi)^*\mathcal{L}_{j^2\tilde{\xi}}(L\eta), \tag{19}$$

where ϕ is a solution of the nonholonomic Euler–Lagrange equations. This equation holds both for ‘genuine’ symmetries as well as for symmetries along the projection.

5.2.1. Translations in time. Consider the action of \mathbb{R} on X by translations in time defined by the map $\Phi : \mathbb{R} \times X \rightarrow X$, with $\Phi(\alpha, (s, t)) = (s, t + \alpha)$. As the bundle π is trivial, this action naturally induces an action on Y , and, by prolongation, also on $J^1\pi$. Clearly, this action is not vertical; the fundamental vector field associated with a Lie algebra element $\xi \in \mathbb{R}$ is given by

$$j^1\xi_Y = \xi \frac{\partial}{\partial t}. \tag{20}$$

This vector field is a nonholonomic symmetry: it is π_1 -related to a vector field on X and its contraction with the elements of F vanishes along \mathcal{C} : for the n -form Φ^1 defined in (18), we have

$$\frac{\partial}{\partial t} \lrcorner \Phi^1 = -(\dot{x} + R\dot{\theta}y') ds,$$

which vanishes on \mathcal{C} , and a similar result holds for Φ^2 .

Note that the nonholonomic symmetry (20) is a horizontal symmetry (definition 4.2); in general, this will not be the case. In this special case, however, we have that $\mathcal{L}_{j^1\xi_Y}(L\eta) = 0$, expressing the infinitesimal invariance of the Lagrangian. The momentum map now becomes

$$(j^3\phi)^*J_1^{\text{n.h.}} = [-Kx'''\dot{x} - Ky'''\dot{y} + \beta\theta'\dot{\theta} + K(x''\dot{x}' + y''\dot{y}')] dt + \mathcal{E} ds,$$

where we have introduced the *energy density*

$$\mathcal{E} = \frac{\rho}{2}(\dot{x}^2 + \dot{y}^2) + \frac{\alpha}{2}\dot{\theta}^2 + \frac{K}{2}((x'')^2 + (y'')^2) + \frac{\beta}{2}(\theta')^2.$$

The nonholonomic momentum lemma then states that $d(j^3\phi)^*J_1^{\text{n.h.}} = 0$. This equation expresses local conservation of energy; by integrating over a hypersurface of constant s , we may then obtain a law expressing global conservation of energy.

Remark 5.1. This case was also treated in [23] by use of a different method, to which we refer for further details. Note, however, that the method used in that paper is less general as it is only valid for horizontal nonholonomic symmetries.

5.2.2. Spatial translations. Consider the action of $\mathbb{R}^2 \times \mathbb{S}^1$ on Y by translations; i.e. for each (a, b, φ) we consider the map $\Phi_{(a,b,\varphi)} : (s, t, x, y, \theta) \mapsto (s, t, x + a, y + b, \theta + \varphi)$. Let $\xi = (v_1, v_2, v_\theta)$ be an element of the Lie algebra of $\mathbb{R}^2 \times \mathbb{S}^1$. The corresponding fundamental vector field is given by

$$\xi_Y = v_1 \frac{\partial}{\partial x} + v_2 \frac{\partial}{\partial y} + v_\theta \frac{\partial}{\partial \theta}.$$

When no constraints are present, this symmetry implies the conservation of linear momentum. In the presence of nonholonomic constraints, a modified conservation law holds: it is easy to see that the following vector field annihilates F along \mathcal{C} :

$$\tilde{\xi} = -Ry' \frac{\partial}{\partial x} + Rx' \frac{\partial}{\partial y} + \frac{\partial}{\partial \theta}.$$

(Any scalar multiple of the above vector field is also allowed.) This generalized vector field corresponds to the section $\tilde{\xi} = (-Ry', Rx', 1)$ of $\pi_{1,0}^* \mathfrak{g}^F$. As $\tilde{\xi}$ is vertical, the nonholonomic momentum lemma 4.9 can be applied.

The right-hand side of (19) is

$$j^2 \phi^* \mathcal{L}_{\tilde{\xi}}(L\eta) = [-R\rho\dot{y}'\dot{x} + R\rho\dot{x}'\dot{y} - KRx'''y'' + K Ry'''x'']\eta.$$

The nonholonomic momentum map $J^{\text{n.h.}}$, on the other hand, is given by

$$J_{\tilde{\xi}}^{\text{n.h.}} = -[\rho(Rx'\dot{y} - Ry'\dot{x}) + \alpha\dot{\theta}] ds - [KR(y'x''' - x'y''')] + \beta\theta'] dt,$$

and the nonholonomic momentum equation hence becomes

$$Ry'(\rho\ddot{x} + Kx''''') - Rx'(\rho\ddot{y} + Ky''''') = \alpha\ddot{\theta} - \beta\theta''. \tag{21}$$

This conservation law can also be derived from the nonholonomic field equations (17) by subtracting the second equation multiplied by x' from the first equation multiplied by y' , and using the third equation to eliminate the Lagrange multipliers λ and μ . Unfortunately, the knowledge of this nonholonomic conservation law does not help us in solving the field equations (in contrast with the situation for the vertical rolling disc; see [4]).

6. Conclusions

In this paper, we presented a geometric framework for nonholonomic field theories with symmetries. We showed that the momentum map associated with a group action satisfies a certain momentum equation, which we proved in a number of cases. On the one hand, there exists a momentum equation for nonholonomic symmetries which act nontrivially on the base space (examples being energy conservation for the nonholonomic rod), while on the other hand a similar result exists for generalized symmetries associated with a vertical group action. The Benenti system is an example of the latter.

It is likely that the results in this paper can be generalized still further. In particular, there seems to be no reason why there should not be a momentum equation for generalized symmetries associated with a non-vertical group action, up to some technical restrictions (for instance, that the generalized vector field should be related to a regular vector field on the base space). This generalization would encompass both momentum equations derived in this paper. Moreover, it would be interesting to study further examples of such a momentum equation.

Acknowledgments

The authors would like to thank F Cantrijn and D Saunders for stimulating discussions and comments. The first author is a Postdoctoral Fellow from the Research Foundation—Flanders (FWO-Vlaanderen), and a Fulbright Research Scholar at the California Institute of Technology. Additional financial support from the Fonds Professor Wuytack is gratefully acknowledged. The second author is supported by MEC (Spain) grants MTM 2004-7832 and MTM2007-62478, project ‘Ingenio Mathematica’ (i-MATH) no. CSD 2006-00032 (Consolider-Ingenio 2010) and S-0505/ESP/0158 of the CAM.

Appendix A. Noncovariant nonholonomic constraints

A special class of nonholonomic field theories consists of those where the base space X is $\mathbb{R} \times M$, where the first factor represents time, and such that π is trivial. In other words, there exists a canonical distinction between time and space. Accordingly, one can show that in this

case, the jet bundle $J^1\pi$ is isomorphic to the product bundle $\mathbb{R} \times [J^1(M, S) \times_S TS]$, thus providing a canonical distinction between derivatives of the fields with respect to time and space.

This class of field theories was discussed in detail in [23] and includes among others also theories of nonrelativistic elasticity (see also [17]). The example studied in section 5.2 is also among these field theories.

As shown in [23], the reaction forces for this kind of field theory are not the ones obtained from the Chetaev prescription in remark 3.6. The problem is that this form of the principle is fully covariant, in the sense that no distinction is made between spatial derivatives and derivatives with respect to time. However, in some cases (the nonholonomic Cosserat rod being one of them) the time derivatives *do* play a distinguished role, and therefore a different, ‘noncovariant’ Chetaev principle is needed.

The main objective of this appendix is to propose such a principle, using the geometry of $J^1\pi$, and especially the isomorphism with $\mathbb{R} \times [J^1(M, S) \times_S TS]$. Using this modified principle, we derive the correct form of the reaction forces for the nonholonomic Cosserat rod, which were obtained in [23] using a different approach.

A.1. A new vertical endomorphism

Recall the coordinate expression (1) of the vertical endomorphism S_η on $J^1\pi$. This tensor field was constructed by Saunders [20] using a map assigning to each 1-form ω on X the vector-valued 1-form S_ω on $J^1\pi$ given in coordinates by (see [20, p 156])

$$S_\omega = \omega_\mu (dy^a - y^a_\nu dx^\nu) \otimes \frac{\partial}{\partial y^a_\mu}, \quad \text{where} \quad \omega = \omega_\mu dx^\mu. \quad (\text{A.1})$$

Roughly speaking, the vertical endomorphism S_η then arises, once a volume form on X is chosen, by putting

$$S_\eta = S_{dx^\mu} \hat{\wedge} (\pi_1^* d^n x_\mu), \quad (\text{A.2})$$

where the wedge operator ‘ $\hat{\wedge}$ ’ is defined as follows: if Φ is a vector-valued k -form on $J^1\pi$, and α is a regular (i.e. \mathbb{R} -valued) l -form, then $\Phi \hat{\wedge} \alpha$ is the vector-valued $(k+l)$ -form given by $\langle \Phi \hat{\wedge} \alpha, \beta \rangle = \langle \Phi, \beta \rangle \wedge \alpha$ for all $\beta \in \Omega^1(J^1\pi)$.

It is obvious that (A.2) is fully covariant, in the sense that no distinction is made between the variables on the base space. However, in elastodynamics, this is not always desirable, as we have a distinguished direction of time. Therefore, we propose the following ‘non-covariant’ vertical endomorphism:

Definition 1.1. *The non-covariant vertical endomorphism is the vector-valued $(n+1)$ -form $S_{n.c.}$ defined as $S_{n.c.} := S_{dt} \hat{\wedge} (\pi_1^* \eta_M)$, where S_{dt} is the vector-valued 1-form associated with dt as in (A.1).*

Here η_M denotes the volume form on M defined by restriction to M of $d^n x_0$

Note that dt is a well-defined 1-form on $X = \mathbb{R} \times M$; therefore, $S_{n.c.}$ is an intrinsic object. In coordinates, $S_{n.c.}$ is given by

$$S_{n.c.} = (dy^a - y^a_\mu dx^\mu) \wedge d^n x_0 \otimes \frac{\partial}{\partial y^a_0}.$$

Remark 1.2. In [20], it is shown that the action of the vertical endomorphism S_η is related to the fact that $\pi_{1,0} : J^1\pi \rightarrow Y$ is an affine bundle; i.e. there exists an affine action of $\pi^*T^*X \otimes V\pi$ on $J^1\pi$. The noncovariant vertical endomorphism can be understood in a similar vein, by restricting this affine action to $\pi^*T^*\mathbb{R} \otimes V\pi$.

A.2. The bundle of constraint forces

Let $\iota : \mathcal{C} \hookrightarrow J^1\pi$ be a constraint manifold. In section 3, we required that reaction forces be n -horizontal and 1-contact. This leads to local expressions of the form (4). For noncovariant constraints, we attribute a special status to the time coordinate, and therefore we require also that the following holds:

$$i_v i_w \Phi = 0$$

for all tangent vectors v, w on $J^1\pi$ such that $T(\text{pr}_1 \circ \pi_1)(v) = T(\text{pr}_1 \circ \pi_1)(w) = 0$, where $\text{pr}_1 : \mathbb{R} \times M \rightarrow \mathbb{R}$ is the projection onto the first factor (this condition expresses that v and w do not contain a component proportional to $\frac{\partial}{\partial t}$). In coordinates, this implies that Φ has the following form:

$$\Phi = A_a (dy^a - y_\mu^a dx^\mu) \wedge d^n x_0, \quad (\text{A.3})$$

where the A_a are local functions on \mathcal{C} . Compare this with (4) and note that only the ‘time’ component is left (i.e. the component proportional to $\theta^a \wedge (\frac{\partial}{\partial t} \lrcorner \eta)$).

Using the noncovariant vertical endomorphism of section appendix A.1, one can construct from \mathcal{C} a natural candidate for the bundle F . This is a field-theoretic generalization of the well-known Chetaev principle from mechanics.

Recall that \mathcal{C} is assumed to be given by the vanishing of k functions φ^α and define the associated bundle of reaction forces as the bundle F locally spanned by the $(n+1)$ -forms $\Phi^\alpha := S_{\text{n.c.}}^*(d\varphi^\alpha)$, or in coordinates

$$\Phi^\alpha = \frac{\partial \varphi^\alpha}{\partial y_0^a} (dy^a - y_\mu^a dx^\mu) \wedge d^n x_0. \quad (\text{A.4})$$

The $(n+1)$ -form Φ^α is therefore of the form outlined in (A.3), with $A_a^\alpha = \frac{\partial \varphi^\alpha}{\partial y_0^a}$.

References

- [1] Bates L and Śniatycki J 1993 Nonholonomic reduction *Rep. Math. Phys.* **32** 99–115
- [2] Benenti S 1996 Geometrical aspects of the dynamics of non-holonomic systems *Rend. Sem. Mat. Univ. Politec. Torino* **54** 203–12
- [3] Binz E, de León M, Martín de Diego D and Socolescu D 2002 Nonholonomic constraints in classical field theories *Rep. Math. Phys.* **49** 151–66
- [4] Bloch A 2003 Nonholonomic mechanics and control *Interdisciplinary Applied Mathematics* vol 24 (Berlin: Springer)
- [5] Bloch A, Krishnaprasad P, Marsden J E and Murray R 1996 Nonholonomic mechanical systems with symmetry *Arch. Rat. Mech. Anal.* **136** 21–99
- [6] Cantrijn F, de León M, Marrero J C and Martín de Diego D 1998 Reduction of nonholonomic mechanical systems with symmetry *Rep. Math. Phys.* **42** 25–45
- [7] Cantrijn F, de León M, Marrero J C and Martín de Diego D 1999 Reduction of constrained systems with symmetries *J. Math. Phys.* **40** 795–820
- [8] Cariñena J F, Crampin M and Ibrat L A 1991 On the multisymplectic formalism for first order field theories *Diff. Geom. Appl.* **1** 345–74
- [9] Cortés J 2002 Geometric, control and numerical aspects of nonholonomic systems *Lecture Notes in Mathematics* vol 1793 (Berlin: Springer)
- [10] de León M, Martín de Diego D and Santamaría-Merino A 2004 Symmetries in classical field theory *Int. J. Geom. Meth. Mod. Phys.* **1** 651–710
- [11] de León M, McLean M, Norris L, Roca A and Salgado M 2002 Geometric structures in field theory *Preprint math-ph/0208036*
- [12] de León M and Rodrigues P R 1985 Generalized classical mechanics and field theory *A Geometrical Approach of Lagrangian and Hamiltonian Formalisms Involving Higher Order Derivatives* (Amsterdam: North-Holland) p 112

- [13] Gotay M, Isenberg J and Marsden J 1998 Momentum maps and classical relativistic fields: part I. Covariant field theory *Preprint physics/9801019*
- [14] Ibrort A, de León M, Marmo G and Martín de Diego D 1996 Non-holonomic constrained systems as implicit differential equations *Rend. Sem. Mat. Univ. Politec. Torino* **54** 295–317
- [15] Krupková O and Volný P 2005 Euler–Lagrange and Hamilton equations for nonholonomic systems in field theory *J. Phys. A: Math. Gen.* **38** 8715–45
- [16] Marle C M 1998 Various approaches to conservative and nonconservative nonholonomic systems *Rep. Math. Phys.* **42** 211–29
- [17] Marsden J E, Pekarsky S, Shkoller S and West M 2001 Variational methods, multisymplectic geometry and continuum mechanics *J. Geom. Phys.* **38** 253–84
- [18] Neimark J and Fufaev N 1972 *Dynamics of Nonholonomic Systems (Translations of Mathematical Monographs vol 33)* (Providence, RI: American Mathematical Society)
- [19] Olver P 1986 *Applications of Lie Groups to Differential Equations (Graduate Texts in Mathematics vol 107)* (New York: Springer)
- [20] Saunders D J 1989 *The Geometry of Jet Bundles (London Mathematical Society Lecture Note Series vol 142)* (Cambridge: Cambridge University Press)
- [21] Vankerschaver J, Cantrijn F, de León M and Martín de Diego D 2005 Geometric aspects of nonholonomic field theories *Rep. Math. Phys.* **56** 387–411 (*Preprint math-ph/0506010*)
- [22] Vankerschaver J 2005 The momentum map for nonholonomic field theories with symmetry *Int. J. Geom. Meth. Mod. Phys.* **2** 1029–41 (*Preprint math-ph/0507059*)
- [23] Vankerschaver J 2007 A class of nonholonomic kinematic constraints in elasticity *J. Phys. A: Math. Theor.* **40** 3889–913