Journal of Geometry and Physics 35 (2000) 126–144

Vakonomic mechanics versus non-holonomic mechanics: a unified geometrical approach

Manuel de Leon [a,*], Juan C. Marrero [b], David Martin de Diego [c]

[a] *Instituto de Matemáticas y Física Fundamental, CSIC, Serrano 123, 28006 Madrid, Spain*
[b] *Dpto. de Matemática Fundamental, Facultad de Matemáticas, Universidad de La Laguna, Tenerife, Canary Islands, Spain*
[c] *Departamento de Economía Aplicada (Matemáticas), Facultad de CC Económicas y Empresariales, Avda. Valle Esgueva 6, 47011 Valladolid, Spain*

Received 15 October 1999

Abstract

In this paper some geometrical aspects of constrained mechanics are studied. A symplectic setting for vakonomic mechanics is given, and the conservation of the energy is discussed. This formulation needs constraint forces involving accelerations. A reduction procedure is given for Čaplygin vakonomic systems. Finally, a unified geometrical framework for non-holonomic and vakonomic mechanics is described. © 2000 Elsevier Science B.V. All rights reserved.

MSC: 58F05; 70H30; 70H33

PACS: 03.20.+i.

Subj. Class: Dynamic systems

Keywords: Vakonomic mechanics; Non-holonomic mechanics; Symmetries; Reduction

1. Introduction

Non-holonomic mechanics have gained much attention in the last years. Indeed, there are many efforts to put it in the stream of the so-called geometric mechanics (see for instance [1,3–9,12–18,20–24,27–34] and references therein).

A non-holonomic mechanical system consists of a Lagrangian function L defined on the space of velocities TQ of a configuration manifold Q subjected to constraints given

* Corresponding author.
E-mail addresses: mdeleon@imaff.cfmac.csis.es, mdeleon@pinar1.csis.es (M. de Leon), jcmarrer@ull.es (J.C. Marrero), dmartin@esgueva.eco.uva.es (D.M.de. Diego)

0393-0440/00/$ – see front matter
PII: S0393-0440(00)00004-8

by a submanifold M of TQ. If a compatibility condition is assumed one can obtain the constrained dynamics by projecting the unconstrained one.

A different approach to constrained mechanics is the so-called vakonomic mechanics studied by Kosloz [19] (see also [2,27]). In contrast with non-holonomic mechanics, vakonomic mechanics comes from a variational principle. Roughly speaking, in non-holonomic mechanics one consider extremals for the action which are compatible with the constraints; on the other hand, in vakonomic mechanics one looks for extremals in the class of curves which satisfy the constraints. Both mechanics lead to different equations of motion, and the solutions coincide if the constraints are holonomic.

Our purpose in this paper is to describe a geometrical (symplectic) setting for vakonomic mechanics. We first discuss in Section 2 some aspects of the above variational issues concerning unconstrained, non-holonomic and vakonomic mechanics. Our discussion is inspired in that by Lewis and Murray [27], but we deal with non-linear constraints. The main conclusion is that the action is always the same, but the allowable virtual displacements are different according to the cases. In Section 3, we recall the symplectic formalism for unconstrained mechanics, and in Section 4, we develop a symplectic setting for non-holonomic mechanics. Apart from the Lagrangian forces we have the reaction forces due to the constraints, which are understood here as a suitable 1-form taking values in the sometimes called Chetaev bundle $S^*(TM^0)$, where S is the vertical endomorphism on TQ. As we said before, the compatibility condition allows us to obtain the constrained dynamics in an intrinsic way. Moreover, if the constraint submanifold M is homogeneous, then the energy is a conserved quantity. The main results are contained in Sections 5 and 6. In Section 5 we construct a geometrical framework for vakonomic mechanics. The bundle of forces acting on the system is given by the space of semibasic 1-forms $\lambda^i \delta\Phi_i + \mathrm{d}_T(\lambda^i) S^*(\mathrm{d}\Phi_i)$, where δ is the Euler–Lagrange operator [35], Φ_i a family of independent constraints, d_T the total derivative with respect to time, and λ^i are arbitrary functions on TQ. Therefore, the forces involve accelerations. In fact, they are 1-forms defined along the canonical projection $\tau_{21} : T^2Q \to TQ$ of the second or der tangent bundle T^2Q onto TQ along M. Even if the compatibility condition is assumed, the existence of a solution on M is not guaranteed. Indeed, a vakonomic system is equivalent to a pre-symplectic one (see [2,10]). A solution, if exists, is necessarily a SODE, and in fact it is a vector field along the canonical projection $\tau_{21} : T^2Q \to TQ$. Moreover, we prove that the energy is a conserved quantity if and only if the work performed for any "non-holonomic" reaction force $\mu^i S^*(\mathrm{d}\Phi_i)$ is constant along the motions. In Section 6 we discuss an example (a skate on an inclined plane) in order to clarify the ideas and results contained in this paper. A unified symplectic setting is proposed in Section 7. Finally, in Section 8 we develop a reduction procedure for Čaplygin vakonomic systems, and apply it to reduce the vakonomic equations of motion of the rolling disk.

2. Some remarks on variational principles in mechanics

Let Q be a configuration n-dimensional manifold, and $L : TQ \to \mathbb{R}$ an autonomous Lagrangian function. If (q^A) are coordinates on Q, we denote by $(q^A, \dot{q}^A)$ the fibred coordinates on TQ. Then the tangent bundle projection $\tau_Q : TQ \to Q$ reads as $\tau_Q(q^A, \dot{q}^A) = (q^A)$.

We will denote by $S = \mathrm{d}q^A \otimes (\partial/\partial\dot{q}^A)$ the canonical vertical endomorphism on TQ, and by $\Delta = \dot{q}^A(\partial/\partial\dot{q}^A)$ the Liouville vector field of TQ (see [26] for the intrinsic definitions). In what follows, S^* denotes the adjoint operator defined by $S^*(\alpha) = \alpha \circ S$, for any 1-form α on TQ.

If f is a function on the tangent bundle of order k of Q, i.e., $f : T^kQ \to \mathbb{R}$, we denote by $\mathrm{d}_T f$ the total derivative of f which is a function on the tangent bundle $T^{k+1}Q$ of order $k+1$ of Q; $\mathrm{d}_T f$ is locally defined by

$$\mathrm{d}_T f = \dot{q}^A \frac{\partial f}{\partial q^A} + \ddot{q}^A \frac{\partial f}{\partial \dot{q}^A} + \cdots + q^{(k+1)A} \frac{\partial f}{\partial q^{(k)A}},$$

where $(q^A, \dot{q}^A, \ldots, q^{(r)A})$ stand for the induced coordinates on the tangent bundle of order r (see [25,35] for more details). The operator d_T can be extended to differential forms in a very natural way by requiring that $\mathrm{d}_T\,\mathrm{d} = \mathrm{d}\,\mathrm{d}_T$.

The operator d_T was used by Tulczyjew [35] to define the Lagrange differential

$$\delta(\Psi) = \mathrm{d}\Psi - \mathrm{d}_T(S^*(\mathrm{d}\Psi))$$

for all function $\Psi \in C^\infty(TQ)$. In local coordinates we have

$$\delta(\Psi) = \left(\frac{\partial \Psi}{\partial q^A} - \mathrm{d}_T\left(\frac{\partial \Psi}{\partial \dot{q}^A}\right)\right) \mathrm{d}q^A,$$

which is a 1-form on T^2Q.

Given two points $x, y \in Q$ we define the manifold of twice piecewise differentiable curves which connect x and y as

$$\mathcal{C}^2(x, y) = \{c : [0,1] \to Q | c \text{ is } C^2,\ c(0) = x, \text{ and } c(1) = y\}.$$

Let c be a curve in $\mathcal{C}^2(x, y)$. As is well known the tangent space of $\mathcal{C}^2(x, y)$ at c is given by

$$\begin{aligned} T_c\mathcal{C}^2(x, y) = \{&X : [0,1] \to TQ \mid X \text{ is } C^1, X(t) \in T_{c(t)}Q, \\ &X(0) = 0, \text{ and } X(1) = 0\}. \end{aligned}$$

We assume that L is subjected to constraints given by a submanifold M of TQ. M is locally defined as the zero set of m functions $\{\Phi_1, \ldots, \Phi_m\}$, where m is the codimension of M in TQ. The bundle of Chetaev or constraint forces is $S^*((TM)^0)$ (see [14,21]), where $(TM)^0$ denotes the annihilator of the tangent bundle TM. Notice that the Chetaev bundle is only defined along M. The Chetaev bundle is locally generated by the semibasic 1-forms

$$S^*(\mathrm{d}\Phi_i) = \frac{\partial \Phi_i}{\partial \dot{q}^A}\,\mathrm{d}q^A, \quad 1 \le i \le m.$$

It is tacitly assumed that the 1-forms $S^*(\mathrm{d}\Phi_i)$ are indeed linearly independent. Therefore, an arbitrary constraint force is a linear combination of them: $\lambda^i(\partial\Phi_i/\partial\dot{q}^A)\,\mathrm{d}q^A$, which will be referred to as a non-holonomic force.

Now, we introduce the submanifold of $\mathcal{C}^2(x, y)$ which consists of those curves which are compatible with the constraint submanifold M

$$\tilde{\mathcal{C}}^2(x, y) = \{\tilde{c} \in \mathcal{C}^2(x, y) | \dot{\tilde{c}}(t) \in M \ \ \forall t \in [0, 1]\}.$$

Given a curve $\tilde{c} \in \tilde{\mathcal{C}}^2(x, y)$, the constraints allow us to consider a special vector subspace of $T_{\tilde{c}}\mathcal{C}^2(x, y)$

$$\mathcal{V}_{\tilde{c}} = \{X \in T_{\tilde{c}}\mathcal{C}^2(x, y) | S^*(\mathrm{d}\Phi_i)(\bar{X}) = 0 \ \ \forall i,$$

for all vector fields $\bar{X}$ on TQ along $\dot{\tilde{c}}$ such that $T\tau_Q(\bar{X}) = X\}$. Therefore, if $X = X^A(\partial/\partial q^A)$ we deduce that $X \in \mathcal{V}_{\tilde{c}}$ if and only if

$$X^A \frac{\partial \Phi_i}{\partial \dot{q}^A} = 0 \quad \forall i \tag{1}$$

along the curve $\tilde{c}$.

Next, we will describe the tangent space to $\tilde{\mathcal{C}}^2(x, y)$ at a curve $\tilde{c}$. If $\tilde{X} \in T_{\tilde{c}}\tilde{\mathcal{C}}^2(x, y)$, then there exists a family of curves $\tilde{c}_s(t)$ in $\tilde{\mathcal{C}}^2(x, y)$ passing through $\tilde{c}$ (say, $\tilde{c}_0 = \tilde{c}$) such that

$$\tilde{X}(t) = \left.\frac{\mathrm{d}\tilde{c}_s(t)}{\mathrm{d}s}\right|_{s=0}. \tag{2}$$

In local coordinates we have $\tilde{c}_s(t) = (q^A(t, s))$, and (2) becomes

$$X^A = \left.\frac{\partial q^A(t, s)}{\partial s}\right|_{s=0}, \tag{3}$$

where $\tilde{X} = X^A(\partial/\partial q^A)$. But the curves $\dot{\tilde{c}}_s(t)$ lie in M, so that

$$\Phi_i \left(q^A(t, s), \frac{\partial q^A(t, s)}{\partial t} \right) = 0, \tag{4}$$

and differentiating with respect to s we obtain

$$\frac{\partial q^A(t, s)}{\partial s} \frac{\partial \Phi_i}{\partial q^A} + \frac{\partial \dot{q}^A(t, s)}{\partial s} \frac{\partial \Phi_i}{\partial \dot{q}^A} = 0, \tag{5}$$

where the dot means derivative with respect to t. From (3) and (5) we deduce that $\tilde{X}$ has to satisfy the following equation:

$$X^A \frac{\partial \Phi_i}{\partial q^A} + (\mathrm{d}_T X^A) \frac{\partial \Phi_i}{\partial \dot{q}^A} = 0 \quad \forall i \tag{6}$$

along the curve $\tilde{c}$. Notice that in fact the derivative d_T is meaningful along a curve.

We define the functional $\mathcal{J}$ by

$$\mathcal{J} : \mathcal{C}^2(x, y) \to \mathbb{R},$$

$$c \mapsto \int_0^1 L(\dot{c}(t))\, \mathrm{d}t.$$

If $c \in \mathcal{C}^2(x, y)$ and $X \in T_c\mathcal{C}^2(x, y)$, a direct computation using integration by parts shows that (see [27])

$$\mathrm{d}\mathcal{J}(c)(X) = \int_0^1 \left(\frac{\partial L}{\partial q^A} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^A}\right)\right) X^A \,\mathrm{d}t.$$

2.1. Unconstrained systems

In this case, $M = TQ$, and then $S^*((TM)^0) = 0$. The Hamilton principle states that a curve $c \in \mathcal{C}^2(x, y)$ is a motion of the Lagrangian system defined by L if and only if c is a critical point of $\mathcal{J}$, i.e., $\mathrm{d}\mathcal{J}(c)(X) = 0$ for all $X \in T_c\mathcal{C}^2(x, y)$ which in turn is equivalent to the condition

$$\int_0^1 \left(\frac{\partial L}{\partial q^A} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^A}\right)\right) X^A \,\mathrm{d}t = 0$$

for all values of X^A. Therefore, we have

$$\int_0^1 \left(\frac{\partial L}{\partial q^A} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^A}\right)\right) fX^A \,\mathrm{d}t = 0$$

for all functions f defined along c. In particular, we can take

$$f = \left(\frac{\partial L}{\partial q^A} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^A}\right)\right) X^A,$$

and we deduce that

$$\left(\frac{\partial L}{\partial q^A} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^A}\right)\right) X^A = 0$$

for all X^A. The converse is obvious. Hence, c satisfies $\mathrm{d}\mathcal{J}(c)(X) = 0$ for all $X \in T_c\mathcal{C}^2(x, y)$ if and only if we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A} = 0, \quad 1 \leq A \leq n,$$

i.e., c satisfies the Euler–Lagrange equations.

2.2. Non-holonomic mechanics

The Hölder principle [2] states that a curve $\tilde{c} \in \tilde{\mathcal{C}}^2(x, y)$ is a motion if and only if it satisfies $\mathrm{d}\mathcal{J}(\tilde{c})(X) = 0$ for all $X \in \mathcal{V}_{\tilde{c}}$, i.e.,

$$\int_0^1 \left(\frac{\partial L}{\partial q^A} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^A}\right)\right) X^A \,\mathrm{d}t = 0$$

for all X^A satisfying Eq. (1). Since Eq. (1) is linear, we deduce that

$$\int_0^1 \left(\frac{\partial L}{\partial q^A} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^A}\right)\right) fX^A \,\mathrm{d}t = 0$$

for all functions f along $\tilde{c}$. As above, if we take

$$f = \left(\frac{\partial L}{\partial q^A} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}^A} \right) \right) X^A,$$

we deduce that $\tilde{c}$ is a motion if and only if

$$\left(\frac{\partial L}{\partial q^A} - \frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}^A} \right) \right) X^A = 0 \tag{7}$$

for all X^A satisfying Eq. (1), which is just the statement of D'Alembert's principle. If we denote by $\mathcal{F}_{nh}^0$ the annihilator of the space

$$\mathcal{F}_{nh} = \langle X^A \mid X^A \text{ satisfies}(1) \rangle,$$

then it is locally generated by the constraint forces, say

$$\mathcal{F}_{nh}^0 = \langle S^*(\mathrm{d}\Phi_i) \rangle.$$

Therefore, the 1-form $\delta(L) = ((\partial L/\partial q^A) - (\mathrm{d}/\mathrm{d}t)(\partial L/\partial \dot{q}^A))\,\mathrm{d}q^A$ is in $\mathcal{F}_{nh}^0$, i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}^A} \right) - \frac{\partial L}{\partial q^A} = -\lambda^i \frac{\partial \Phi_i}{\partial \dot{q}^A}, \quad 1 \le A \le n, \tag{8}$$

for some Lagrange multipliers $\lambda^1, \dots, \lambda^m$. Conversely, if $\delta(L) \in \mathcal{F}_{nh}^0$ along $\tilde{c}$, then $\tilde{c}$ is a motion for the non-holonomic problem.

Next, we will discuss the principle of virtual work.

Principle of virtual work The work done by the forces of constraint is zero on motions allowed by the constraints.

This statement cannot be derived from the non-holonomic equations of motion. Indeed, the principle of virtual work states that the work performed by the constraint force $(\lambda^i(\partial \Phi_i/\partial \dot{q}^A))\,\mathrm{d}q^A$ (which is obtained after the determination of the Lagrange multipliers) is zero, i.e.,

$$\lambda^i \frac{\partial \Phi_i}{\partial \dot{q}^A} \dot{q}^A = 0.$$

Therefore, if the Liouville vector field Δ is tangent to the constraint submanifold, i.e., $\Delta_{|M} \in TM$, our system will satisfy the principle of virtual work. When Δ is tangent to M we will say that the constraints are homogeneous. Clearly, a linear constraint is homogeneous.

2.3. Vakonomic mechanics

In vakonomic mechanics, a motion is a curve $\tilde{c} \in \tilde{\mathcal{C}}^2(x, y)$ such that $\mathrm{d}\mathcal{J}(\tilde{c})(\tilde{X}) = 0$ for all $\tilde{X} \in T_{\tilde{c}}\tilde{\mathcal{C}}^2(x, y)$. Using the Lagrange multipliers theorem in an infinite dimensional context we deduce (see [2,27]) that $\tilde{c}$ is a motion if and only if there exist m functions $\lambda^1, \dots, \lambda^m$ such that

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial L}{\partial \dot{q}^A} \right) - \frac{\partial L}{\partial q^A} = -\lambda^i \left(\frac{\mathrm{d}}{\mathrm{d}t} \left(\frac{\partial \Phi_i}{\partial \dot{q}^A} \right) - \frac{\partial \Phi_i}{\partial q^A} \right) - \frac{\mathrm{d}\lambda^i}{\mathrm{d}t} \frac{\partial \Phi_i}{\partial \dot{q}^A}, \quad 1 \le A \le n. \tag{9}$$

Indeed, we have to use the natural pairing

$$(\alpha, \tilde{X}) \mapsto \int_0^1 \langle \alpha, \tilde{X} \rangle \, \mathrm{d}t,$$

where α is a 1-form on T^2Q along $\tau_{21} : T^2Q \to TQ$ and $\tilde{X} \in T_{\tilde{c}}\tilde{\mathcal{C}}^2(x, y)$.

It should be noticed that Eq. (9) can be written as

$$(\delta L)_A = -\lambda^i (\delta \Phi_i)_A + (\mathrm{d}_T \lambda^i) \frac{\partial \Phi_i}{\partial \dot{q}^A}, \quad 1 \leq A \leq n. \tag{10}$$

An alternative approach to vakonomic mechanics is the following. We can prove that a curve $\tilde{c} = (q^A(t))$ in $\tilde{\mathcal{C}}^2(x, y)$ is a solution of the vakonomic equations if and only if there exist local functions $\lambda^1, \ldots, \lambda^m$ on TQ such that $\bar{c}(t) = (q^A(t), \lambda^i(t))$ is an extremal for the extended Lagrangian

$$\mathcal{L} = L + \lambda^i \Phi_i,$$

i.e., it satisfies the Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^A}\right) - \frac{\partial \mathcal{L}}{\partial q^A} = 0, \quad 1 \leq A \leq n$$

(see [2,27] for details).

3. Unconstrained mechanics: a geometrical approach

Let us recall the symplectic formulation for unconstrained mechanics (see [26]).

Using S and Δ we construct the Poincaré–Cartan 1-form $\alpha_L = S^*(\mathrm{d}L)$, the Poincaré–Cartan 2-form $\omega_L = -\mathrm{d}\alpha_L$, and the energy function $E_L = \Delta(L) - L$ associated with L. As we know, ω_L is symplectic if and only if L is regular. In such a case, the symplectic geometry allows us to derive the equations of motion in a geometrical way. Indeed, the equation

$$i_X \omega_L = \mathrm{d}E_L \tag{11}$$

has a unique solution Γ_L, which is the Hamiltonian vector field X_{E_L}; Γ_L is usually called the Euler–Lagrange vector field. Denote by $\flat_L : T(TQ) \to T^*(TQ)$ and $\sharp_L : T^*(TQ) \to T(TQ)$ the musical isomorphisms defined by ω_L, i.e., $\flat_L(Y) = i_Y \omega_L$, and $\sharp_L = \flat_L^{-1}$. In local coordinates we have

$$\flat_L\left(\frac{\partial}{\partial q^A}\right) = \left(\frac{\partial \hat{p}_A}{\partial q^B} - \frac{\partial \hat{p}_B}{\partial q^A}\right) \mathrm{d}q^B + \frac{\partial \hat{p}_A}{\partial \dot{q}^B} \mathrm{d}\dot{q}^B,$$

$$\flat_L\left(\frac{\partial}{\partial \dot{q}^A}\right) = -\frac{\partial \hat{p}_B}{\partial \dot{q}^A} \mathrm{d}q^B,$$

where $\hat{p}_A = (\partial L / \partial \dot{q}^A)$, $1 \leq A \leq n$, are the generalized momenta.

Therefore, Γ_L is a second order differential equation (SODE, for short), i.e., $S(\Gamma_L) = \Delta$. So, we have

$$\Gamma_L = \dot{q}^A \frac{\partial}{\partial q^A} + \xi^A(q, \dot{q}) \frac{\partial}{\partial \dot{q}^A},$$

which implies that the solutions of Γ_L (i.e., the projections of its integral curves onto Q) are the solutions of the Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A} = 0, \quad 1 \le A \le n.$$

If we contract Eq. (11) by Γ_L we have

$$0 = \Gamma_L(E_L),$$

which proves that the energy is a conserved quantity.

4. Non-holonomic mechanics: a geometrical approach

Assume that L is subjected to a constraint submanifold M of TQ locally defined by the equations $\Phi_i(q^A, \dot{q}^A) = 0$, $1 \le i \le m$. We denote by F_{nh} the vector bundle over M defined by prescribing its annihilator be $S^*((TM)^0)$, i.e., $F_{nh}^0 = S^*((TM)^0)$. Notice that $\flat_L(F_{nh}^\perp) = F_{nh}^0$, where $\perp$ is used to denote the complement with respect to the symplectic form ω_L.

Consider the following set of equations:

$$i_X \omega_L - \mathrm{d}E_L \in F_{nh}^0, \qquad X \in TM. \tag{12}$$

Henceforth, we will assume that the following conditions are fulfilled (see [36,37]):
1. codim M = rank F_{nh}^0 (*admissibility condition*),
2. $TM \cap F_{nh}^\perp = 0$ (*compatibility condition*).

The admissibility condition simply states that the 1-forms $\{S^*(\mathrm{d}\Phi_i)\}$ are linearly independent (see Section 2). The meaning of the compatibility condition will become clear below.

If X is a solution of Eq. (12) then it is a SODE. Moreover, if $c(t) = (q^A(t))$ is a solution of X we have that it satisfies the non-holonomic equations (8). Thus, Eq. (12) are the geometrical version of (8).

A solution of Eq. (12) will be of the form $X = \Gamma_L + \lambda^i Z_i$, where $Z_i = \sharp_L(S^*(\mathrm{d}\Phi_i))$. In order to determine the Lagrange multipliers λ^i we need to use the tangency condition $X(\Phi_i) = 0$, for all i. Thus, we get

$$0 = X(\Phi_j) = \Gamma_L(\Phi_j) + \lambda^i Z_i(\Phi_j).$$

Therefore, we obtain a system of linear equations. If the matrix $(Z_i(\Phi_j))$ is regular the system has a unique solution. This happens, for instance, if the Hessian matrix of L is positive or negative definite (see [21,24]).

If $(Z_i(\Phi_j))$ is regular then we have

$$\lambda^i = -\mathcal{C}^{ji}\Gamma_L(\Phi_j), \tag{13}$$

where $(\mathcal{C}^{ij})$ is the inverse matrix of $(Z_i(\Phi_j))$. In [21,24] it was proved that the regularity of these local matrices is equivalent to the compatibility condition. By a simple counting of dimensions we obtain a Withney sum decomposition

$$T(TQ)_{|M} = TM \oplus F_{nh}^{\perp}$$

with complementary projectors $\mathcal{P} : T(TQ)_{|M} \to TM$ and $\mathcal{Q} : T(TQ)_{|M} \to F_{nh}^{\perp}$ such that $\Gamma_{L,M} = \mathcal{P}(\Gamma_L)$ is the constrained dynamics (see [7,21,23,24]). We explicitly obtain

$$\Gamma_{L,M} = \mathcal{P}(\Gamma_L) = \Gamma_L - \mathcal{C}^{ij}\Gamma_L(\Phi_j)Z_i. \tag{14}$$

Remark 4.1. *If L is a Lagrangian of mechanical type, say $L = T - V$, where T is the kinetic energy derived from a Riemannian metric on Q and V is a potential energy, then L is always compatible with any constraint submanifold.*

Remark 4.2. *It should be noted that the determination of the Lagrange multipliers implies the knowledge of the dynamics as well as the knowledge of the constraint force $\lambda^i(\partial\Phi_i/\partial\dot{q}^A)\,\mathrm{d}q^A$.*

Assume that the non-holonomic system (L, M) is compatible and let $\Gamma_{L,M}$ be the solution of the equations of motion (12). If we contract the first equation in Eq. (12) by $\Gamma_{L,M}$, we obtain

$$\Gamma_{L,M}(E_L) = \lambda^i \frac{\partial\Phi_i}{\partial\dot{q}^A}\dot{q}^A,$$

since $\Gamma_{L,M}$ is a SODE. Therefore, we have the following result.

Theorem 4.3. *If the constraint submanifold M is homogeneous then the energy is conserved.*

5. Vakonomic mechanics: a geometrical approach

In this section we consider the vakonomic case.

Consider the projection $(\tau_{21})_{|\tau_{21}^{-1}(M)} : \tau_{21}^{-1}(M) \to M$, where $\tau_{21} : T^2Q \to TQ$ is the canonical bundle projection. We denote by F_{vk} the real vector space defined by prescribing its annihilator consists of the 1-forms

$$\{\lambda^i\alpha_i + (\mathrm{d}_T\lambda^i)S^*(\mathrm{d}\Phi_i) \mid \lambda^1, \ldots, \lambda^m \in C^\infty(TQ)\},$$

where

$$\alpha_i = \mathrm{d}_T(S^*(\mathrm{d}\Phi_i)) - \mathrm{d}\Phi_i, \quad 1 \le i \le m.$$

In local coordinates we obtain

$$\alpha_i = \left(\mathrm{d}_T\left(\frac{\partial\Phi_i}{\partial\dot{q}^A}\right) - \frac{\partial\Phi_i}{\partial q^A}\right)\mathrm{d}q^A, \quad S^*(\mathrm{d}\Phi_i) = \frac{\partial\Phi_i}{\partial\dot{q}^A}\,\mathrm{d}q^A.$$

In fact, α_i is just the 1-form $\alpha_i = -\delta\Phi_i$ on T^2Q.

The above local expressions show that the set F_{vk} consists of semibasic 1-forms along the canonical projection $(\tau_{21})_{|\tau_{21}^{-1}(M)} : \tau_{21}^{-1}(M) \to M$.

Notice that F_{vk} is a global object. Indeed, define an operator from 1-forms on TQ into 1-forms on T^2Q by

$$\mathcal{K} = \mathrm{d}_T S^* - \tau_{21}^*,$$

i.e., if α is a 1-form on TQ then $\mathcal{K}(\alpha) = \mathrm{d}_T(S^*(\alpha)) - \tau_{21}^*\alpha$ is a 1-form on T^2Q. Now, a simple computation shows that

$$F_{vk}^0 = \mathcal{K}((TM)^0).$$

Consider the following set of equations

$$i_X\omega_L - \mathrm{d}E_L \in F_{vk}^0, \qquad X \in TM. \tag{15}$$

Let us say some words about the meaning of Eq. (15). X is a vector field along τ_{21}, i.e., $X(z) \in T_{\tau_{21}(z)}(TQ)$, for any $z \in \tau_{21}^{-1}(M)$. Thus, $i_X\omega_L - \mathrm{d}E_L$ is a 1-form along τ_{21}.

If X is a solution of Eq. (15) then it is a SODE, since the elements in F_{vk}^0 are semibasic 1-forms. Moreover, if $c(t) = (q^A(t))$ is a solution of X we have that it satisfies the vakonomic equation (9). Thus, Eq. (15) are the geometrical version of (9).

Any solution of Eq. (15) will be of the form $X = \Gamma_L + \lambda^i U_i + (\mathrm{d}_T\lambda^i)Z_i$, where $U_i = \sharp_L(\alpha_i)$ and $Z_i = \sharp_L(S^*(\mathrm{d}\Phi_i))$. Now, using the tangency condition, we get

$$0 = X(\Phi_j) = \Gamma_L(\Phi_j) + \lambda^i U_i(\Phi_j) + (\mathrm{d}_T\lambda^i)Z_i(\Phi_j).$$

Let X be a solution of Eq. (15), hence it is a SODE. A curve $c(t) = (q^A(t))$ in Q is a solution of X if and only if

$$\frac{\mathrm{d}q^A}{\mathrm{d}t} = \dot{q}^A, \qquad \frac{\mathrm{d}^2q^A}{\mathrm{d}t^2} = X^A\left(q^B, \dot{q}^B, \ddot{q}^B, \lambda^i, \frac{\mathrm{d}\lambda^i}{\mathrm{d}t}\right), \tag{16}$$

where $X = \dot{q}^A(\partial/\partial q^A) + X^A(q, \dot{q}, \ddot{q}, \lambda^i, \dot{\lambda}^i)(\partial/\partial\dot{q}^A)$. Eq. (16) are in general implicit differential equations.

Now, assume that, as in the precedent section, the matrix $(Z_i(\Phi_j))$ is regular. Then we have

$$\mathrm{d}_T\lambda^i = -\mathcal{C}^{ji}U_k(\Phi_j)\lambda^k - \mathcal{C}^{ji}\Gamma_L(\Phi_j), \tag{17}$$

where $(\mathcal{C}^{ij})$ is the inverse matrix of $(Z_i(\Phi_j))$. As for non-holonomic mechanics, the geometrical characterization of the regularity of the local matrices $(Z_i(\Phi_j))$ is the condition

$$F_{nh}^{\perp} \cap TM = 0.$$

Eq. (17) is a differential equation involving the Lagrange multipliers. We thus have to give initial conditions for them in order to solve Eqs. (16) and (17). Note the difference with respect to the non-holonomic case, in which one obtains the Lagrange multipliers by an algebraic procedure.

Coming back to Eq. (16), if the associated non-holonomic system is compatible we can implement them as follows:

$$\frac{\mathrm{d}q^A}{\mathrm{d}t} = \dot{q}^A, \qquad \frac{\mathrm{d}^2 q^A}{\mathrm{d}t^2} = X^A\left(q^B, \dot{q}^B, \ddot{q}^B, \lambda^i, \frac{\mathrm{d}\lambda^i}{\mathrm{d}t}\right),$$
$$\frac{\mathrm{d}\lambda^i}{\mathrm{d}t} = -\mathcal{C}^{ji} U_k(\Phi_j)\lambda^k - \mathcal{C}^{ji}\Gamma_L(\Phi_j).$$

Assume that X is a solution of Eq. (15). If we contract the first equation in Eq. (15) by X we obtain

$$-X(E_L) = \lambda^i \alpha_i(X) + (\mathrm{d}_T\lambda^i) S^*(\mathrm{d}\Phi_i)(X).$$

Taking into account that X is a SODE we have

$$\begin{aligned} -X(E_L) &= \lambda^i\, \mathrm{d}_T\left(\frac{\partial\Phi_i}{\partial\dot{q}^A}\right)\dot{q}^A - \lambda^i \frac{\partial\Phi_i}{\partial q^A}\dot{q}^A + \mathrm{d}_T(\lambda^i)\frac{\partial\Phi_i}{\partial\dot{q}^A}\dot{q}^A \\ &= \lambda^i\, \mathrm{d}_T\left(\frac{\partial\Phi_i}{\partial\dot{q}^A}\right)\dot{q}^A + \lambda^i \frac{\partial\Phi_i}{\partial\dot{q}^A}\ddot{q}^A + \mathrm{d}_T(\lambda^i)\frac{\partial\Phi_i}{\partial\dot{q}^A}\dot{q}^A, \end{aligned}$$

since

$$\dot{q}^A \frac{\partial\Phi_i}{\partial q^A} + \ddot{q}^A \frac{\partial\Phi_i}{\partial\dot{q}^A} = 0.$$

Therefore, we deduce that

$$-X(E_L) = \mathrm{d}_T\left(\lambda^i \frac{\partial\Phi_i}{\partial\dot{q}^A}\dot{q}^A\right).$$

Then, we have proved the following result.

Theorem 5.1. *The energy of the vakonomic system* (M, L) *is a conserved quantity if and only if the work done by the "reaction non-holonomic forces"* $\lambda^i(\partial\Phi_i/\partial\dot{q}^A)\,\mathrm{d}q^A$ *is a constant of the motion*.

Remark 5.2. *We can compare with the result for non-holonomic mechanics. If the constraint submanifold M is homogeneous, then the energy* E_L *is a conserved quantity for both mechanical systems, say non-holonomic and vakonomic*.

Therefore, the following could be considered as a principle of virtual work for vakonomic mechanics:

Principle of virtual work for vakonomic mechanics The work done by any non-holonomic force is constant along the motions allowed by the constraints.

6. An example: a skate on an inclined plane

Consider a skate on an inclined plane Π with Cartesian coordinates x, y. We assume that the y-axis is horizontal, while the x-axis is directed downward (see [2]). Denote by (x, y) the coordinates of the point of contact of the skate with Π, and let ϕ be the angle measured from the x-axis. The Lagrangian function is

$$L = \tfrac{1}{2}(\dot{x}^2 + \dot{y}^2 + \dot{\phi}^2) + x$$

with an appropriate choice of unities. The system is also subjected to the constraint

$$\Phi = \dot{x}\sin\phi - \dot{y}\cos\phi.$$

A direct computation shows that

$$\Gamma_L = \dot{x}\frac{\partial}{\partial x} + \dot{y}\frac{\partial}{\partial y} + \dot{\phi}\frac{\partial}{\partial \phi} + \frac{\partial}{\partial \dot{x}}.$$

6.1. *Non-holonomic equations of motion*

From (13) we deduce that $\lambda = \dot{x}\dot{\phi}\cos\phi + \dot{y}\dot{\phi}\sin\phi + \sin\phi$, and then (14) implies that

$$\begin{aligned} X = {} & \dot{x}\frac{\partial}{\partial x} + \dot{y}\frac{\partial}{\partial y} + \dot{\phi}\frac{\partial}{\partial \phi} + (1 - \sin\phi(\dot{\phi}(\dot{x}\cos\phi + \dot{y}\sin\phi) + \sin\phi))\frac{\partial}{\partial \dot{x}} \\ & + \cos\phi(\dot{\phi}(\dot{x}\cos\phi + \dot{y}\sin\phi) + \sin\phi)\frac{\partial}{\partial \dot{y}}. \end{aligned}$$

In other words, a motion $(x(t), y(t), \phi(t))$ satisfies the following system of second order differential equations:

$$\begin{aligned} &\ddot{x} = 1 - \sin\phi(\dot{\phi}(\dot{x}\cos\phi + \dot{y}\sin\phi) + \sin\phi), \\ &\ddot{y} = \cos\phi(\dot{\phi}(\dot{x}\cos\phi + \dot{y}\sin\phi) + \sin\phi), \qquad \ddot{\phi} = 0, \end{aligned}$$

which can be explicitly integrated (see [2]).

6.2. *Vakonomic equations of motion*

In this case, (17) yields

$$\dot{\lambda} = \dot{\phi}(\dot{x}\cos\phi + \dot{y}\sin\phi) + \sin\phi. \tag{18}$$

Therefore, a solution has the following form:

$$\begin{aligned} X = {} & \dot{x}\frac{\partial}{\partial x} + \dot{y}\frac{\partial}{\partial y} + \dot{\phi}\frac{\partial}{\partial \phi} + (1 - \lambda\dot{\phi}\cos\phi - \dot{\lambda}\sin\phi)\frac{\partial}{\partial \dot{x}} \\ & + (-\lambda\dot{\phi}\sin\phi + \dot{\lambda}\cos\phi)\frac{\partial}{\partial \dot{y}} + \lambda(\dot{x}\cos\phi + \dot{y}\sin\phi)\frac{\partial}{\partial \dot{\phi}}, \end{aligned}$$

where λ is given by (18).

Hence, a motion $(x(t), y(t), \phi(t))$ satisfies the following system of second order differential equations:

$$\ddot{x} = 1 - \lambda\dot{\phi}\cos\phi - \dot{\lambda}\sin\phi, \qquad \ddot{y} = -\lambda\dot{\phi}\sin\phi + \dot{\lambda}\cos\phi,$$
$$\ddot{\phi} = \lambda(\dot{x}\cos\phi + \dot{y}\sin\phi), \qquad \dot{\lambda} = \dot{\phi}(\dot{x}\cos\phi + \dot{y}\sin\phi) + \sin\phi.$$

Remark 6.1. *Notice that in both cases the energy is conserved since the constraints are linear in the velocities.*

7. A unified geometrical approach

The purpose of this section is to describe a general geometrical setting which includes non-holonomic and vakonomic mechanics as particular cases.

Let $L : TQ \to \mathbb{R}$ be a regular Lagrangian function, and M a submanifold of codimension m of TQ.

Let $\tilde{F}^0$ be a set of semibasic 1-forms along the projection $(\tau_{21})_{|\tau_{21}^{-1}(M)} : \tau_{21}^{-1}(M) \to M$, and denote by $\tilde{F}$ its annihilator. In other words, for each $z \in T^2Q$ such that $y = \tau_{21}(z) \in M$, $\tilde{F}_z$ is a subset of $T_y(TQ)$. Therefore, if $\gamma \in \tilde{F}^o$, we have $\gamma = \gamma_A(q^B, \dot{q}^B, \ddot{q}^B)\, \mathrm{d}q^A$.

Consider the following system of equations:

$$i_X\omega_L - \mathrm{d}E_L \in \tilde{F}^0, \qquad X \in TM. \tag{19}$$

One can imagine that the Lagrangian L is subjected to the forces given by $\tilde{F}^0$ and some constraints given by M.

A solution X of Eq. (19), if it exists, has to be of the form

$$X = \Gamma_L + Y,$$

where $Y = \sharp_L(\gamma)$, and $\gamma \in \tilde{F}^0$. Now, we impose the second equation in Eq. (19) (the tangency condition) and we get

$$0 = X(\Phi_j) = \Gamma_L(\Phi_j) + Y(\Phi_j).$$

A direct computation shows that the solutions of a solution X of Eq. (19) satisfy the following equations of motion:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A} = \gamma_A,$$

where $\gamma = \gamma_A\, \mathrm{d}q^A$.

In order to recover the non-holonomic and vakonomic cases, we only need to take as $\tilde{F}^0$ either the natural lift of F^0_{nh} to T^2Q or F^0_{vk}.

8. Čaplygin vakonomic systems: reduction of the vakonomic equations

We will consider a Lagrangian function subjected to constraints given by a connection.

Suppose that Q is a principal G-bundle over a manifold $\bar{Q}$ with projection $\rho : Q \to \bar{Q}$, and let Γ be a principal connection in $\rho : Q \to \bar{Q}$. The Lagrangian $L : TQ \to \mathbb{R}$ is supposed to be G-invariant and in addition it is subjected to the constraints given by the horizontal distribution $\mathcal{H}$ of Γ, i.e., the allowable motions have to be horizontal curves with respect to that connection. Thus, the constraint submanifold M is just the total space of the vector bundle $\mathcal{H} \to Q$, with $TQ = \mathcal{H} \oplus V\rho$, where $V\rho$ denotes the vertical subbundle with respect to the projection ρ. We denote by $\mathbf{h} : TQ \to \mathcal{H}$ the horizontal projector. We will assume that the resultant non-holonomic system is compatible in the sense of Section 4.

Take bundle coordinates $(q^A) = (q^a, q^i)$, $1 \le a \le n - m$, $1 \le i \le m$, $n = \dim Q$. The horizontal distribution is locally spanned by the local vector fields

$$H_a = \left(\frac{\partial}{\partial q^a}\right)^{\mathrm{H}} = \frac{\partial}{\partial q^a} - \Gamma^i_a(q^A)\frac{\partial}{\partial q^i},$$

where Y^{H} stands for the horizontal lift to Q of a vector field Y on $\bar{Q}$, and $\Gamma^i_a = \Gamma^i_a(q^b, q^j)$ are the Christoffel components of Γ. Thus, we obtain a local basis of vector fields on Q

$$\left\{H_a, V_i = \frac{\partial}{\partial q^i}\right\}.$$

Its dual basis of 1-forms is

$$\{\eta_a = \mathrm{d}q^a, \eta_i = \Gamma^i_a\,\mathrm{d}q^a + \mathrm{d}q^i\},$$

and the constraints are

$$\Phi_i = \Gamma^i_a\dot{q}^a + \dot{q}^i.$$

We deduce that

$$F^0_{nh} = \mathrm{span}\,\{\eta_i\}, \qquad F^0_{vk} = \{\lambda^i\alpha_i + \mathrm{d}_T(\lambda^i)\eta_i\},$$

where $\alpha_i = (\delta\Phi_i)_A\,\mathrm{d}q^A$. A direct computation shows that

$$\alpha_i = \left[\mathrm{d}_T(\Gamma^i_a) - \frac{\partial\Gamma^i_b}{\partial q^a}\dot{q}^b\right]\mathrm{d}q^a - \frac{\partial\Gamma^i_b}{\partial q^j}\dot{q}^b\,\mathrm{d}q^j.$$

The curvature of Γ is the tensor field of type (1,2) on Q given by $R = \frac{1}{2}[\mathbf{h}, \mathbf{h}]$. Since

$$\mathbf{h}\left(\frac{\partial}{\partial q^a}\right) = \frac{\partial}{\partial q^a} - \Gamma^i_a\frac{\partial}{\partial q^i}, \qquad \mathbf{h}\left(\frac{\partial}{\partial q^i}\right) = 0,$$

we obtain

$$R\left(\frac{\partial}{\partial q^a}, \frac{\partial}{\partial q^b}\right) = R^i_{ab}\frac{\partial}{\partial q^i},$$

with

$$R^i_{ab} = \frac{\partial\Gamma^i_a}{\partial q^b} - \frac{\partial\Gamma^i_b}{\partial q^a} + \Gamma^j_a\frac{\partial\Gamma^i_b}{\partial q^j} - \Gamma^j_b\frac{\partial\Gamma^i_a}{\partial q^j}.$$

As we know, Γ is flat if and only if the curvature R identically vanishes and, in this case, the constrained system is holonomic.

Since L is G-invariant one can define a Lagrangian function $L^* : T\bar{Q} \to \mathbb{R}$ given by

$$L^*(Y) = L((Y^H)_q)$$

for any $Y \in T_{\bar{q}}\bar{Q}$, where q is an arbitrary point in the fiber over $\bar{q}$. In local coordinates we have

$$L^*(q^a, \dot{q}^a) = L(q^a, q^i, \dot{q}^a, -\Gamma^i_a \dot{q}^a).$$

Since L^* does not depend on q^i we deduce that

$$\frac{\partial L}{\partial q^i} = \frac{\partial L}{\partial \dot{q}^j} \frac{\partial \Gamma^j_a}{\partial q^i} \dot{q}^a. \tag{20}$$

The vakonomic equations for L are the following:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^a}\right) - \frac{\partial L}{\partial q^a} = -\lambda^i \left(\frac{\mathrm{d}}{\mathrm{d}t}(\Gamma^i_a) - \frac{\partial \Gamma^i_b}{\partial q^a}\dot{q}^b\right) - \dot{\lambda}^i \Gamma^i_a,$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}^i}\right) - \frac{\partial L}{\partial q^i} = \lambda^j \frac{\partial \Gamma^j_a}{\partial q^i}\dot{q}^a - \dot{\lambda}^i.$$

After some calculations, and using (20) we obtain that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L^*}{\partial \dot{q}^a}\right) - \frac{\partial L^*}{\partial q^a} = -\left(\lambda^i + \frac{\partial L}{\partial \dot{q}^i}\right)\dot{q}^b R^i_{ab},$$

where the Lagrange multipliers satisfy Eq. (17).

The term $(\partial L/\partial \dot{q}^i)\dot{q}^b R^i_{ab}$ was intrinsically defined in [24] as follows. Define a 1-form $\alpha_{L,\Gamma}$ on $T\bar{Q}$ by putting

$$(\alpha_{L,\Gamma})_u(U) = -(\alpha_L)_x(\tilde{X})$$

for any $U \in T_u(T\bar{Q})$, for any $u \in T_{\bar{q}}\bar{Q}$, where $\tilde{X} \in T_x(TQ)$ is a tangent vector which projects onto the tangent vector $\gamma_u(U) = R((u^H)_q, (T\tau_{\bar{Q}}(U))^H_q) \in T_q Q$, $\rho(q) = \bar{q}$, and $x \in M$ with $\tau_Q(x) = q$. In local coordinates, we get

$$\alpha_{L,\Gamma} = \left(\frac{\partial L}{\partial \dot{q}^i}\dot{q}^b R^i_{ab}\right) \mathrm{d}q^a.$$

Therefore, $\alpha_{L,\Gamma}$ encodes the curvature of the connection.

Let us now examine the term $\lambda^i \dot{q}^b R^i_{ab}$. Using the same notations as above we have a linear mapping

$$\gamma_u : T_u(T\bar{Q}) \to V_q\rho,$$

where $V_q\rho$ denotes the vector subspace of $T_q Q$ consisting of the vertical vectors with respect to ρ. From now on we will assume that the Lie group G is abelian. In this case, we have a natural identification $V_q\rho \cong \mathfrak{g}$ which induces another one between the corresponding dual spaces: $V_q^*\rho \cong \mathfrak{g}^*$. Therefore, the transpose mapping to γ_u can be understood as a linear mapping $\gamma_u^* : \mathfrak{g}^* \to T_u^*(T\bar{Q})$. Then we can define a real vector space $\tilde{F}_{vk}$ on $T\bar{Q}$ by prescribing its annihilator to be $\mathrm{Im}\,\gamma^*$. A direct computation proves that $\tilde{F}^0_{vk}$ is locally generated by the 1-forms $R^i_{ab}\dot{q}^b\,\mathrm{d}q^a$.

Since G is abelian, we have that Γ^i_a is G-invariant, from which deduce that all the coefficients involved in Eq. (17) are G-invariant. Therefore, we could look for G-invariant Lagrange multipliers. We also remark that if λ is G-invariant then $\mathrm{d}_T\lambda$ is so also, where now the action of G on T^2Q is the natural prolongation. Thus, one could consider a family $\bar{\lambda}^i$ of Lagrange multipliers satisfying the differential equation

$$\mathrm{d}_T\bar{\lambda}^i = -\bar{\mathcal{C}}^{ji}\overline{U_k(\Phi_j)}\bar{\lambda}^k - \bar{\mathcal{C}}^{ji}\overline{\Gamma_L(\Phi_j)}, \tag{21}$$

where the bar over a term means that we are considering its projection onto $T\bar{Q}$. Then the lifts of $\bar{\lambda}^i$ to TQ would satisfy Eq. (17). Conversely, each G-invariant solution of Eq. (17) projects onto a solution of Eq. (21).

Now, take a Lagrangian system with Lagrangian function L and external force $\alpha_{L,\Gamma}$. Consider the following motion equation:

$$i_Y\,\omega_{L^*} - \mathrm{d}E_{L^*} - \alpha_{L,\Gamma} \in \tilde{F}^0_{vk} \tag{22}$$

on $T\bar{Q}$.

The above discussion can be summarized as follows.

Theorem 8.1. *If G is abelian, then the vakonomic Čaplygin system* (L,Γ) *is equivalent to the system defined by the Lagrangian* L^* *and external force* $\alpha_{L,\Gamma} + \beta$, *with* $\beta \in \tilde{F}^0_{vk}$, *plus the differential equation* (21). *This means that the solutions of both systems are related by projection and lifting*.

Indeed, if Y is a solution of Eq. (22) with Lagrange multipliers $\bar{\lambda}^i$ then its horizontal lift Y^H to TQ is a solution of Eq. (15) with Lagrange multipliers $\lambda^i = \bar{\lambda}^i \circ T\rho$. Conversely, if X is a G-invariant solution of Eq. (15) then its projection $T\rho(X)$ is a solution of Eq. (22).

Remark 8.2. *We have developed a reduction procedure which works as follows. First, we solve Eq.* (21) *and obtain the Lagrange multipliers* $\bar{\lambda}^i$. *Next, we solve the reduced dynamics Y, and finally we lift Y to TQ by using the tangent connection* $T\Gamma$ *in the principal bundle* $TQ \to T\bar{Q}$ (*see* [11] *for tangent prolongations of principal connections*).

8.1. An example: the rolling disk

Consider a rolling disk of radius R and mass m constrained to remain vertical on a plane. We introduce coordinates $(x, y, \theta_1, \theta_2)$ in the configuration manifold $Q = \mathbb{R}^2 \times S^1 \times S^1$, where x, y are the Cartesian coordinates of the center of mass, θ_1 the angle between the tangent of the disk at the point of contact and the axis x and θ_2 is the angle given by a fixed diameter and the vertical.

The system is described by a Lagrangian

$$L = \tfrac{1}{2}m(\dot{x}^2 + \dot{y}^2) + \tfrac{1}{2}I_1\dot{\theta}_1^2 + \tfrac{1}{2}I_2\dot{\theta}_2^2,$$

where I_1 and I_2 are the moments of inertia, and the constraints

$$\Phi_1 = \dot{x} - R\dot{\theta}_2\cos\theta_1, \qquad \Phi_2 = \dot{y} - R\dot{\theta}_2\sin\theta_1.$$

The vakonomic equations of motion are

$$m\ddot{x} = -\dot{\lambda}^1, \qquad m\ddot{y} = -\dot{\lambda}^2, \qquad I_1\ddot{\theta}_1 = \lambda^1\dot{\theta}_2 R\sin\theta_1 - \lambda^2\dot{\theta}_2 R\cos\theta_1,$$
$$(I_2 + mR^2)\ddot{\theta}_2 = -\lambda^1\dot{\theta}_1 R\sin\theta_1 + \lambda^2\dot{\theta}_1 R\cos\theta_1,$$

with

$$\dot{\lambda}^1 = -\frac{R^2m\dot{\theta}_1\sin\theta_1\cos\theta_1}{I_2 + R^2m}\lambda^1 - \frac{R^2m\dot{\theta}_1\cos^2\theta_1}{I_2 + R^2m}\lambda^2 + Rm\dot{\theta}_1\dot{\theta}_2\sin\theta_1,$$
$$\dot{\lambda}^2 = -\frac{R^2m\dot{\theta}_1\sin^2\theta_1}{I_2 + R^2m}\lambda^1 + \frac{R^2m\dot{\theta}_1\sin\theta_1\cos\theta_1}{I_2 + R^2m}\lambda^2 - Rm\dot{\theta}_1\dot{\theta}_2\cos\theta_1.$$

Next, we will apply the reduction method discussed above.

$\mathbb{R}^2 \times S^1 \times S^1$ is a principal $\mathbb{R}^2$-bundle over $\mathbb{T}^2 = S^1 \times S^1$ with projection $\rho : \mathbb{R}^2 \times \mathbb{T}^2 \to \mathbb{T}^2$, $\rho(x, y, \theta_1, \theta_2) = (\theta_1, \theta_2)$. A principal connection Γ is defined by prescribing its horizontal distribution to be given by

$$\mathcal{H}^0 = \mathrm{span}\,\{\mathrm{d}x - R\cos\theta_1\,\mathrm{d}\theta_2, \mathrm{d}y - R\sin\theta_1\,\mathrm{d}\theta_2\}.$$

So (L, Γ) is a Čaplygin system.

The Christoffel components are

$$\Gamma_1^x = \Gamma_1^y = 0, \qquad \Gamma_2^x = -R\cos\theta_1, \qquad \Gamma_2^y = -R\sin\theta_1$$

with the obvious notations. Therefore, the reduced Lagrangian is

$$L^*(\theta_1, \theta_2, \dot{\theta}_1, \dot{\theta}_2) = \tfrac{1}{2}I_1\dot{\theta}_1^2 + \tfrac{1}{2}(I_2 + mR^2)\dot{\theta}_2^2.$$

Since the curvature of Γ is given by

$$R_{12}^x = -R_{21}^x = -R\sin\theta_1, \qquad R_{12}^y = -R_{21}^y = R\cos\theta_1,$$

the other components being zero, we deduce that $\alpha_{\mathrm{L},\Gamma}$ identically vanishes.

A long but straightforward computation shows that the reduced vakonomic equations of motion are

$$I_1\ddot{\theta}_1 = \bar{\lambda}^1\dot{\theta}_2 R\sin\theta_1 - \bar{\lambda}^2\dot{\theta}_2 R\cos\theta_1,$$
$$(I_2 + mR^2)\ddot{\theta}_2 = -\bar{\lambda}^1\dot{\theta}_1 R\sin\theta_1 + \bar{\lambda}^2\dot{\theta}_1 R\cos\theta_1,$$

with

$$\dot{\bar{\lambda}}^1 = -\frac{R^2 m\dot{\theta}_1 \sin\theta_1 \cos\theta_1}{I_2 + R^2 m}\bar{\lambda}^1 - \frac{R^2 m\dot{\theta}_1 \cos^2\theta_1}{I_2 + R^2 m}\bar{\lambda}^2 + Rm\dot{\theta}_1\dot{\theta}_2 \sin\theta_1,$$
$$\dot{\bar{\lambda}}^2 = -\frac{R^2 m\dot{\theta}_1 \sin^2\theta_1}{I_2 + R^2 m}\bar{\lambda}^1 + \frac{R^2 m\dot{\theta}_1 \sin\theta_1 \cos\theta_1}{I_2 + R^2 m}\bar{\lambda}^2 - Rm\dot{\theta}_1\dot{\theta}_2 \cos\theta_1.$$

Acknowledgements

This work has been partially supported through grants DGICYT (Spain) (Projects PB97-1257 and PB97-1487) and UNED (Spain). ML and DMD wish to express his gratitude for the hospitality offered to them in the Departamento de Matemática Fundamental (University of La Laguna) where part of this work was done.

References

[1] P. Aimé, Sur la dynamique des systèmes mécaniques, Ann. Inst. Henri Poincaré 64 (1996) 153–176.
[2] V.I. Arnold, Dynamical Systems, Vol. III, Springer, New York, 1988.
[3] L. Bates, H. Graumann, C. MacDonnell, Examples of gauge conservation laws in nonholonomic systems, Rep. Math. Phys. 37 (3) (1996) 295–308.
[4] L. Bates, J. Śniatycki, Nonholonomic reduction, Rep. Math. Phys. 32 (1) (1992) 99–115.
[5] A.M. Bloch, P.S. Krishnaprasad, J.E. Marsden, R.M. Murray, Nonholonomic mechanical systems with symmetry, Arch. Rational Mech. Anal. 136 (1996) 21–99.
[6] F. Cantrijn, M. de León, D. Martín de Diego, The momentum equation for non-holonomic systems with symmetry, in: Proceedings of the National Conference on Theoretical and Applied Mechanics, Leuven, May 22–23, 1997, pp. 31–34.
[7] F. Cantrijn, M. de León, J.C. Marrero, D. Martín de Diego, Reduction of nonholonomic mechanical systems with symmetries, in: Proceedings of the Workshop on Non-Holonomic Constraints in Dynamics, Calgary, August 26–29, 1997; Rep. Math. Phys. 42 (1–2) (1998) 25–45.
[8] F. Cantrijn, M. de León, D. Martín de Diego, On pseudo-Poisson structures in nonholonomic mechanics, Nonlinearity 12 (1999) 721–737.
[9] J.F. Cariñena, M.F. Rañada, Lagrangians systems with constraints: a geometric approach to the method of Lagrange multipliers, J. Phys. A 26 (1993) 1335–1351.
[10] J.F. Cariñena, M.F. Rañada, Comments on the presymplectic formalism and the theory of regular Lagrangians with constraints, J. Phys. A 28 (1995) L91–L97.
[11] L.A. Cordero, C.T.J. Dodson, M. de León, Differential Geometry of Frame Bundles, Kluwer Academic Publishers, Dordrecht, 1989.
[12] R. Cushman, D. Kemppainen, J. Śniatycki, L. Bates, Geometry of nonholonomic constraints, Rep. Math. Phys. 36 (2/3) (1995) 275–286.
[13] P. Dazord, Mécanique Hamiltonienne en présence de constraintes, Illinois J. Math. 38 (1) (1994) 148–175.
[14] A. Ibort, M. de León, G. Marmo, D. Martín de Diego, Non-holonomic constrained systems as implicit differential equations, in: Geometrical Structures for Physical Theories, I, Rend. Sem. Mat. Univ. Pol. Torino 54 (3) (1996) 295–317.
[15] A. Ibort, M. de León, J.C. Marrero, D. Martín de Diego, Dirac brackets in constrained dynamics, Forschritte der Physik 47 (5) (1999) 459–492.
[16] J. Koiller, Reduction of some classical non-holonomic systems with symmetry, Arch. Rational Mech. Anal. 118 (1992) 113–148.
[17] W.S. Koon, J.E. Marsden, The Hamiltonian and Lagrangian approaches to the dynamics of nonholonomic systems, Rep. Math. Phys. 40 (1997) 21–62.
[18] W.S. Koon, J.E. Marsden, Poisson reduction of nonholonomic mechanical systems with symmetry, in: Proceedings of the Workshop on Non-Holonomic Constraints in Dynamics, Calgary, August 26–29, 1997; Rep. Math. Phys. 42 (1/2) (1998) 101–134.

[19] V.V. Kosloz, Realization of nonintegrable constraints in classical mechanics, Dokl. Akad. Nauk SSSR 272 (3) (1983) 550–554 (Russian); English transl.: Sov. Phys. Dokl. 28 (9) (1983) 735–737.
[20] O. Krupková, Mechanical systems with non-holonomic constraints, J. Math. Phys. 38 (1997) 5098–5126.
[21] M. de León, J.C. Marrero, D. Martín de Diego, Mechanical systems with non-linear constraints, Int. J. Theoret. Phys. 36 (4) (1997) 973–989.
[22] M. de León, J.C. Marrero, D. Martín de Diego, Time-dependent mechanical systems with non-linear constraints, in: J. Szenthe (Ed.), Proceedings of the Conference on Differential Geometry, Budapest, Hungary, July 21–27, 1996, Kluwer Academic Publishers, Dordrecht, 1999, pp. 221–234.
[23] M. de León, D. Martín de Diego, Solving non-holonomic Lagrangian dynamics in terms of almost product structures, Extracta Mathematicae 11 (2) (1996) 325–347.
[24] M. de León, D. Martín de Diego, On the geometry of non-holonomic Lagrangian systems, J. Math. Phys. 37 (7) (1996) 3389–3414.
[25] M. de León, P.R. Rodrigues, Generalized Classical Mechanics and Field Theory, North-Holland Mathematical Series, Vol. 112, North-Holland, Amsterdam, 1985.
[26] M. de León, P.R. Rodrigues, Methods of Differential Geometry in Analytical Mechanics, North-Holland Mathematical Series, Vol. 152, North-Holland, Amsterdam, 1989.
[27] A.D. Lewis, R.M. Murray, Variational principles for constrained systems: theory and experiment, Int. J. Nonlinear Mech. 30 (6) (1995) 793–815.
[28] C.M. Marle, Reduction of constrained mechanical systems and stability of relative equilibria, Commun. Math. Phys. 174 (1995) 295–318.
[29] C.M. Marle, Various approaches to conservative and nonconservative nonholonomic systems, in: Proceedings of the Workshop on Non-Holonomic Constraints in Dynamics, Calgary, August 26–29, 1997; Rep. Math. Phys. 42 (1/2) (1998) 211–229.
[30] E. Massa, E. Pagani, Classical dynamics of non-holonomic systems: a geometric approach, Ann. Inst. Henri Poincaré 55 (1991) 511–544.
[31] E. Massa, E. Pagani, A new look at classical mechanics of constrained systems, Ann. Inst. Henri Poincaré 66 (1997) 1–36.
[32] W. Sarlet, F. Cantrijn, D.J. Saunders, A geometrical framework for the study of non-holonomic Lagrangian systems, J. Phys. A 28 (1995) 3253–3268.
[33] D.J. Saunders, W. Sarlet, F. Cantrijn, A geometrical framework for the study of non-holonomic Lagrangian systems: II, J. Phys. A 29 (1996) 4265–4274.
[34] J. Śniatycki, Non-holonomic Noether theorem and reduction of symmetries, in: Proceedings of the Workshop on Non-Holonomic Constraints in Dynamics, Calgary, August 26–29, 1997, Rep. Math. Phys. 42 (1/2) (1998) 5–23.
[35] W.M. Tulczyjew, Sur la différentielle de Lagrange, C. R. Acad. Sci. Paris A 280 (1975) 1295–1298.
[36] A.M. Vershik, Classical and non-classical dynamics with constraints, Global Analysis — Studies and Applications I, Lecture Notes in Mathematics, Vol. 1108, Springer, Berlin, 1984, pp. 278–301.
[37] A.M. Vershik, L.D. Faddeev, Differential geometry and Lagrangian mechanics with constraints, Soviet Phys. Dokl. 17 (1) (1972) 34–36.