GEOMETRIC DESCRIPTION OF VAKONOMIC AND NONHOLONOMIC DYNAMICS. COMPARISON OF SOLUTIONS*

J. CORTÉS[†‡¶], M. DE LEÓN[‡], D. MARTÍN DE DIEGO [‡], AND S. MARTÍNEZ[§‡¶]

Abstract. We treat the vakonomic dynamics with general constraints within a new geometric framework which can be useful in the study of optimal control problems. We compare our formulation with the one of Vershik and Gershkovich in the case of linear constraints. We show how nonholonomic mechanics also admits a new geometrical description which allows to develop an algorithm of comparison between the solutions of both dynamics. Examples illustrating the theory are treated.

Key words. vakonomic dynamics, nonholonomic dynamics, optimal control, symplectic geometry

AMS subject classifications. 34A26, 49K15, 70F25

1. Introduction. As is well known, the application of tools from modern differential geometry in the fields of mechanics and control theory has caused an important progress in these research areas. For example, the study of the geometrical formulation of the nonholonomic equations of motion has led to a better comprehension of locomotion generation, controllability, motion planning and trajectory tracking, raising new interesting questions in these subjects (see [4, 5, 26, 28, 35, 47, 49, 50] and references therein). On the other hand, there are by now many papers in which optimal control problems are addressed using geometric techniques (references [8, 23, 24, 59, 60] are good examples).

In this context, we present a unified geometrical formulation of the dynamics of nonholonomic and vakonomic systems. Both kinds of systems have the same mathematical "ingredients": a Lagrangian function and a set of nonintegrable constraints. But the way in which the equations of motion are derived differs. In the case of vakonomic systems, the dynamics is obtained through the application of a constrained variational principle [1]. In particular, an optimal control problem can be seen as a vakonomic one. The term "vakonomic" ("variational axiomatic kind") is inherited from Kozlov [29], who proposed this mechanics as an alternative set of equations of motion for a physical system in the presence of nonholonomic constraints. Nonholonomic equations of motion are deduced using d'Alembert principle when the constraints are linear or affine.

The two approaches have received a lot of attention in recent years (see [1, 2, 10, 14, 33, 31, 36, 41, 64, 68] and references therein). Vakonomic mechanics (also called dynamical optimization subject to nonholonomic constraints) is used in mathematical economics (growth economic theory), sub-Riemannian geometry, motion of

*This research was partially supported by Spanish DGICYT grants PB97-1257 and PGC2000-2191-E.

[†]Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, 1308 W. Main St., Urbana, IL 61801, USA, jorge@motion.csl.uiuc.edu, Ph. +1-217-333-0656, Fax. +1-217-244-1653

[‡]Laboratory of Dynamical Systems, Mechanics and Control, Instituto de Matemáticas y Física Fundamental, CSIC, Serrano 123, Madrid 28006, SPAIN, mdeleon@imaff.cfmac.csic.es, d.martin@imaff.cfmac.csic.es, Ph. +34 915616800, ext. 1115 and 2408, Fax. +34 915854894

[§]Escola Universitària Politèctnica de Vilanova i la Geltrú, Universidad Politécnica de Cataluña, Av. V. Balaguer s/n, Vilanova i la Geltrú 08800, SPAIN, soniam@mat.upc.es, Ph. +34 938967713, Fax. +34 938967700

[¶]Research partially supported by FPU and FPI grants from the Spanish Ministerio de Educación y Cultura and Ministerio de Ciencia y Tecnología, respectively

microorganisms, etc., while nonholonomic mechanics provides the evolution equations for wheeled and autonomous vehicles, robotic systems, etc.

Several authors have discussed the domains of validity of both approaches [1, 29, 36, 68]. The solutions of the resulting dynamical systems do not coincide, in general, though there are examples in which the nonholonomic solutions can be seen as solutions of the constrained variational problem. In recent papers [20, 36] the characterization of this situation has been studied. In [36] Lewis and Murray considered the example of a ball on a rotating table and showed that the subset of solutions of the nonholonomic problem is not included in the set of vakonomic ones. In [20] Favretti obtains conditions in some particular cases for the equivalence between both formulations.

Our project of unifying the comparative studies of both types of dynamics from a geometrical point of view has brought us to develop a new geometric setting for vakonomic and nonholonomic mechanics, strongly inspired on the Skinner and Rusk formulation for singular Lagrangians systems [58]. Herewith, we are able to compare them using an algorithm which gives rise, under appropriate conditions, to a final constraint submanifold containing all the nonholonomic solutions which are also vakonomic. As an application of the proposed algorithm, we extend several known results [4, 20, 36]. In particular, we prove that any solution of the unconstrained problem which verifies the constraints, is simultaneously a solution of the nonholonomic and the vakonomic problems. This allows us to generalize to arbitrary metrics a result proven in [20] for bundle-like metrics and kinetic energy Lagrangians, $L = \frac{1}{2}g$.

The paper is structured as follows. In §2, we obtain the equations of motion for vakonomic mechanics, assuming an admissibility condition, which permits us to present them in terms of the restriction of the Lagrangian to the constraint submanifold M. Let us recall that from a geometrical point of view, the Lagrangian L is defined on the tangent bundle TQ of the configuration manifold Q, and M represents the submanifold of TQ determined by the vanishing of the nonholonomic constraint functions. We will deal here with arbitrary submanifolds, that is, the constraints may be nonlinear. It should also be pointed out that we do not consider abnormal solutions. It is interesting to note that our derivation of the equations of motion shows that the information provided by L outside M is completely irrelevant for the vakonomic problem. This fact is not clearly seen in the classical way of writing the equations for vakonomic systems [1, 29].

Section 3 is devoted to a reformulation of vakonomic mechanics in geometric terms. In this section we will use as ambient space the fibred manifold $W_0 = T^*Q \times_Q M$, which is in fact a subbundle of the Whitney sum $T^*Q \oplus TQ$ (the phase space in the Skinner and Rusk approach). Since T^*Q is equipped with a canonical symplectic form we can induce a presymplectic structure ω on $T^*Q \times_Q M$. Moreover, we can consider the Hamiltonian function $H_{W_0} = \langle \pi_1, \pi_2 \rangle - \pi_2^* \tilde{L}$, where π_1 and π_2 are the canonical projections, $\langle \cdot, \cdot \rangle$ denotes the natural pairing between vectors and covectors on Q, and $\tilde{L}$ is the restriction of L to M. Then, we prove that the equations of motion of vakonomic mechanics are intrinsically represented by the presymplectic Hamiltonian equation $i_X \omega = dH_{W_0}$. Since the 2-form ω is presymplectic, a constraint algorithm must be applied in order to obtain well-defined solutions of the dynamics. If the algorithm stabilizes, we obtain a family of explicit solutions on the final constraint submanifold. In addition, a compatibility condition is found which determines when the first constraint submanifold W_1 is symplectic (and therefore the algorithm stabilizes at the first step). We illustrate in Subsections 3.1 and 3.2 how this framework can be of use in the analysis of optimal control problems.

In §4, we compare our approach with the one of Vershik-Gershkovich [64] for vakonomic systems with linear constraints. We prove that both are related by a convenient presymplectomorphism, so that our approach could be considered as a generalization to the case of nonlinear constraints.

Since we want to compare vakonomic and nonholonomic dynamics, it is necessary to construct a geometrical framework for nonholonomic mechanics using a closed phase space. Indeed, in §5 it is proved that the nonholonomic dynamics lives on a submanifold $\tilde{M}$ of W_0. In general, we have again a presymplectic system and a constraint algorithm is needed to obtain the dynamics on the final constraint submanifold.

In §6, assuming that the vakonomic and the nonholonomic dynamics live on W_1 and $\tilde{M}$, respectively, we can compare their solutions by means of the map $\Upsilon : W_1 \to \tilde{M}$, $(\alpha, v) \mapsto (Leg_L(v), v)$. We present here an algorithm that selects those solutions of the nonholonomic problem that can be seen as solutions of the constrained variational one. Several illustrative examples are worked out in order to illustrate the different behaviours, showing that our framework provides a generalization and common context for the equivalence results in [4, 20, 36]. In particular, in the example of the planar mobile robot, we prove that, under an appropriate design of the system, every solution of the nonholonomic problem can be seen as a solution of the vakonomic one.

2. Variational approach to constrained mechanics. Let Q be the configuration manifold with dimension n and $L : TQ \longrightarrow \mathbb{R}$ an autonomous Lagrangian function. If (q^A), $1 \leq A \leq n$, are coordinates on Q, we denote by $(q^A, \dot{q}^A)$ the natural bundle coordinates on TQ in terms of which the tangent bundle projection $\tau_Q : TQ \longrightarrow Q$ reads as $\tau_Q(q^A, \dot{q}^A) = (q^A)$.

Let us suppose that the system is subject to some constraints given by a $(2n-m)$-dimensional submanifold M of TQ, locally defined by $\Phi^\alpha = 0$, $1 \leq \alpha \leq m$, where $\Phi^\alpha : TQ \longrightarrow \mathbb{R}$. Throughout the paper, we will assume the following admissibility condition for the submanifold $M \subseteq TQ$: for all $x \in M$, $\dim T_xM^o = \dim S^*T_xM^o$, where $S = dq^A \otimes \frac{\partial}{\partial \dot{q}^A}$ is the canonical vertical endomorphism (see [34]). This is equivalent to saying that the rank of the matrix

$$\frac{\partial(\Phi^1, \dots, \Phi^m)}{\partial(\dot{q}^1, \dots, \dot{q}^n)}$$

is m for any choice of coordinates $(q^A, \dot{q}^A)$ in TQ. Consequently, by the implicit function theorem, we can locally express the constraints (reordering coordinates if necessary) as

$$\dot{q}^\alpha = \Psi^\alpha(q^A, \dot{q}^a)\,, \tag{2.1}$$

where $1 \leq \alpha \leq m$, $m+1 \leq a \leq n$ and $1 \leq A \leq n$. Then, $(q^A, \dot{q}^a)$ are local coordinates for the submanifold M of TQ.

We denote the set of twice differentiable curves connecting two points $x, y \in Q$ as

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

This set is a differentiable infinite-dimensional manifold [3].

Let c be a curve in $\mathcal{C}^2(x,y)$. A variation of c is a curve c_s in $\mathcal{C}^2(x,y)$, that is a differentiable mapping $c_s : (-\epsilon, \epsilon) \to \mathcal{C}^2(x,y)$, $s \mapsto c_s(t)$, such that $c_0 = c$. An infinitesimal variation of c is the tangent vector of a variation of c, that is,

$$u(t) = \left.\frac{dc_s(t)}{ds}\right|_{s=0} \in T_{c(t)}Q\,.$$

The tangent space of $\mathcal{C}^2(x,y)$ at c is then given by

$$T_c\mathcal{C}^2(x,y) = \{u : [0,1] \longrightarrow TQ \mid u \text{ is } C^1,\ u(t) \in T_{c(t)}Q,\ u(0) = 0 \text{ and } u(1) = 0\}\,.$$

Now, we introduce a special subset $\tilde{\mathcal{C}}^2(x,y)$ of $\mathcal{C}^2(x,y)$ which consists of those curves whose velocities belong to the constraint submanifold M

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

Finally, let us consider the action functional $\mathcal{J}$ defined by

$$\mathcal{J} : \mathcal{C}^2(x,y) \longrightarrow \mathbb{R}\,, \quad c \mapsto \mathcal{J}(c) = \int_0^1 L(\dot{c}(t))\,dt\,.$$

DEFINITION 2.1. *The* **vakonomic problem** *associated with* (Q, L, M, x, y) *consists of extremizing the functional* $\mathcal{J}$ *among the curves satisfying the constraints imposed by* M, $c \in \tilde{\mathcal{C}}^2(x,y)$. *Hence, a curve* $c \in \tilde{\mathcal{C}}^2(x,y)$ *will be a* **solution of the vakonomic problem** *if* c *is a critical point of* $\mathcal{J}_{|\tilde{\mathcal{C}}^2(x,y)}$.

REMARK 2.2. In this paper, we will assume that the solution curves $c \in \tilde{\mathcal{C}}(x,y)$ admit enough nontrivial variations in $\tilde{\mathcal{C}}(x,y)$. These solutions are called normal in the literature, in contrast to the abnormal ones, which are pathological curves which do not admit sufficient nontrivial variations [1]. Several investigators have shown the existence of C^1, stable under perturbations abnormal $\mathcal{J}$-minimizing solutions [37, 45].

Now, we find a characterization for the solutions of the vakonomic problem.

PROPOSITION 2.3. *A curve* $c \in \tilde{\mathcal{C}}^2(x,y)$ *is a normal solution of the vakonomic problem if and only if there exists* $\mu : [0,1] \to \mathbb{R}^m$ *such that*

$$\left\{\begin{array}{l} \dfrac{d}{dt}\left(\dfrac{\partial \tilde{L}}{\partial \dot{q}^a}\right) - \dfrac{\partial \tilde{L}}{\partial q^a} = \mu_\alpha \left[\dfrac{d}{dt}\left(\dfrac{\partial \Psi^\alpha}{\partial \dot{q}^a}\right) - \dfrac{\partial \Psi^\alpha}{\partial q^a}\right] + \dot{\mu}_\alpha \dfrac{\partial \Psi^\alpha}{\partial \dot{q}^a}, \\ \dot{\mu}_\alpha = \dfrac{\partial \tilde{L}}{\partial q^\alpha} - \mu_\beta \dfrac{\partial \Psi^\beta}{\partial q^\alpha}, \\ \dot{q}^\alpha = \Psi^\alpha(q^A, \dot{q}^a), \end{array}\right. \tag{2.2}$$

where $\tilde{L} : M \to \mathbb{R}$ *is the restriction of* L *to* M.

Proof. The condition for a curve to be a solution of the vakonomic problem is

$$0 = d\mathcal{J}(c) \cdot u = \frac{d}{ds}\mathcal{J}(c_s)\Big|_{s=0}\,,$$

for any variation c_s in $\tilde{\mathcal{C}}^2(x,y)$ of c, where $u = \dfrac{dc_s}{ds}\Big|_{s=0}$. Then, we have that

$$0 = \frac{d}{ds}\mathcal{J}(c_s)\Big|_{s=0} = \frac{d}{ds}\left(\int_0^1 L(\dot{c}_s(t))\,dt\right)\Big|_{s=0} = \int_0^1 \frac{d}{ds}L(\dot{c}_s(t))\Big|_{s=0} dt\,.$$

In local coordinates, we obtain

$$\begin{aligned} 0 &= \int_0^1 \left(\frac{\partial L}{\partial q^A}u^A + \frac{\partial L}{\partial \dot{q}^a}\dot{u}^a + \frac{\partial L}{\partial \dot{q}^\alpha}\frac{\partial \Psi^\alpha}{\partial q^A}u^A + \frac{\partial L}{\partial \dot{q}^\alpha}\frac{\partial \Psi^\alpha}{\partial \dot{q}^a}\dot{u}^a\right) dt \\ &= \int_0^1 \left(\left[\frac{\partial L}{\partial q^A} + \frac{\partial L}{\partial \dot{q}^\alpha}\frac{\partial \Psi^\alpha}{\partial q^A}\right]u^A + \left[\frac{\partial L}{\partial \dot{q}^a} + \frac{\partial L}{\partial \dot{q}^\alpha}\frac{\partial \Psi^\alpha}{\partial \dot{q}^a}\right]\dot{u}^a\right) dt \\ &= \int_0^1 \left(\frac{\partial \tilde{L}}{\partial q^A}u^A + \frac{\partial \tilde{L}}{\partial \dot{q}^a}\dot{u}^a\right) dt\,. \end{aligned} \tag{2.3}$$

From (2.1) we know that the infinitesimal variations u^A, $1 \leq A \leq n$, are not arbitrary. Consider the functions μ_α defined as the solutions of the following system of first order differential equations

$$\dot{\mu}_\alpha = \frac{\partial \tilde{L}}{\partial q^\alpha}\Big|_c - \mu_\beta \frac{\partial \Psi^\beta}{\partial q^\alpha}\Big|_c ,\ 1 \leq \alpha \leq m\,.$$

Then, using the fact that $\dot{u}^\alpha = \dfrac{\partial \Psi^\alpha}{\partial q^A} u^A + \dfrac{\partial \Psi^\alpha}{\partial \dot{q}^a}\dot{u}^a$, we get

$$\frac{d}{dt}(\mu_\alpha u^\alpha) = \mu_\alpha \dot{u}^\alpha + \left(\frac{\partial \tilde{L}}{\partial q^\alpha} - \mu_\beta \frac{\partial \Psi^\beta}{\partial q^\alpha}\right) u^\alpha = u^\alpha \frac{\partial \tilde{L}}{\partial q^\alpha} + \mu_\alpha \frac{\partial \Psi^\alpha}{\partial q^a} u^a + \mu_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a}\dot{u}^a\ ,$$

or, equivalently, $u^\alpha \dfrac{\partial \tilde{L}}{\partial q^\alpha} = \dfrac{d}{dt}(\mu_\alpha u^\alpha) - \mu_\alpha \dfrac{\partial \Psi^\alpha}{\partial q^a} u^a - \mu_\alpha \dfrac{\partial \Psi^\alpha}{\partial \dot{q}^a}\dot{u}^a$. Substituting the last expression in (2.3) and integrating by parts, we obtain

$$d\mathcal{J}(c)\cdot u = \int_0^1 \left(\left[\frac{\partial \tilde{L}}{\partial q^a} - \mu_\alpha \frac{\partial \Psi^\alpha}{\partial q^a}\right] u^a + \left[\frac{\partial \tilde{L}}{\partial \dot{q}^a} - \mu_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a}\right]\dot{u}^a \right) dt\,.$$

Now, since

$$\left[\frac{\partial \tilde{L}}{\partial \dot{q}^a} - \mu_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a}\right]\dot{u}^a = \frac{d}{dt}\left(\left[\frac{\partial \tilde{L}}{\partial \dot{q}^a} - \mu_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a}\right] u^a\right) - \frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial \dot{q}^a} - \mu_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a}\right) u^a\ ,$$

using again integration by parts, we can write

$$0 = \int_0^1 \left[\frac{\partial \tilde{L}}{\partial q^a} - \mu_\alpha \frac{\partial \Psi^\alpha}{\partial q^a} - \frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial \dot{q}^a} - \mu_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a}\right)\right] u^a\, dt\ .$$

As the infinitesimal variations u^a are arbitrary, the fundamental lemma of the Calculus of Variations applies and we can assert that $d\mathcal{J}(c)\cdot u = 0$ if and only if c and μ_α satisfy equations (2.2). □

REMARK 2.4. The usual way in which the equations of motion for vakonomic mechanics are presented is the following

$$\begin{cases} \dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{q}^A}\right) - \dfrac{\partial L}{\partial q^A} = \dot{\lambda}_\alpha \dfrac{\partial \Phi^\alpha}{\partial \dot{q}^A} + \lambda_\alpha \left[\dfrac{d}{dt}\left(\dfrac{\partial \Phi^\alpha}{\partial \dot{q}^A}\right) - \dfrac{\partial \Phi^\alpha}{\partial q^A}\right], \\ \Phi^\alpha(q,\dot{q}) = 0,\ 1 \leq \alpha \leq m\,, \end{cases} \tag{2.4}$$

where $\Phi^\alpha = \Psi^\alpha - \dot{q}^\alpha$ and $\lambda_\alpha = \dfrac{\partial L}{\partial \dot{q}^\alpha} - \mu_\alpha$, $1 \leq \alpha \leq m$. Observe that, in contrast to equations (2.2), equations (2.4) are expressed in terms of the ambient Lagrangian $L : TQ \to \mathbb{R}$. Equations (2.2) stress how the information given by L outside M is irrelevant to obtain the vakonomic equations, a fact that it is not promptly deduced from equations (2.4). This is in contrast with what happens in nonholonomic mechanics (see § 5 below).

Equations (2.4) can be seen as the Euler-Lagrange equations for the extended Lagrangian $\mathcal{L} = L + \lambda_\alpha \Phi^\alpha$. We will not follow this approach here, which has been exploited successfully in [20, 28, 42, 43]. Finally, note that if we consider the extended Lagrangian $\lambda_0 L + \lambda_\alpha \Phi^\alpha$, with $\lambda_0 = 1$ or 0, then we recover all the solutions, both the normal and the abnormal ones [1].

3. Geometric approach to vakonomic mechanics. We will develop a geometric characterization of vakonomic mechanics following an approach similar to the formulation given by Skinner and Rusk [58] for singular Lagrangians (see also [15, 22, 38]). This characterization is specially interesting, for it enables us to study both linear and nonlinear constraints in an intrinsic way. Moreover, as we shall discuss later, this formalism will allow to use ideas from Geometric Mechanics in the treatment of optimal control problems.

Consider the Whitney sum of T^*Q and TQ, $T^*Q \oplus TQ$, and its canonical projections $pr_1 : T^*Q \oplus TQ \longrightarrow T^*Q$, $pr_2 : T^*Q \oplus TQ \longrightarrow TQ$. Let us take the submanifold $W_0 = pr_2^{-1}(M)$, where M is the constraint submanifold, locally determined by the constraint equations $\Phi^\alpha = 0$, $1 \leq \alpha \leq m$. We will denote $W_0 = T^*Q \times_Q M$ and $\pi_1 = pr_{1|W_0}$, $\pi_2 = pr_{2|W_0}$. Now, define on $T^*Q \times_Q M$ the presymplectic 2-form $\omega = \pi_1^* \omega_Q$, where ω_Q is the canonical symplectic form on T^*Q. Observe that the rank of this presymplectic form is equal to $2n$ everywhere. Define also the function

$$H_{W_0} = \langle \pi_1, \pi_2 \rangle - \pi_2^* \tilde{L} \,,$$

where $\langle \cdot, \cdot \rangle$ denotes the natural pairing between vectors and covectors on Q.

If (q^A) are local coordinates on a neighborhood U of Q, $(q^A, \dot{q}^a)$ coordinates on $TU \cap M$ and (q^A, p_A) the induced coordinates on T^*U, then we have induced coordinates $(q^A, p_A, \dot{q}^a)$ on $T^*U \times_Q (TU \cap M)$. Locally, the Hamiltonian function H_{W_0} reads as

$$H_{W_0}(q^A, p_A, \dot{q}^a) = p_a \dot{q}^a + p_\alpha \Psi^\alpha - \tilde{L}(q^A, \dot{q}^a) \,,$$

and the 2-form ω is $\omega = dq^A \wedge dp_A$.

Now, we will see how the dynamics of the vakonomic system (2.2) is determined by the solutions of the equation

$$i_X \omega = dH_{W_0} \,. \tag{3.1}$$

This then justifies the use of the following terminology:

Definition 3.1. *The presymplectic Hamiltonian system* $(T^*Q \times_Q M, \omega, H_{W_0})$ *will be called* **vakonomic Hamiltonian system.**

The system $(T^*Q \times_Q M, \omega, H_{W_0})$ being presymplectic, we may apply the Gotay-Nester constraint algorithm [21]. First we consider the set of points W_1 of $T^*Q \times_Q M$ where (3.1) has a solution. This first constraint submanifold is determined by

$$W_1 = \{x \in T^*Q \times_Q M \mid dH_{W_0}(x)(V) = 0, \ \forall V \in \ker \omega(x)\} \,.$$

Locally, $\ker \omega = \operatorname{span}\langle \partial / \partial \dot{q}^a \rangle$. Therefore, the constraint submanifold W_1 is locally characterized by the vanishing of the constraints

$$\varphi_a = p_a + p_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a} - \frac{\partial \tilde{L}}{\partial \dot{q}^a} = 0 \,, \ m+1 \leq a \leq n \,,$$

or, equivalently,

$$p_a = \frac{\partial \tilde{L}}{\partial \dot{q}^a} - p_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a} \,, \ m+1 \leq a \leq n \,. \tag{3.2}$$

Expanding the expressions in (3.1) the equations of motion along W_1 are

$$\dot{q}^A = \frac{\partial H_{W_0}}{\partial p_A} \,, \quad \dot{p}_A = -\frac{\partial H_{W_0}}{\partial q^A} \,,$$

which is equivalent to

$$\dot{q}^\alpha = \Psi^\alpha(q^A, \dot{q}^a) \,, \tag{3.3}$$

$$\dot{p}_\alpha = \frac{\partial \tilde{L}}{\partial q^\alpha} - p_\beta \frac{\partial \Psi^\beta}{\partial q^\alpha} \,, \tag{3.4}$$

$$\frac{d}{dt}\left(\frac{\partial \tilde{L}}{\partial \dot{q}^a} - p_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a}\right) = \frac{\partial \tilde{L}}{\partial q^a} - p_\beta \frac{\partial \Psi^\beta}{\partial q^a} \,. \tag{3.5}$$

Observe that these equations are precisely the vakonomic equations of motion (2.2), where now $p_\alpha = \mu_\alpha$, $1 \leq \alpha \leq m$.

REMARK 3.2. The momenta p_α, $1 \leq \alpha \leq m$, play the role of the Lagrange multipliers, but they do not have any physical meaning (see [61]).

Therefore, a vector field X solution of equation (3.1) will generally be of the form

$$\begin{aligned} X = \dot{q}^a \left(\frac{\partial}{\partial q^a} + \left(\frac{\partial^2 \tilde{L}}{\partial q^a \partial \dot{q}^b} - p_\gamma \frac{\partial^2 \Psi^\gamma}{\partial q^a \partial \dot{q}^b} \right) \frac{\partial}{\partial p_b} \right) + \\ + \Psi^\alpha \left(\frac{\partial}{\partial q^\alpha} + \left(\frac{\partial^2 \tilde{L}}{\partial q^\alpha \partial \dot{q}^b} - p_\gamma \frac{\partial^2 \Psi^\gamma}{\partial q^\alpha \partial \dot{q}^b} \right) \frac{\partial}{\partial p_b} \right) + \\ + \bar{X}^a \left(\frac{\partial}{\partial \dot{q}^a} + \left(\frac{\partial^2 \tilde{L}}{\partial \dot{q}^a \partial \dot{q}^b} - p_\gamma \frac{\partial^2 \Psi^\gamma}{\partial \dot{q}^a \partial \dot{q}^b} \right) \frac{\partial}{\partial p_b} \right) + \left(\frac{\partial \tilde{L}}{\partial q^\alpha} - p_\beta \frac{\partial \Psi^\beta}{\partial q^\alpha} \right) \left(\frac{\partial}{\partial p_\alpha} - \frac{\partial \Psi^\alpha}{\partial \dot{q}^b} \frac{\partial}{\partial p_b} \right) , \end{aligned}$$

where the coefficients $\bar{X}^a$ are still undetermined. The solution on W_1 may not be tangent to W_1. In such a case, we have to restrict W_1 to the submanifold W_2 where this solution is tangent to W_1. Proceeding further, we obtain a sequence of submanifolds (we are assuming that all the subsets generated by the algorithm are submanifolds)

$$\cdots \hookrightarrow W_k \hookrightarrow \cdots \hookrightarrow W_2 \hookrightarrow W_1 \hookrightarrow W_0 = T^*Q \times_Q M \,.$$

Algebraically, these constraint submanifolds may be described as

$$W_i = \{ x \in T^*Q \times_Q M \mid dH_{W_0}(x)(v) = 0 \,,\ \forall v \in T_x W_{i-1}^\perp \} \,, \quad i \geq 1 \,, \tag{3.6}$$

where $T_x W_{i-1}^\perp = \{ v \in T_x(T^*Q \times_Q M) \mid \omega(x)(u,v) = 0 \ ,\ \forall u \in T_x W_{i-1} \}$. If this constraint algorithm stabilizes, i.e., if there exists a positive integer $k \in \mathbb{N}$ such that $W_{k+1} = W_k \neq W_{k-1}$ and $\dim W_k \neq 0$, then we will have obtained a final constraint submanifold $W_f = W_k$ on which a vector field X exists such that

$$(i_X \omega = dH_{W_0})_{|W_f} \,.$$

Note that on W_f we will have an explicit solution of the vakonomic dynamics. A very important particular case is when the final constraint submanifold is the first one, i.e. $W_f = W_1$. Observe that the dimension of W_1 is even, $\dim W_1 = 2n$. In the sequel, we will investigate when this constraint submanifold is equipped with a symplectic 2-form in order to determine a unique solution X of the vakonomic equations. Obviously, this geometrical study is related to the explicit or implicit character of the second order differential equations obtained in (2.2).

Denote by ω_{W_1} the restriction of the presymplectic 2-form ω to W_1.

PROPOSITION 3.3. *(W_1, ω_{W_1}) is a symplectic manifold iff for any point of W_1,*

$$\det\left(\frac{\partial^2 \tilde{L}}{\partial \dot{q}^a \partial \dot{q}^b} - p_\alpha \frac{\partial^2 \Psi^\alpha}{\partial \dot{q}^a \partial \dot{q}^b}\right) \neq 0\,. \tag{3.7}$$

Proof. ω_{W_1} is symplectic if and only if $T_x W_1 \cap (T_x W_1)^\perp = 0$, for all $x \in W_1$. This condition is satisfied if and only if the matrix $d\varphi_a(\frac{\partial}{\partial \dot{q}^b})$ is regular, that is,

$$\det\left(\frac{\partial^2 \tilde{L}}{\partial \dot{q}^a \partial \dot{q}^b} - p_\alpha \frac{\partial^2 \Psi^\alpha}{\partial \dot{q}^a \partial \dot{q}^b}\right) \neq 0\,,$$

for all $x \in W_1$ □

In this case, equations (3.5) can be rewritten in explicit form as

$$\ddot{q}^a = -\bar{\mathcal{C}}^{ab}\left[\dot{q}^A \frac{\partial^2 \tilde{L}}{\partial q^A \partial \dot{q}^b} - \dot{q}^A p_\alpha \frac{\partial^2 \Psi^\alpha}{\partial q^A \partial \dot{q}^b} - \frac{\partial \tilde{L}}{\partial q^b} + p_\alpha \frac{\partial \Psi^\alpha}{\partial q^b} - \left(\frac{\partial \tilde{L}}{\partial q^\gamma} - p_\beta \frac{\partial \Psi^\beta}{\partial q^\gamma}\right)\frac{\partial \Psi^\gamma}{\partial \dot{q}^b}\right], \tag{3.8}$$

where

$$\bar{\mathcal{C}}_{ab} = \frac{\partial^2 \tilde{L}}{\partial \dot{q}^a \partial \dot{q}^b} - p_\alpha \frac{\partial^2 \Psi^\alpha}{\partial \dot{q}^a \partial \dot{q}^b}\,, \tag{3.9}$$

and $(\bar{\mathcal{C}}^{ab})$ denotes the inverse matrix of $(\bar{\mathcal{C}}_{ab})$.

REMARK 3.4. The characterization found in Proposition 3.3 for the symplectic nature of the manifold (W_1, ω_{W_1}) implies, by the implicit function theorem, that the constraint equations

$$\varphi_a = p_a + p_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a} - \frac{\partial \tilde{L}}{\partial \dot{q}^a} = 0\,, m+1 \leq a \leq n\,,$$

locally determine the variables $\dot{q}^a$, $m+1 \leq a \leq n$. That is, we have $\dot{q}^a = \varsigma^a(q^A, p_A)$, $m+1 \leq a \leq n$. Therefore, we can also consider local coordinates (q^A, p_A) on W_1. In such a case, the symplectic form and the restriction of the Hamiltonian H_{W_0} to W_1 have the following local expressions

$$\omega_{W_1} = dq^A \wedge dp_A\,, \quad H_{W_1} = p_a \varsigma^a + p_\alpha \Psi^\alpha - \bar{L}(q^A, p_A)\,,$$

where $\bar{L}(q^A, p_A) = \tilde{L}(q^A, \varsigma^a(q^A, p_A))$. Consequently, equations (3.3)-(3.5) can be rewritten in Hamiltonian form as

$$\dot{q}^A = \frac{\partial H_{W_1}}{\partial p_A}\,, \quad \dot{p}_A = -\frac{\partial H_{W_1}}{\partial q^A}\,.$$

This choice of coordinates is common in optimal control theory.

Now, observe that, if the constraints are linear in the velocities, we can write $\dot{q}^\alpha = \Psi^\alpha_a(q)\dot{q}^a$. Then, from Proposition 3.3, ω_{W_1} is symplectic if and only if

$$\det\left(\frac{\partial^2 \tilde{L}}{\partial \dot{q}^a \partial \dot{q}^b}\right) \neq 0\,.$$

PROPOSITION 3.5. *Suppose that the constraints are given by* $\dot{q}^\alpha = \Psi^\alpha_a(q)\dot{q}^a$, $1 \leq \alpha \leq m$ *and the Lagrangian* L *is regular. Denote by* (W^{AB}) *the inverse matrix of the Hessian matrix of* L. *In this case,* ω_{W_1} *is symplectic on* W_1 *if and only if the constraints are compatible, that is, the matrix whose entries are*

$$\mathcal{C}^{\alpha\beta} = W^{ab}\Psi^\alpha_a\Psi^\beta_b - W^{\alpha b}\Psi^\beta_b - W^{a\beta}\Psi^\alpha_a + W^{\alpha\beta}\,,$$

is nonsingular.

Proof. See the geometrical proof of Theorem IV. 3 in reference [33]. □

REMARK 3.6. The compatibility condition guarantees the existence and uniqueness of the solutions for the nonholonomic problem with Lagrangian L and constraint submanifold M [33, 57].

Before ending this section, we would like to make some remarks concerning this geometric approach to vakonomic dynamics. First of all, we must say that it provides an intrinsic formulation of variational problems subject to both linear and *nonlinear* constraints on *manifolds*. In addition, this formulation belongs to the context of Symplectic Geometry and Geometric Mechanics, following previous work by Bloch and Crouch [4, 8, 9], Jurdjevic [23, 24] and others. There is a whole collection of ideas and methods ensuing from these fields that have been used in the treatment of optimal control problems. Apart from being of use as a tool for an algorithmic study of the existence of optimal solutions and their domains of definition, we think that this formulation has something to contribute in at least three directions: the study of the symmetry properties of constrained problems [8, 18, 24, 43] (infinitesimal, Noether and Cartan symmetries, dynamical symmetries,...), the study of higher order variational problems [6] (since a generalization of our approach to the higher order case seems to be straightforward) and the development of numerical integrators [19, 55, 65, 66, 67] that take into account the geometry of the problem (2-form, Hamiltonian, momentum) and are competitive with the traditional methods.

An immediate outcome of the formulation on $T^*Q \times_Q M$ is that for the study of problems subject to nonlinear constraints we can use similar techniques to those used for the linear case. Finally, this framework will allow us in § 6 to compare vakonomic dynamics with nonholonomic dynamics within a common setting.

In the following, we aim to illustrate some of the above ideas on two examples.

3.1. Applications in economy. The variational calculus is an indispensable tool in many economic problems [25, 39, 52]. In fact, a typical optimization problem in modern economics deals with extremizing the functional

$$\int_0^T D(t)U[f(t,k,\dot{k})]\,dt$$

subject or not to constraints. Here, $D(t)$ is a discount rate factor, U an utility function, f a consumption function and k the capital-labor ratio. It is common to find dynamical economic models with nonholonomic constraints.

EXAMPLE 3.7 (Closed von Neumann System [53, 54, 56]). Consider the transformation function F on $\mathbb{R}^{2n}$ which relates n capital goods $K_1, K_2, \ldots, K_n$ and the net capital formations $\dot{K}_1, \dot{K}_2, \ldots, \dot{K}_n$ as

$$F(K_1,\ldots,K_n,\dot{K}_1,\ldots,\dot{K}_n) = K_1^{\alpha_1}K_2^{\alpha_2}\cdots K_n^{\alpha_n} - \left[\dot{K}_1^2 + \ldots + \dot{K}_n^2\right]^{1/2}\,,$$

with $\alpha_1+\alpha_2+\ldots+\alpha_n=1$. The von Neumann problem consists of maximizing

$$\int_0^T \dot K_n\,dt \quad \text{subject to} \quad F(K_1,\ldots,K_n,\dot K_1,\ldots,\dot K_n)=0\,,$$

with appropriate initial conditions.

Our formalism makes it possible to write this problem as a presymplectic system on $W_0=\mathbb{R}^{3n-1}$. The constraint $F=0$ can be rewritten as

$$\dot K_1=\pm\left(K_1^{2\alpha_1}\cdots K_n^{2\alpha_n}-\sum_{i=2}^n \dot K_i^2\right)^{1/2}=\pm\Psi(K_1,\ldots,K_n,\dot K_2,\ldots,\dot K_n)\,.$$

Here, we restrict the analysis to the component $\dot K_1=\Psi$. Taking coordinates $(K_1,\ldots,K_n,\dot K_2,\ldots,\dot K_n,P^1,\ldots,P^n)$ we have that

$$\omega=\sum_{j=1}^n dK_j\wedge dP^j\,,\ H_{W_0}=\sum_{i=2}^n P^i\dot K_i+P^1\cdot\left(K_1^{2\alpha_1}K_2^{2\alpha_2}\cdots K_n^{2\alpha_n}-\sum_{i=2}^n\dot K_i^2\right)^{1/2}-\dot K_n\,.$$

Applying the Gotay and Nester algorithm, new constraints arise,

$$\begin{aligned}P^i&=P^1\dot K_i\left(K_1^{2\alpha_1}K_2^{2\alpha_2}\cdots K_n^{2\alpha_n}-\sum_{i=2}^n\dot K_i^2\right)^{-1/2}, \quad 2\le i\le n-1\\ P^n&=1+P^1\dot K_n\left(K_1^{2\alpha_1}K_2^{2\alpha_2}\cdots K_n^{2\alpha_n}-\sum_{i=2}^n\dot K_i^2\right)^{-1/2}.\end{aligned}$$

Therefore, from (3.3-3.5) the initial system is determined by solving the following n differential equations on the variables $(K_1,\ldots,K_n,\dot K_2,\ldots,\dot K_n,P^1)$

$$\begin{aligned}\dot P^1&=-P^1\alpha_1\left(K_1^{2\alpha_1-1}K_2^{2\alpha_2}\cdots K_n^{2\alpha_n}\right)G\\ 0&=\dot P^1\dot K_iG\\ &\quad+P^1\left[\left(\ddot K_i+\alpha_i\left(K_1^{2\alpha_1}\cdots K_i^{2\alpha_i-1}\cdots K_n^{2\alpha_n}\right)\right)G+\dot K_i\frac{d}{dt}(G)\right]\,,\ 2\le i\le n\,,\end{aligned}\tag{3.10}$$

where $G=1/\Psi$. The presymplectic context for these optimal equations provides us with some new insights into the problem. On the one hand, the existence of well-defined solutions to equations (3.10) is not guaranteed in general. It can occur, for instance, that an optimal curve starting from a point in W_1 "escapes" from this phase space after some time because the dynamical vector field is no longer tangent to W_1. But one can indeed eliminate this possibility. Consider the case $n=2$ for simplicity. Assume $\Psi\neq 0$. Otherwise the dynamics is fully determined and the optimization problem is trivial (we have abnormal solutions). The determinant (3.7) is equal to

$$\frac{P^1}{\Psi^3}(\Psi^2+\dot K_2^2)=\frac{H_{W_1}}{\Psi^2}\,.\tag{3.11}$$

Therefore, if the optimal curve starts from any point in $x\in W_1$ such that $H_{W_1}(x)\neq 0$, equation (3.11) guarantees that the dynamics of the vakonomic problem remains tangent to W_1. On the other hand, the optimal solutions with $H_{W_1}=0$ are stationary curves, $K_1=\text{const}$, $K_2=\text{const}$ and $K_1K_2=0$.

This formulation can also shed light on the aspect of symmetries and conservation laws. It is known [53, 54, 56] that the closed von Neumann system possesses, besides the Hamiltonian H, another conservation law, which is usually found by *ad hoc* methods. However, it is not difficult to define in our context the notion of *Noether symmetry* and verify that the vector field

$$Y = \sum_{j=1}^{n} K_i \frac{\partial}{\partial K^i} \in \mathfrak{X}(Q)$$

indeed corresponds to such a symmetry. The associated conservation law is precisely given by $\Phi = P^1(K_1\Psi + \sum_{j=2}^{n} K_j \dot{K}_j)/\Psi$. In the same way, one can explore the presence of other types of symmetries, like Cartan symmetries for example [32, 34, 48].

Finally, obtaining explicit solutions of equations (3.10) is in general a very difficult task. The use of numerical integrators can help in analyzing the behavior of the system. In the last years there has been an increasing activity in the development of integrators that take into account the geometric structures associated with the problem [19, 55, 65, 66, 67]. The proposed formalism offers the possibility of designing such methods for a variety of optimal control problems.

3.2. LC-circuits. The dynamics of nonlinear LC electric circuits [44] can be given a variational interpretation, as discussed in [46]. Here, we treat this class of systems under our vakonomic formalism and study the well-posedness of the optimal equations.

Consider a circuit consisting of capacitors and inductors, which are charge and current controlled. Let $\mathcal{C}$ be the collection of n-capacitor branches and $\mathcal{L}$ the m-inductor branches. Denote by $q \in Q_{\mathcal{C}}$ the vector of capacitor charges, and by $i \in Q_{\mathcal{L}}$ the inductor currents. Kirchhoff's current and voltage laws require that $\dot{q} = A_{\mathcal{C}}u$, $i = A_{\mathcal{L}}u$, where $A_{\mathcal{C}}$ and $A_{\mathcal{L}}$ are appropriate linear maps from a vector space $\mathcal{U}$ to $Q_{\mathcal{C}}$ and $Q_{\mathcal{L}}$ characterizing, respectively, the topology of the network and the chosen current reference directions. The new variables $u \in \mathcal{U}$ are usually thought of as a vector of some independent loop currents. The generality of the interconnection structure of the circuit relies on how general the matrices $A_{\mathcal{C}}$, $A_{\mathcal{L}}$ can be. In the following, we will assume that $A_{\mathcal{L}}$ is non-singular and then the space $\mathcal{U}$ will be identified with $Q_{\mathcal{L}}$ through $A_{\mathcal{L}}$. Finally, denote by $W_e : Q_{\mathcal{C}} \to \mathbb{R}$ the electric energy and by $W_m^* : Q_{\mathcal{L}} \to \mathbb{R}$ the magnetic coenergy of the circuit.

The dynamics of the circuit is governed by the element equations, the equations arising from Kirchhoff's current law and those arising from Kirchhoff's voltage law. After some manipulations, these equations may be reduced to

$$\dot{q} = A_{\mathcal{C}}u\,, \quad A_{\mathcal{L}}^* \frac{d}{dt}\left(dW_m^*(A_{\mathcal{L}}u)\right) = -A_{\mathcal{C}}^* dW_e(q)\,, \tag{3.12}$$

where the star superscript denotes the transpose of the corresponding matrix operator. However, the well-posedness of this mathematical model for the electric circuit is not guaranteed in general. It could be, for instance, that some specifications of initial conditions $(q(0), u(0))$ turn out to be incompatible with the algebraic constraints embedded in equations (3.12).

The theoretical setting described above can bring some new insight into this question. Consider as configuration space the product manifold $Q = Q_{\mathcal{C}} \times Q_{\mathcal{L}}$ with coordinates (q^α, u^a). Let $L : TQ \to \mathbb{R}$, $L = W_m^*(A_{\mathcal{L}}u) - W_e(q)$, be the Lagrangian and define $M \subseteq TQ$ by $\dot{q}^\alpha = (A_{\mathcal{C}})^\alpha_b u^b$, as the submanifold of constraints. Then, the dynamics of the LC circuit is found to be defined on the tertiary constraint submanifold of the presymplectic Hamiltonian system $(T^*Q \times_Q M, \omega, H)$. This means that all initial conditions in W_3 are compatible in the sense of the previous paragraph.

Let $(q^\alpha, u^a, \xi_\alpha, \zeta_a, \dot{u}^a)$ be the local coordinates in $W_0 = T^*Q \times_Q M$. Then,

$$\omega = dq^\alpha \wedge d\xi_\alpha + du^a \wedge d\zeta_a \,, \quad H = \zeta_a \dot{u}^a + \xi_\alpha A^\alpha_{\mathcal{C}b} u^b - W^*_m(A_{\mathcal{L}} u) + W_e(q) \,.$$

The first submanifold of constraints is given by $W_1 = \{x \in W_0 \mid dH_x(\frac{\partial}{\partial \dot{u}^a}) = \zeta_a = 0\}$. After some computations, we find that

$$TW_1 \cap TW_1^\perp = \operatorname{span}\left\{\frac{\partial}{\partial u^a}, \frac{\partial}{\partial \dot{u}^a}\right\} ,$$

and hence we must continue with the constraint algorithm. Following (3.6), we have that W_2 is described by the new constraints

$$\frac{\partial H}{\partial u^a} = [A^*_{\mathcal{C}}\xi - A^*_{\mathcal{L}} dW^*_m(A_{\mathcal{L}} u)]_a = 0 \,. \tag{3.13}$$

Under the additional assumption of invertibility of dW^*_m, or, equivalently, under the assumption that the LC circuit is also flux controlled, we can ensure that there exists a magnetic energy W_m such that $dW^*_m(A_{\mathcal{L}} u) = \phi \iff A_{\mathcal{L}} u = dW_m(\phi)$. Then, we can rewrite (3.13) as

$$u = A^{-1}_{\mathcal{L}} dW_m((A^{-1}_{\mathcal{L}})^* A^*_{\mathcal{C}} \xi) \equiv F(\xi) \,,$$

and consider $(q^\alpha, \xi_\alpha, \dot{u}^a)$ as a set of coordinates on W_2. The following step of the algorithm leads us to the constraints

$$\frac{\partial H}{\partial \zeta_a} + \frac{\partial F^a}{\partial \xi_\alpha}\frac{\partial H}{\partial q^\alpha} = \dot{u}^a + \frac{\partial F^a}{\partial \xi_\alpha} dW_e(q) = 0 \,.$$

In this way, we have (q^α, ξ_α) as coordinates on W_3, which turns out to be the final constraint submanifold. The dynamics of the system is described on W_3 by the differential equations

$$\dot{q}^\alpha = A_{\mathcal{C}} F(\xi) \,, \quad \dot{\xi}_\alpha = -dW_e(q) \,. \tag{3.14}$$

Thus, the application of the algorithm allows us to say that, under the given assumptions, the initial conditions in W_3 provide us with consistent optimal solutions of the dynamics of the LC-circuit.

There are of course other optimal control problems that can be interpreted in a vakonomic setting and for which this formulation can be of some help. We mention here the optimal control for nonholonomic systems with symmetry, with interesting applications to the locomotion of kinematic, and mixed kinematic and dynamic systems [18, 28, 49] or sub Riemannian geometry [11].

4. Comparison of the Vershik-Gershkovich and the vakonomic Hamiltonian approaches. In the preceding section we have found an intrinsic geometric approach to vakonomic dynamics. It is possible to give an alternative geometric formulation of the vakonomic equations of motion, related to the one of Vershik and Gershkovich [64]. A key element to obtain this alternative description will be the next fibred morphism

$$\begin{array}{rccl} F: & T^*Q \oplus TQ & \longrightarrow & T^*Q \oplus TQ \\ & (\alpha, v) & \longmapsto & (\alpha - Leg_L(v), v) \,, \end{array}$$

for any $\alpha \in T_x^*Q$, $v \in T_xQ$ and $x \in Q$. Here, $Leg_L : TQ \to T^*Q$ denotes the Legendre transformation associated with the Lagrangian L, which in local coordinates reads $Leg_L(q^A, \dot{q}^A) = (q^A, \dfrac{\partial L}{\partial \dot{q}^A})$. It is clear that $F(T^*Q \times_Q M) = T^*Q \times_Q M$. We will see how in the case of linear constraints, we "recover" the Vershik-Gershkovich formulation. As a by-product, we will have obtained a generalization of their formulation to the case of nonlinear constraints.

Consider on $T^*Q \oplus TQ$ the presymplectic 2-form $\Omega = pr_1^*\omega_Q$. Let $\omega_L = -dS^*dL$ be the Poincaré-Cartan 2-form on TQ associated with $L : TQ \to \mathbb{R}$ and E_L its energy function. Take also the presymplectic 2-form $pr_2^*\omega_L$ on $T^*Q \oplus TQ$, and define the functions

$$H = \langle pr_1, pr_2 \rangle - pr_2^* L \,, \quad \bar{H} = \langle pr_1, pr_2 \rangle - pr_2^* E_L \,.$$

Lemma 4.1. *The morphism $F : T^*Q \oplus TQ \to T^*Q \oplus TQ$ is a presymplectomorphism from $(T^*Q \oplus TQ, \Omega)$ onto $(T^*Q \oplus TQ, \Omega + pr_2^*\omega_L)$, i.e., $F^*(\Omega + pr_2^*\omega_L) = \Omega$. Moreover, it verifies $F^*\bar{H} = H$.*

Proof. F is clearly invertible with inverse

$$\begin{array}{rccc} F^{-1}: & T^*Q \oplus TQ & \longrightarrow & T^*Q \oplus TQ \\ & (\alpha, v) & \longmapsto & (\alpha + Leg(v), v) \,. \end{array}$$

A direct computation shows that $H \circ F^{-1} = \bar{H}$. Moreover, in local coordinates,

$$(F^{-1})^*(dq^A \wedge dp_A) = dq^A \wedge \left[dp_A + d\left(\frac{\partial L}{\partial \dot{q}^A} \right) \right] = dq^A \wedge dp_A + dq^A \wedge d\left(\frac{\partial L}{\partial \dot{q}^A} \right) ,$$

which implies $F^*(\Omega + pr_2^*\omega_L) = \Omega$. □

Denote by $j : T^*Q \times_Q M \hookrightarrow T^*Q \oplus TQ$ and $i : M \hookrightarrow TQ$ the respective canonical inclusions. Let us define $\bar{\omega} = j^*(\Omega + pr_2^*\omega_L)$. Since $pr_2 \circ j = i \circ \pi_2$, we have that

$$\bar{\omega} = \omega + (i \circ \pi_2)^* \omega_L \,.$$

Proposition 4.2. *The solutions of the equations*

$$i_X \omega = dH_{W_0} \,, \tag{4.1}$$

and

$$i_Y \bar{\omega} = d(j^* \bar{H}) \,, \tag{4.2}$$

are $F_{|W_0}$-related, that is, if $x \in T^*Q \times_Q M$ is a point where a solution Y of equation (4.2) exists, then $TF^{-1}(Y)$ is a solution of equation (4.1) at $F^{-1}(x)$ and, conversely, if X is a solution of equation (4.1) at $F^{-1}(x)$, then $TF(X)$ is a solution of equation (4.2) at x.

Proof. It readily follows from Lemma 4.1. □

An immediate consequence is the following

Corollary 4.3. *F preserves the constraint submanifolds provided by the presymplectic systems $(T^*Q \times_Q M, \omega, H_{W_0})$ and $(T^*Q \times_Q M, \bar{\omega}, j^*\bar{H})$. That is, if*

$$\cdots \hookrightarrow W_k \cdots \hookrightarrow W_1 \hookrightarrow W_0 = T^*Q \times_Q M \quad \textit{and}$$

$$\cdots \hookrightarrow P_k \cdots \hookrightarrow P_1 \hookrightarrow P_0 = T^*Q \times_Q M \,,$$

are the sequences of submanifolds generated by the Gotay and Nester algorithm for the first and the second presymplectic Hamiltonian system, respectively, then $F_i = F_{|W_i} : W_i \longrightarrow P_i$*, are diffeomorphisms for all* i.

In conclusion, Proposition 4.2 and Corollary 4.3 show that solving the vakonomic Hamiltonian equations (4.1) as in § 3 is equivalent to solving equations (4.2). Locally, if $(q^A(t), p_A(t), \dot{q}^a(t))$ is an integral curve of X then

$$(q^A(t), p_A - i^* \frac{\partial L}{\partial \dot{q}^A}(q^B(t), \dot{q}^b(t)), \dot{q}^a(t))$$

is an integral curve of Y.

4.1. Vershik-Gershkovich approach. In [64], Vershik and Gershkovich gave a formulation for the "nonholonomic variational problem", i.e., the vakonomic problem, within the framework of the so-called mixed bundle picture, which we briefly review in the following (see also [7]).

If $\mathcal{D} : Q \longrightarrow TQ$ is a differentiable distribution along Q, then the mixed bundle over Q associated with $\mathcal{D}$ is given by $\mathcal{D} \oplus \mathcal{D}^o$, where $\mathcal{D}^o$ is the codistribution annihilating $\mathcal{D}$; the fibres of $\mathcal{D} \oplus \mathcal{D}^o \longrightarrow Q$ are $\mathcal{D}_q \oplus \mathcal{D}^o_q$.

Let $\{ \Phi^\alpha(q^A, \dot{q}^A) = \Psi^\alpha_a(q)\dot{q}^a - \dot{q}^\alpha \, , 1 \leq \alpha \leq m \}$ be a set of independent functions whose annihilation defines the distribution $\mathcal{D}$, and let $\{ \eta^\alpha = \Psi^\alpha_a dq^a - dq^\alpha \, , 1 \leq \alpha \leq m \}$ be the corresponding basis of $\mathcal{D}^o$. Regarding $\mathcal{D} \subset TQ$ as the set of admissible velocities, Vershik and Gershkovich write the equations of motion (2.4) for the vakonomic problem $(L, \mathcal{D})$ as follows

$$\left\{ \begin{array}{l} \left(\dfrac{d}{dt}\left(\dfrac{\partial L}{\partial \dot{q}^A}\right) - \dfrac{\partial L}{\partial q^A} \right) dq^A = \dot{\lambda}_\alpha \eta^\alpha + \lambda_\alpha (i_{\dot{q}} d\eta^\alpha) \, , \\ \\ \langle \dot{q}, \eta^\alpha \rangle = 0 \, , \; 1 \leq \alpha \leq m \, . \end{array} \right. \tag{4.3}$$

In this particular case, we obtain that P_1, the first constraint submanifold for the presymplectic Hamiltonian system $(T^*Q \times_Q M, \bar{\omega}, j^* \bar{H})$, is just $\mathcal{D}^o \oplus \mathcal{D}$, since we get $\lambda_a + \lambda_\alpha \Psi^\alpha_a = 0, \; 1 \leq \alpha \leq m$.

If $(P_1 = \mathcal{D}^o \oplus \mathcal{D}, \omega_{P_1})$ is a symplectic manifold (see Proposition 3.5), then the equations of motion (4.3) determine a unique vector field on $\mathcal{D}^o \oplus \mathcal{D}$ and the Lagrange multipliers λ_α are coordinates in $\mathcal{D}^o$ with respect to the basis η^α.

Consequently, the geometrical picture we have developed in § 3 is equivalent to the Vershik-Gershkovich approach. As said above, we have obtained a generalization of the Vershik-Gershkovich formulation to the case of nonlinear constraints, just "translating" things from our approach by the diffeomorphism F.

In the nonlinear case, under the admissibility condition, one can verify that the first constraint submanifold $P_1 = F(W_1)$ can be identified with the manifold $S^*(TM^o) \times_Q M$. In fact, we have that $S^*(TM^o)$ is generated by the 1-forms

$$S^* d\Phi^\alpha = dq^\alpha - \frac{\partial \Psi^\alpha}{\partial \dot{q}^a} dq^a \, , \; 1 \leq \alpha \leq m \, .$$

If $(q^A, \lambda_A, \dot{q}^a) \in P_1$, then the 1-form $\lambda_A dq^A$ is a linear combination of the 1-forms $S^* d\Phi^\alpha$ in the following manner: $\lambda_A dq^A = \lambda_\alpha S^* d\Phi^\alpha$.

5. Geometric approach to nonholonomic mechanics. A nonholonomic Lagrangian system consists of a Lagrangian $L : TQ \to \mathbb{R}$ subject to nonholonomic constraints defined by m local functions $\Phi^\alpha(q^A, \dot{q}^A)$, $1 \leq \alpha \leq m$. The equations of motion for nonholonomic mechanics are derived assuming that the constraints satisfy d'Alembert's principle, in the linear or affine case. In the nonlinear case, there does not seem to exist a general consensus concerning the correct principle to adopt [41, 51]. The most widely used model is based on Chetaev's principle, which will also be adopted in the present paper. The equations of motion are then given by

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^A}\right) - \frac{\partial L}{\partial q^A} = \lambda_\alpha \frac{\partial \Phi^\alpha}{\partial \dot{q}^A} \ , \tag{5.1}$$

together with the algebraic equations $\Phi^\alpha(q^A, \dot{q}^A) = 0$. The functions λ_α, $1 \leq \alpha \leq m$, are some Lagrange multipliers to be determined. As in the vakonomic case, we assume the admissibility condition, so it is possible to write the constraints as $\dot{q}^\alpha = \Psi^\alpha(q^A, \dot{q}^a)$, where $1 \leq \alpha \leq m$, $m+1 \leq a \leq n$ and $1 \leq A \leq n$.

The study of nonholonomic systems in the realm of Geometric Mechanics started with the work by Vershik and Faddeev [62, 63] and has been an active area of research since then, with many contributions from different authors (see [16] for a recent survey). In particular, the role of symmetry has been treated extensively in the literature, starting with the work by Koiller [27] and going through the use of the Hamiltonian formalism [2], Lagrangian reduction [10], the geometry of the tangent bundle [12, 13, 17, 33] or Poisson methods [40], among others.

Nonholonomic mechanics also admits a nice geometrical description on the space $T^*Q \oplus TQ$ inspired on the Skinner and Rusk formalism [58]. In addition, this novel description will be appropriate to compare the solutions of the dynamics between the vakonomic and nonholonomic mechanics. In the following, we will prove that equations (5.1) can be intrinsically written as

$$\left\{ \begin{array}{l} (i_X\Omega - dH)_{|T^*Q\times_Q M} \in F^o \ , \\ X_{|T^*Q\times_Q M} \in T(T^*Q \times_Q M) \ , \end{array} \right. \tag{5.2}$$

where Ω is the presymplectic 2-form $\Omega = pr_1^*\omega_Q$ on $T^*Q \oplus TQ$, H the Hamiltonian function $H = \langle pr_1, pr_2\rangle - pr_2^*L$ and F^o the subbundle of $T^*(T^*Q\oplus TQ)$ along $T^*Q \times_Q M$ defined by $F^o = pr_2^*(S^*(TM^o))$, representing the constraint forces.

Indeed we have in local coordinates

$$\Omega = dq^A \wedge dp_A \ , \quad dH = \dot{q}^A dp_A + p_A d\dot{q}^A - \frac{\partial L}{\partial q^A} dq^A - \frac{\partial L}{\partial \dot{q}^A} d\dot{q}^A \ ,$$

and F^o is generated by the 1-forms

$$\frac{\partial \Phi^\alpha}{\partial \dot{q}^A} dq^A = \frac{\partial \Psi^\alpha}{\partial \dot{q}^a} dq^a - dq^\alpha \ , \ 1 \leq \alpha \leq m \ .$$

If $X = X^A \dfrac{\partial}{\partial q^A} + Y^A \dfrac{\partial}{\partial \dot{q}^A} + Z_A \dfrac{\partial}{\partial p_A}$ was a solution of equations (5.2), then

$$X^A = \dot{q}^A \ , \quad Z_A = \frac{\partial L}{\partial q^A} + \lambda_\alpha \frac{\partial \Phi^\alpha}{\partial \dot{q}^A} \ , \tag{5.3}$$

along with the constraints

$$p_A - \frac{\partial L}{\partial \dot{q}^A} = 0 \ , \quad \Phi^\alpha(q^A, \dot{q}^A) = 0 \ . \tag{5.4}$$

Observe that these constraints determine a submanifold $\tilde{M}$ of $T^*Q \times_Q M$. The submanifold $\tilde{M}$ is diffeomorphic to M since

$$\begin{array}{ccl} M & \longrightarrow & \tilde{M} \\ m & \longmapsto & (Leg_L(m), m)\,, \end{array}$$

is a diffeomorphism. $\tilde{M}$ is the first constraint submanifold provided by the constraint algorithm applied to equations (5.2). This algorithm will lead to a final constraint submanifold on which there exists a well-defined dynamics. Obviously, equations (5.3) and (5.4) are equivalent to the nonholonomic equations of motion (5.1).

In terms of the Ψ^α's the above equations can be written as

$$X^A = \dot{q}^A\,, \quad Z_a = \frac{\partial L}{\partial q^a} + \lambda_\alpha \frac{\partial \Psi^\alpha}{\partial \dot{q}^a}\,, \quad Z_\beta = \frac{\partial L}{\partial q^\beta} - \lambda_\beta\,,$$

together with the constraints

$$p_A - \frac{\partial L}{\partial \dot{q}^A} = 0\,, \quad \dot{q}^\alpha - \Psi^\alpha(q^A, \dot{q}^a) = 0\,. \tag{5.5}$$

Therefore, a solution X of (5.2) is of the form

$$\begin{aligned} X = \dot{q}^a &\left(\frac{\partial}{\partial q^a} + \frac{\partial \Psi^\alpha}{\partial q^a}\frac{\partial}{\partial \dot{q}^\alpha} + \left(\frac{\partial^2 L}{\partial \dot{q}^A \partial q^a} + \frac{\partial \Psi^\alpha}{\partial q^a}\frac{\partial^2 L}{\partial \dot{q}^A \partial \dot{q}^\alpha} \right)\frac{\partial}{\partial p_A} \right) \\ +\Psi^\gamma &\left(\frac{\partial}{\partial q^\gamma} + \frac{\partial \Psi^\alpha}{\partial q^\gamma}\frac{\partial}{\partial \dot{q}^\alpha} + \left(\frac{\partial^2 L}{\partial \dot{q}^A \partial q^\gamma} + \frac{\partial \Psi^\alpha}{\partial q^\gamma}\frac{\partial^2 L}{\partial \dot{q}^A \partial \dot{q}^\alpha} \right)\frac{\partial}{\partial p_A} \right) \\ +Y^a &\left(\frac{\partial}{\partial \dot{q}^a} + \frac{\partial \Psi^\alpha}{\partial \dot{q}^a}\frac{\partial}{\partial \dot{q}^\alpha} + \left(\frac{\partial^2 L}{\partial \dot{q}^A \partial \dot{q}^a} + \frac{\partial \Psi^\alpha}{\partial \dot{q}^a}\frac{\partial^2 L}{\partial \dot{q}^A \partial \dot{q}^\alpha} \right)\frac{\partial}{\partial p_A} \right)\,. \end{aligned}$$

Under the regularity assumption, which here means that the matrix

$$\tilde{\mathcal{C}}_{ab} = \frac{\partial^2 \tilde{L}}{\partial \dot{q}^a \partial \dot{q}^b} - i^*\!\left(\frac{\partial L}{\partial \dot{q}^\alpha}\right)\frac{\partial^2 \Psi^\alpha}{\partial \dot{q}^a \partial \dot{q}^b}\,, \tag{5.6}$$

is invertible (see [57]), there is a unique solution of the dynamics on $\tilde{M}$. In particular, after some computations we obtain

$$Y^a = -\tilde{\mathcal{C}}^{ab}\left[\dot{q}^A \frac{\partial^2 \tilde{L}}{\partial q^A \partial \dot{q}^b} - \dot{q}^A i^*\!\left(\frac{\partial L}{\partial \dot{q}^\alpha}\right)\frac{\partial^2 \Psi^\alpha}{\partial q^A \partial \dot{q}^b} - \frac{\partial \tilde{L}}{\partial q^b} + i^*\!\left(\frac{\partial L}{\partial q^\alpha}\right)\left(\frac{\partial \Psi^\alpha}{\partial q^b} - \frac{\partial \Psi^\alpha}{\partial \dot{q}^b}\right)\right]\,,$$

where $i : M \to TQ$ is the canonical inclusion and $(\tilde{\mathcal{C}}^{ab})$ the inverse matrix of $(\tilde{\mathcal{C}}_{ab})$.

Taking coordinates $(q^A, \dot{q}^a)$ on $\tilde{M}$, the equations of motion for a nonholonomic system will be

$$\left\{ \begin{aligned} \dot{q}^\alpha &= \Psi^\alpha(q^A, \dot{q}^a)\,, \\ \ddot{q}^a &= -\mathcal{C}^{ab}\left[\dot{q}^A \frac{\partial^2 \tilde{L}}{\partial q^A \partial \dot{q}^b} - \dot{q}^A i^*\!\left(\frac{\partial L}{\partial \dot{q}^\alpha}\right)\frac{\partial^2 \Psi^\alpha}{\partial q^A \partial \dot{q}^b} - \frac{\partial \tilde{L}}{\partial q^b} + i^*\!\left(\frac{\partial L}{\partial q^\alpha}\right)\left(\frac{\partial \Psi^\alpha}{\partial q^b} - \frac{\partial \Psi^\alpha}{\partial \dot{q}^b}\right)\right] \end{aligned} \right. \tag{5.7}$$

6. Vakonomic and nonholonomic mechanics: Equivalence of dynamics. In this section, we shall investigate the relation between vakonomic and nonholonomic dynamics. Consider a physical system with Lagrangian $L : TQ \to \mathbb{R}$ and constraint submanifold $M \subset TQ$. Let us assume that the vakonomic problem lives on the first constraint submanifold, W_1, and that the nonholonomic one lives on $\tilde{M}$ (this will be the case if the constraints are linear and the admissibility and compatibility conditions are satisfied). As a consequence, we have well defined vector fields X_{vk} on W_1 and X_{nh} on $\tilde{M}$. It is clear that the mapping $(\pi_2)_{|W_1} : W_1 \to M$ is a surjective submersion and that we can define the mapping $\Upsilon : W_1 \to \tilde{M}$ as

$$\begin{array}{rccc} \Upsilon : & W_1 & \longrightarrow & \tilde{M} \\ & (\alpha, v) & \longmapsto & (Leg_L(v), v) \end{array}$$

In coordinates, Υ reads as $\Upsilon(q^A, \dot{q}^a, p_\alpha) = (q^A, \dot{q}^a)$.

Our aim is to know whether, given a solution of the nonholonomic problem, we can find initial conditions in the vakonomic Lagrange multipliers, p_α, so that the curve can also be seen as a solution of the vakonomic problem. In order to capture the common solutions to both systems, we have developed the following algorithm. It is inspired on the idea of the Υ-relation of X_{vk} and X_{nh} and the constraint algorithm developed by O. Krupková [30]. If both fields were Υ-related, then the projection to $\tilde{M}$ of all the vakonomic solutions would be nonholonomic. So, selecting those points where both vector fields are related, we are picking up all the possible good candidates. We write $W_1 = S_0$ and define

$$S_1 = \{ w \in S_0 \mid T_w\Upsilon(X_{vk}(w)) = X_{nh}(\Upsilon(w)) \} .$$

In general S_1 is not a submanifold. If $S_1 = \emptyset$, there is no relation between the vakonomic and nonholonomic dynamics. If $S_1 \neq \emptyset$, we apply the following algorithm:

- **Step 1**: For any $w \in S_1$, consider $C_{(w)} = \cup_i C_{(w)i}$, the union of all connected submanifolds $C_{(w)i}$ of maximal dimension lying in S_1, contained in a neighborhood U of w and passing through w (maximal dimension means that if N is a connected submanifold lying in $S_1 \cap U$ passing through w and $C_{(w)i} \subseteq N$, then $C_{(w)i} = N$).
  Suppose that $C_{(w)} \neq \{w\}$. For each i we consider the subset of $C_{(w)i}$
  $$\tilde{C}_{(w)i} = \{ v \in C_{(w)i} \mid X_{vk}(v) \in T_v C_{(w)i} \} .$$
  If $\tilde{C}_{(w)i} = C_{(w)i}$ then we call the submanifold $C_{(w)i}$ a **final constraint submanifold at** w. If $\tilde{C}_{(w)i} = \emptyset$, we exclude $C_{(w)i}$ from the collection $C_{(w)}$. If $\emptyset \subsetneq \tilde{C}_{(w)i} \subsetneq C_{(w)i}$, then we proceed to the next step.
- **Step 2**: Repeat the Step 1 with $\tilde{C}_{(w)i}$ instead of S_1.

After a sufficient number of steps in this algorithm we either obtain a collection of final constraint submanifolds at w, or we find that there is no final constraint submanifold passing through w. Collecting all the points where there exist such final constraint submanifolds, we obtain the subset of W_1 where there is equivalence between vakonomic and nonholonomic dynamics.

Suppose that the constraints Φ^α, $1 \leq \alpha \leq m$, are linear in the velocities so we can write them as $\dot{q}^\alpha = \Psi^\alpha_a(q)\dot{q}^a$. In such a case, the matrices $\mathcal{C}$ and $\tilde{\mathcal{C}}$ defined in (3.9) and (5.6), respectively, are the same (even for constraints affine in the velocities).

PROPOSITION 6.1. *S_1 is locally characterized by the vanishing of the $n-m$ constraints functions on W_1*

$$g_b = \dot{q}^a \left(p_\alpha - i^* \frac{\partial L}{\partial \dot{q}^\alpha} \right) \left[\frac{\partial \Psi^\alpha_b}{\partial q^a} - \frac{\partial \Psi^\alpha_a}{\partial q^b} + \Psi^\beta_a \frac{\partial \Psi^\alpha_b}{\partial q^\beta} - \Psi^\beta_b \frac{\partial \Psi^\alpha_a}{\partial q^\beta} \right], \quad m+1 \leq b \leq n \,. \tag{6.1}$$

Proof. The comparison between the vector fields X_{vk} and X_{nh} consists of taking the difference between $\ddot{q}^a$'s in the expressions (3.8) and (5.7) and equating the result to zero. □

Consider the local projection $\rho(q^a, q^\alpha) = (q^\alpha)$ and the connection Γ on ρ such that the horizontal distribution $\mathcal{H}$ is given by prescribing its annihilator to be $\mathcal{H}^o = \langle dq^\alpha - \Psi^\alpha_a dq^a, 1 \leq \alpha \leq m \rangle$. Then the curvature R of this connection (see [34]) is given by $R(\frac{\partial}{\partial q^a}, \frac{\partial}{\partial q^b}) = R^\alpha_{ab} \frac{\partial}{\partial q^\alpha}$, where

$$R^\alpha_{ab} = \frac{\partial \Psi^\alpha_b}{\partial q^a} - \frac{\partial \Psi^\alpha_a}{\partial q^b} + \Psi^\beta_a \frac{\partial \Psi^\alpha_b}{\partial q^\beta} - \Psi^\beta_b \frac{\partial \Psi^\alpha_a}{\partial q^\beta} \,.$$

We say that Γ is flat if the curvature R vanishes identically. The tensor R measures the lack of integrability of the horizontal distribution $\mathcal{H}$, which in our case is the constraint manifold. Then, we can write the constraints determining S_1 as

$$g_b = \dot{q}^a \left(p_\alpha - i^* \frac{\partial L}{\partial \dot{q}^\alpha} \right) R^\alpha_{ab}, \quad m+1 \leq b \leq n \,.$$

From this expression we deduce that if the constraints are holonomic, then $R = 0$ and the final constraint submanifold is equal to $S_0 = W_1$. Therefore, every solution of the nonholonomic problem is also a vakonomic solution. Indeed, equations (3.3-3.5) will read as

$$\left\{ \begin{array}{l} \dot{q}^\alpha = \Psi^\alpha_a \dot{q}^a \,, \\ \dot{p}_\alpha = \dfrac{\partial \tilde{L}}{\partial q^\alpha} - p_\beta \dfrac{\partial \Psi^\beta_a}{\partial q^\alpha} \dot{q}^a \,, \\ \dfrac{d}{dt} \left(\dfrac{\partial \tilde{L}}{\partial \dot{q}^a} \right) - \dfrac{\partial \tilde{L}}{\partial q^a} = \Psi^\alpha_a \dfrac{\partial \tilde{L}}{\partial q^\alpha} \,. \end{array} \right. \tag{6.2}$$

The first and the third set of equations determine the trajectory in M. The Lagrange multipliers p_α are determined by the second set of equations once we know the solution in M. This is the typical behavior of the holonomic case [1, 36]. But, in general, for linear constraints, the first constraint subset in the algorithm is determined by

$$S_1 = \{ g_b = 0, \; m+1 \leq b \leq n \} \,,$$

where $g_b(q^A, \dot{q}^a, p_\alpha) = \dot{q}^a R^\alpha_{ab}(q)(p_\alpha - \frac{\partial L}{\partial \dot{q}^\alpha})$. Note that S_1 will not be a submanifold, because 0 is not a regular value of the functions g_b, $b = m+1, ..., n$. Anyway, the geometric context we have developed can be very useful to tackle the problem of the comparison of the two methods.

PROPOSITION 6.2. *If $c(t) = (q^A(t))$ is a solution of the unconstrained problem which, in addition, verifies all the constraints, i.e,*

$$\dot{q}^\alpha(t) = \Psi^\alpha_a(q(t)) \dot{q}^a(t) \,, \; 1 \leq \alpha \leq m \,,$$

then $c(t)$ is a solution of the nonholonomic and vakonomic problems simultaneously.

Proof. Let us consider the submanifold $S := \{p_\alpha = i^*\left(\frac{\partial L}{\partial \dot{q}^\alpha}\right)\}$, which is contained in S_1. A natural question is whether the vakonomic vector field will be tangent to S, that is, $X_{vk} \in TS$. From equations (3.3-3.5), we have along any integral curve of the vakonomic vector field

$$X_{vk} \in TS \Longleftrightarrow \frac{d}{dt}\left(p_\alpha - i^* \frac{\partial L}{\partial \dot{q}^\alpha}\right) = 0 \Longleftrightarrow \dot{p}_\alpha = \dot{q}^A \frac{\partial^2 L}{\partial q^A \partial \dot{q}^\alpha} + \ddot{q}^a \frac{\partial^2 L}{\partial \dot{q}^a \partial \dot{q}^\alpha} .$$

On S, we have that

$$\dot{p}_\alpha = \frac{\partial \tilde{L}}{\partial q^\alpha} - p_\beta \frac{\partial \Psi^\beta}{\partial q^\alpha} = \frac{\partial \tilde{L}}{\partial q^\alpha} - \frac{\partial L}{\partial \dot{q}^\beta} \frac{\partial \Psi^\beta}{\partial q^\alpha} = \frac{\partial L}{\partial q^\alpha} .$$

Then the above condition can be rewritten as

$$\frac{\partial L}{\partial q^\alpha} = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}^\alpha}\right) ,$$

that with equations (3.5) are precisely the Euler-Lagrange equations. Then, we have proved that a solution $c(t)$ of the unconstrained problem satisfies the constraints if and only if

$$(q^A(t), i^*\left(\frac{\partial L}{\partial \dot{q}^\alpha}\right), \dot{q}^a(t)) ,$$

is a solution of the vakonomic equations (3.3). Since the constraints $g_b = 0$ are automatically satisfied for all the points in S we deduce that $c(t)$ is also a solution of the nonholonomic problem. □

Remark 6.3. As a consequence of Proposition 6.2 we obtain that if g is a Riemannian metric on Q , with kinetic energy $L = \frac{1}{2}g$, and if we assume that we are given a distribution $\mathcal{D}$ on Q which is geodesically invariant with respect to the Levi-Civita connection ∇^g, then all the nonholonomic solutions can be seen as vakonomic ones. In fact, they all are solutions of the free problem. This last result was first stated in [20] (Theorem 3.2) with additional hypothesis on the nature of the metric g and the integrability of $\mathcal{D}^{\perp_g}$ which are not essential, as we have seen.

Remark 6.4. Let $\Theta : G \times Q \longrightarrow Q$ be a free and proper action on Q. Then $\pi : Q \longrightarrow Q/G$ is a principal G-bundle. Assume that the Lagrangian $L : TQ \longrightarrow \mathbb{R}$ is G-invariant and is subject to equivariant affine constraints, M, such that its linear part $\mathcal{D}$ is the horizontal distribution of a principal connection γ on $\pi : Q \longrightarrow Q/G$. Then, we have the following result, which is an adaptation of Theorem 3.1 in [20] to our geometric description of vakonomic and nonholonomic mechanics.

Proposition 6.5. *Assume that the admissibility and compatibility conditions hold. Then, the following are equivalent:*

1. *the solution of the nonholonomic problem $(q^A(t), \dot{q}^a(t)) \in \tilde{M}$ verifies the condition $g_b(q^A(t), \dot{q}^a(t), p_0) = 0$ for some p_0, $m + 1 \leq b \leq n$.*
2. *the curve $(q^A(t), \dot{q}^a(t), p_0) \in W_1$ is a vakonomic solution.*

Example 6.6 (Rolling penny [4]). Consider a vertical penny constrained to roll without slipping on a horizontal plane. Let (x, y) denote the position of contact of the disk in the plane, θ the orientation of a chosen material point P with respect to the vertical and ϕ the heading angle of the penny. The configuration space is then

$Q = \mathbb{R}^2 \times \mathbb{S}^1 \times \mathbb{S}^1$. The Lagrangian may be written as $L = (\dot{x}^2 + \dot{y}^2 + \dot{\theta}^2 + \dot{\phi}^2)/2$ and the constraints are given by $\dot{x} = \dot{\theta}\cos\phi$, $\dot{y} = \dot{\theta}\sin\phi$. For simplicity, we assume that the mass m, the moments of inertia I, J and the radius of the penny R are 1.

Applying the algorithm, we obtain the final constraint submanifolds

$$\begin{aligned} &C_{f_1} = \{w \in W_1 \mid \dot{\phi} = 0\}\,, \quad C_{f_2} = \{w \in W_1 \mid 2\dot{\theta} = p_x\cos\phi + p_y\sin\phi\}\,, \\ &C_{f_3} = \{w \in W_1 \mid \dot{\theta} = 0\,,\ \dot{\phi} = 0\}\,. \end{aligned}$$

The nonholonomic solutions living on C_{f_1} are motions of the penny along a straight line in the horizontal plane. The nonholonomic solutions in C_{f_3} are stationary positions. However, any nonholonomic solution can be seen as a vakonomic one contained in C_{f_2}, with Lagrange multipliers $p_x = 2\dot{\theta}\cos\phi$ and $p_y = 2\dot{\theta}\sin\phi$. In terms of the extended Lagrangian formalism mentioned in Remark 2.4, we have the following Lagrange multipliers

$$\lambda_x = \frac{\partial L}{\partial x} - p_x = \dot{x} - p_x = -\dot{\theta}\cos\phi\,, \quad \lambda_y = \frac{\partial L}{\partial y} - p_y = \dot{y} - p_y = -\dot{\theta}\sin\phi\,,$$

which is just the result of Bloch and Crouch [4].

Example 6.7 (Planar mobile robot). Consider the motion of a two-wheeled planar mobile robot which is able to move in the direction in which it points and, in addition, can spin about a vertical axis [26, 29, 33]. Let P be the intersection point of the horizontal symmetry axis of the robot and the horizontal line connecting the centers of the two wheels. The position and orientation of the robot is determined by $(x, y, \theta) \in SE(2)$, where $\theta \in \mathbb{S}^1$ is the heading angle and the coordinates $(x, y) \in \mathbb{R}^2$ locate the point P. Let ψ_1, $\psi_2 \in \mathbb{S}^1$ denote the rotation angles of the wheels which are assumed to be controlled independently and roll without slipping on the floor. The configuration space of this system is $Q = \mathbb{S}^1 \times \mathbb{S}^1 \times SE(2)$.

The Lagrangian function is the kinetic energy of the system

$$L = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) + m_0 l\dot{\theta}(\cos\theta\dot{y} - \sin\theta\dot{x}) + \frac{1}{2}J\dot{\theta}^2 + \frac{1}{2}J_2(\dot{\psi}_1^2 + \dot{\psi}_2^2)\,,$$

where $m = m_0 + 2m_1$, m_0 is the mass of the robot without the wheels, J its moment of inertia with respect to the vertical axis, m_1 the mass of each wheel, J_2 the axial moments of inertia of the wheels, and l the distance between the center of mass C of the robot and the point P. The constraints are induced by the conditions that there is no lateral sliding of the robot and that the motion of the wheels also consists of a rolling without sliding,

$$\dot{x} = -R\cos\theta(\dot{\psi}_1 + \dot{\psi}_2)/2\,,\ \dot{y} = -R\sin\theta(\dot{\psi}_1 + \dot{\psi}_2)/2\,,\ \dot{\theta} = R(\dot{\psi}_2 - \dot{\psi}_1)/(2c)\,,$$

where R is the radius of the wheels and $2c$ the lateral length of the robot.

This example is very interesting because its qualitative behavior changes completely depending on the parameters. If $l = 0$ (namely, the point P is the center of mass of the robot), application of the algorithm yields the following constraint submanifolds

$$\begin{aligned} &C_{f_1} = \{w \in W_1 \mid p_x\sin\theta - p_y\cos\theta = 0\,,\ \dot{\psi}_1 = \dot{\psi}_2\}\,, \\ &C_{f_2} = \{w \in W_1 \mid p_x = 0\,,\ p_y = 0\}\,, \quad C_{f_3} = \{w \in W_1 \mid \dot{\psi}_1 = 0\,,\ \dot{\psi}_2 = 0\}\,. \end{aligned}$$

If $l \neq 0$ and $K_1 \neq K_2^2$, with $K_1 = J_2(J_2 + mR^2/2 + R^2J/2c^2) + mR^3J/4c^2$, $K_2 = m_0lR^2/2c$, we find that

$$C_{f\,1} = \{w \in W_1 \mid p_x \sin\theta - p_y \cos\theta = 0\,,\ \dot\psi_1 = \dot\psi_2\}\,,$$
$$C_{f\,2} = \{w \in W_1 \mid \dot\psi_1 = 0\,,\ \dot\psi_2 = 0\}\,,$$

whereas if $K_1 = K_2^2$, we obtain an additional final constraint submanifold

$$\begin{aligned} C_{f\,3} = \{w \in W_1 \mid p_x &= K_2(\dot\psi_1 - \dot\psi_2)\sin\theta/R - 2K_2^2(\dot\psi_1 + \dot\psi_2)\cos\theta/R(2K_1/J_2 - mR^2)\,,\\ p_y &= -2K_2^2(\dot\psi_1 + \dot\psi_2)\sin\theta/R(2K_1/J_2 - mR^2) - K_2(\dot\psi_1 - \dot\psi_2)\cos\theta/R\}\,,\end{aligned}$$

Therefore, in the cases $l = 0$ and $l \neq 0$, $K_1 = K_2^2$, every nonholonomic solution can be seen as a vakonomic one. This has the following interesting physical interpretation: under an appropriate design of the robot (i.e., choice of the parameters), the trajectories that it describes between two points are optimal, in the sense that they minimize the energy cost among all the other possible trajectories satisfying the constraints and connecting the given points.

EXAMPLE 6.8 (Ball on a rotating table [36]). Applying the algorithm to this example, one can recover the result found in [36]. The configuration space is $Q = \mathbb{R}^2 \times SO(3)$ with coordinates (x, y, R). We denote the spatial angular velocity by $\xi \in \mathbb{R}^3$, where $\hat\xi = \dot R R^T$. The Lagrangian is $L = I((\xi^1)^2 + (\xi^2)^2 + (\xi^3)^2)/2 + m(\dot x^2 + \dot y^2)/2$, where I and m are the moment of inertia and mass of the ball, respectively. The constraints are $\dot x = r\xi^2 - \Omega y$, $\dot y = -r\xi^1 + \Omega x$, where r is the radius of the ball and Ω is the angular velocity of the table.

Applying the algorithm, one finds the following final constraint submanifolds

$$C_{f\,1} = \{w \in W_1 \mid \dot x = \dot y = p_x = p_y = 0\}\,, \quad C_{f\,2} = \{w \in W_1 \mid \xi^3 = \Omega\}\,.$$

There are nonholonomic solutions that can not be seen as vakonomic ones (see [36]).

Acknowledgments. J. Cortés, S. Martínez and D. Martín de Diego would like to thank the Department of Mathematical Physics and Astronomy of the University of Ghent for its kind hospitality. We would like to thank F. Cantrijn and E. Martínez for several helpful conversations, and the referees for their careful reading of the manuscript and their useful remarks that helped us improve the quality of the paper.

REFERENCES

[1] V.I. Arnold: *Dynamical Systems*, vol. III. Springer-Verlag, New York-Heidelberg-Berlin, 1988.

[2] L. Bates, J. Śniatycki: Nonholonomic reduction. *Rep. Math. Phys.* **32** (1) (1992), 99-115.

[3] E. Binz, J. Śniatycki, H. Fisher: *The Geometry of Classical Fields*. North-Holland Math. Ser. 154, Amsterdam, 1988.

[4] A.M. Bloch, P.E. Crouch: Nonholonomic and Vakonomic control systems on Riemannian manifolds. In *Dynamics and Control of Mechanical Systems, The Falling Cat and Related Problems*, Michael J. Enos, ed., Fields Institute Communications, AMS, 1993, pp. 25-52.

[5] A.M. Bloch, P.E. Crouch: Nonholonomic control systems on Riemannian manifolds. *SIAM J. Control Optim.* **33** (1995), 126-148.

[6] A.M. Bloch, P.E. Crouch: On the equivalence of higher order variational problems and optimal control problems. *Proc. IEEE Int. Conf. Decision&Control*, Kobe, Japan, 1996, pp. 1648-1653.

[7] A.M. Bloch, P.E. Crouch: Newton's law and integrability of nonholonomic systems. *SIAM J. Control Optim.* **36** (1998), 2020-2039.

[8] A.M. Bloch, P.E. Crouch: Optimal Control, Optimization, and Analytical Mechanics. In *Mathematical Control Theory*, J. Baillieul, J.C. Willems, eds., Springer Verlag, New York, 1998, pp. 268-321.

[9] A.M. Bloch, P.E. Crouch: Constrained variational principles on manifolds. *Proc. IEEE Int. Conf. Decision&Control*, Phoenix, USA, 1999, pp. 1-6.

[10] 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.

[11] R.W. Brockett: Control Theory and Singular Riemannian Geometry. In *New Directions in Applied Mathematics*, P.J. Hilton, G.S. Young, eds., Springer-Verlag, New York, 1982.

[12] F. Cantrijn, J. Cortés, M. de León, D. Martín de Diego: On the geometry of generalized Chaplygin systems. *Math. Proc. Cambridge Philos. Soc.* **132** (2) (2002), 323-351.

[13] F. Cantrijn, M. de León, J.C. Marrero, D. Martín de Diego: Reduction of nonholonomic mechanical systems with symmetries. *Rep. Math. Phys.* **42** (1998), 25-45.

[14] F. Cardin, M. Favretti: On nonholonomic and vakonomic dynamics of mechanical systems with nonintegrable constraints, *J. Geom. Phys.* **18** (4) (1996), 295-325.

[15] J.F. Cariñena, C. López, M.F. Rañada: Geometric Lagrangian approach to first-order systems and applications, *J. Math. Phys.* **29** (5) (1988), 1134-1142.

[16] H. Cendra, J.E. Marsden, T.S. Ratiu: Geometric mechanics, Lagrangian reduction and nonholonomic systems. In *Mathematics Unlimited-2001 and Beyond*, B. Enguist, W. Schmid, eds., Springer-Verlag, New York, 2001, pp. 221-273.

[17] J. Cortés, M. de León: Reduction and reconstruction of the dynamics of nonholonomic systems. *J. Phys. A: Math. Gen.* **32** (1999), 8615-8645.

[18] J. Cortés, S. Martínez: Optimal control for nonholonomic systems with symmetry. *Proc. IEEE Int. Conf. Decision&Control*, Sydney, Australia, 2000, pp. 5216-5218.

[19] J. Cortés, S. Martínez: Nonholonomic integrators. *Nonlinearity* **14** (5) (2001), 1365-1392.

[20] M. Favretti: Equivalence of dynamics for nonholonomic systems with transverse constraints. *J. Dynam. Differential Equations* **10** (4) (1998), 511-536.

[21] M. Gotay, J. Nester: Presymplectic Lagrangian systems I: the constraint algorithm and the equivalence theorem. *Ann. Inst. Henri Poincaré* **30** 2 (1978), 129-142.

[22] A. Ibort, J. Marín-Solano: A geometric approach to optimal control theory and the inverse problem of the calculus of variations, preprint, 1998.

[23] V. Jurdjevic: *Geometric Control Theory*. Cambridge Studies in Advanced Mathematics, vol. 52, Cambridge University Press, 1997.

[24] V. Jurdjevic: Optimal Control, Geometry and Mechanics. In *Mathematical Control Theory*, J. Baillieul, J.C. Willems, eds., Springer Verlag, New York, 1998, pp. 227-267.

[25] H. Kataoka, H. Hashimoto: New conservation laws in a neoclassical von Neumann model. *J. Math. Econ.* **36** (8) (1995), 271-280.

[26] S.D. Kelly, R.M. Murray: Geometric phases and robotic locomotion. *J. Robotic Systems* **12** (6) (1995), 417-431.

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

[28] W.S. Koon, J.E. Marsden: Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction. *SIAM J. Control Optim.* **35** (1997), 901-929.

[29] V.V. Kozlov: Realization of nonintegrable constraints in classical mechanics. *Dokl. Akad. Nauk SSSR* **272** (3) (1983), 550-554; English transl.: *Sov. Phys. Dokl.* **28** (9) (1983), 735-737.

[30] O. Krupková: *The Geometry of Ordinary Variational Equations.* Lectures Notes in Mathematics Series, vol 1678, Springer-Verlag, New York-Heidelberg-Berlin, 1997.

[31] M. de León, J.C. Marrero, D. Martín de Diego: Vakonomic mechanics versus nonholonomic mechanics: a unified geometrical approach. *J. Geom. Phys.* **35** (2000), 126-144.

[32] M. de León, D. Martín de Diego: Symmetries and constants of the motion for singular Lagrangian systems. *Internat. J. Theoret. Phys.* **35** (5) (1996), 975-1011.

[33] M. de León, D. Martín de Diego: On the geometry of non-holonomic Lagrangian systems. *J. Math. Phys.* **37** (7) (1996), 3389-3414.

[34] M. de León, P.R. Rodrigues: *Methods of Differential Geometry in Analytical Mechanics.* North-Holland Math. Ser. 152, Amsterdam, 1989.

[35] A.D. Lewis: Simple mechanical control systems with constraints. *IEEE Trans. Automat. Control* **45** (8) (2000), 1420-1436.

[36] A.D. Lewis, R.M. Murray: Variational principles for constrained systems: theory and experiments. *International Journal of Nonlinear Mechanics* **30** (6) (1995), 793-815.

[37] W. Liu, H.J. Sussmann: Abnormal sub-Riemannian minimizers. In *Differential equations, dynamical systems, and control science*, Lecture Notes in Pure and Appl. Math., vol 152, Dekker, New York, 1994, pp. 705-716.

[38] C. López, E. Martínez: Sub-Finslerian metric associated to an optimal control system. *SIAM J. Control Optim.* **39** (3) (2000), 798-811.

[39] M.J.P. Magill: *On a General Economic Theory of Motion*. Springer-Verlag, Berlin, 1970.

[40] C.-M. Marle: Reduction of constrained mechanical systems and stability of relative equilibria. *Commun. Math. Phys.* **174** (1995), 295-318.

[41] C.-M. Marle: Various approaches to conservative and nonconservative nonholonomic systems. *Rep. Math. Phys.* **42** (1/2) (1998), 211-229.

[42] S. Martínez, J. Cortés, M. de León: The geometrical theory of constraints applied to the dynamics of vakonomic mechanical systems. The vakonomic bracket. *J. Math. Phys* **41** (2000), 2090-2120.

[43] S. Martínez, J. Cortés, M. de León: Symmetries in vakonomic dynamics. Applications to optimal control. *J. Geom. Phys.* **38** (3-4) (2001), 343-365.

[44] B.M. Maschke, A.J. van der Schaft, P.C. Breedveld: An intrinsic Hamiltonian formulation of the dynamics of LC-circuits. *IEEE Trans. Circuits Systems I Fund. Theory Appl.* **42** (1995), 73-82.

[45] R. Montgomery: Abnormal minimizers. *SIAM J. Control Optim.* **32** (1994), 1605-1620.

[46] L. Moreau, D. Aeyels: A note on the geometric structure of nonlinear inductor-capacitor circuits. Preprint, 2002.

[47] R.M. Murray, S.S. Sastry: Nonholonomic motion planning: steering using sinusoids. *IEEE Trans. Auto. Control* **38** (5) (1993), 700-716.

[48] P.J. Olver: *Applications of Lie groups to differential equations*. Graduate Texts in Mathematics Series, vol. 107, Springer-Verlag, New York, 1986.

[49] J.P. Ostrowski: Optimal controls for kinematic systems on Lie groups. *IFAC World Congress*, Beijing, China, 1999.

[50] J.P. Ostrowski, J.W. Burdick: The geometric mechanics of undulatory robotic locomotion. *Internat. J. Robotics Res.* **17** (7) (1998), 683-702.

[51] Y. Pironneau: Sur les liaisons non holonomes non linéaires, déplacements virtuels à travail nul, conditions de Chetaev. In *Proc. IUTAM-ISIMM Symp. on Modern Develop. in Analytical Mechanics*, *Atti Accad. Sci. Torino Cl. Sci. Fis. Mat. Natur.* **117** (1983), pp. 671-686.

[52] F.P. Ramsey: A mathematical theory of saving. *Economic J.* **38** (1928), 543-559.

[53] P.A. Samuelson: Law of conservation of the capital-output ratio. *Proc. Nat. Acad. Sci. U.S.A.* **67** (1970), 1477-1479.

[54] P.A. Samuelson: Law of conservation of the capital-output ratio in closed von Neumann systems. In *Conservation laws and symmetry*, Kluwer, Boston, 1990, pp. 53-56.

[55] J.M. Sanz-Serna, M. Calvo: *Numerical Hamiltonian problems*. Chapman and Hall, London, 1994.

[56] R. Sato, R.V. Ramachandran (eds.): *Conservation Laws and Symmetry: Applications to Economics and Finance*. Kluwer, Boston, 1990.

[57] D.J. Saunders, F. Cantrijn, W. Sarlet: Regularity aspects and Hamiltonisation of nonholonomic systems, *J. Phys. A: Math. Gen.* **32** (1999), 6869-6890.

[58] R. Skinner, R. Rusk: Generalized Hamiltonian dynamics I. Formulation on $T^*Q \oplus TQ$, *J. Math. Phys.* **24** (11) (1983), 2589-2594.

[59] H.J. Sussmann (ed.): *Nonlinear Controllability and Optimal Control*. Marcel Dekker, New York, 1990.

[60] H.J. Sussmann: Geometry and Optimal Control. In *Mathematical Control Theory*, J. Baillieul, J.C. Willems, eds., Springer Verlag, New York, 1998, pp. 140-198.

[61] I.A. Taimanov: Integrable geodesics flows of non-holonomic metrics. *J. Dynam. Control Systems* **3** (1997), 129-147.

[62] A.M. Vershik, L.D. Faddeev: Differential geometry and Lagrangian mechanics with constraints. *Sov. Phys. Dokl.* **17** (1972), 34-36.

[63] A.M. Vershik, L.D. Faddeev: Lagrangian mechanics in invariant form. *Sel. Math. Sov.* **1** (1981), 339-350.

[64] A.M. Vershik, V.Ya. Gershkovich: "I. Nonholonomic Dynamical Systems, Geometry of Distributions and Variational Problems", *Dynamical Systems*, vol. VII, Springer-Verlag, New York-Heidelberg-Berlin, 1994.

[65] A.P. Veselov: Integrable discrete-time systems and difference operators. *Funkts. Anal. Prilozhen.* **22** (1988), 1-13.

[66] A.P. Veselov: Integrable Lagrangian correspondences and the factorization of matrix polynomials. *Funkts. Anal. Prilozhen.* **25** (1991), 38-49.

[67] J.M. Wendlandt, J.E. Marsden: Mechanical integrators derived from a discrete variational principle. *Physica D* **106** (1997), 223-246.

[68] G. Zampieri: Nonholonomic versus vakonomic dynamics. *J. Differential Equations* **163** (2) (2000), 335-347.