Journal of Dynamics and Differential Equations, Vol. 10, No. 4, 1998

Equivalence of Dynamics for Nonholonomic Systems with Transverse Constraints

Marco Favretti[1]

Received February 5, 1996

This paper is concerned with the dynamics of a mechanical system subject to nonintegrable constraints. In the first part, we prove the equivalence between the classical nonholonomic equations and those derived from the nonholonomic *variational* formulation, proposed by Kozlov in [10–12], for a class of constrained systems with constraints *transverse* to a foliation. This result extends the equivalence between the two formulations, proved for holonomic constraints, to a class of linear nonintegrable ones. In the second part, we derive the nonholonomic variational *reduced* equations for a constrained system with symmetry and constraint transverse to a principal bundle fibration, using a reduction procedure similar to the one developed in [5]. The resulting equations are compared with the nonholonomic reduced ones through mechanical examples.

KEY WORDS: Nonholonomic systems; foliations; reduced equations.

1991 MATHEMATICS SUBJECT CLASSIFICATIONS: 49S05, 70F25, 53C12.

1. INTRODUCTION

The problem of writing the dynamic equations for a mechanical system subject to non holonomic constraints using a variational principle has a long history. It has received new contributions in [10] and subsequent papers. In [10], the author proposes a derivation of the dynamic equations based on an extension of Hamilton' variational Principle. More in detail, the motion for the system is seen as the solution of a constrained variational problem for the Lagrangian functional; using a standard procedure in calculus of variations, one substitutes the constrained variational

[1] Dipartimento di Matematica Pura e Applicata, Università di Padova, via Belzoni 7, 35131 Padova, Italy. E-mail: favretti@math.unipd.it.

1040-7294/98/1000-0511$15.00/0

problem for the Lagrangian L with an unconstrained one for a Lagrangian $\mathscr{L}$ in which the constraints are taken into account by Lagrange multipliers. The nonholonomic variational equations are then the Euler–Lagrange equations for the Lagrangian functional $\mathscr{L}$. This formulation has the advantage that it can handle equally well the linear and nonlinear constraints but, on the other hand, it is troublesome in that, once the dynamic equations are given normal form, the value of the Lagrange multipliers has to be supplied among initial data for the evolution Cauchy Problem for the configuration variables and Lagrange multipliers. This request, which is absent in the classical nonholonomic equations, based on Virtual Work' Principle, amounts to specify the reaction forces of the constraint in the initial phase space configuration. Moreover, while for integrable constraints the two formulations do coincide, for a linear nonintegrable constraint examples are found where the dynamics may or may not be equivalent for a suitable choice of the initial value of the Lagrange multipliers. Both features are explained in [13] using a physical realization procedure of the constraint that leads to the nonholonomic variational equations: the equations are obtained by adding to the kinetic energy tensor an anisotropic inertia tensor term and letting the inertia tend to infinity. As a result, the kinetic energy of the system depends closely on the direction of the motion. On the other hand, classical nonholonomic equations have a physical realization too in terms of anisotropic viscous friction forces and, according to [13], preference between the two models should be accorded upon the nature of the constraint.

In this paper we study the dynamics for systems with linear non integrable constraints using geometric methods. In the first part we introduce a class of constrained systems in which the linear constraint is complementary to (the tangent space of) a foliation of the configuration manifold. Next, we consider the special case in which the configuration manifold has a principal bundle structure and the linear constraint provides a principal connection on it and we give the conditions for the equivalence between the two formulations in this framework (Theorem 3.1).

For systems with symmetry, a better understanding of the dynamics is obtained if one passes to the *reduced* equations of motion, in which the group variables are absent. Roughly speaking, this is physically equivalent to giving the description of the motion using a rotating frame instead of an inertial one. Therefore, in the second part, we derive the reduced nonholonomic variational equations using a procedure similar to the one developed in [5] and we compare the two formulations through mechanical examples.

Now we recall briefly the two sets of dynamic equations that we will discuss in the subsequent sections. For a more detailed derivation of these see, e.g., [7]. Let $(M, L, \mathscr{A})$ be respectively the configuration manifold, the Lagrange function and the linear non integrable constraint distribution $\mathscr{A} \subset TM$ of a constrained mechanical system. Locally, we describe the linear constraint as the null space of $\dim M - \dim \mathscr{A}$ linearly independent one-form ω^α, called characteristic forms of the distribution $\mathscr{A}$. Therefore, locally

$$\mathscr{A} := \{(z, \dot{z}) \in TM : \omega^\alpha(z)\, \dot{z} = 0, \alpha = 1,..., \dim M - \dim \mathscr{A}\}$$

The classical *nonholonomic* equations are (summation over repeated index is understood)

$$[L] = \lambda_\alpha \omega^\alpha, \qquad \omega^\alpha(z)\, \dot{z} = 0 \tag{1}$$

where $[L] = (d/dt)(\partial L/\partial \dot{z}) - \partial L/\partial z$ are the Lagrange brackets of L and the parameters λ are the Lagrange multipliers. For the *variational nonholonomic* equations, we first form the unconstrained Lagrangian in the (z, λ)-variables

$$\mathscr{L}(z, \dot{z}, \lambda) = L(z, \dot{z}) - \lambda_\alpha \omega^\alpha(z)\, \dot{z} \tag{2}$$

and we write the related Euler–Lagrange equations $[\mathscr{L}] = 0$ as

$$[L] = \dot{\lambda}_\alpha \omega^\alpha + \lambda_\alpha \, d\omega^\alpha(\dot{z}, \cdot), \qquad \omega^\alpha(z)\, \dot{z} = 0 \tag{3}$$

These are commonly referred to as *vakonomic* equations (equations of variational axiomatic kind). In the next section we introduce some geometrical tools that will be used in the description of the reaction forces of the constraint.

2. REVIEW OF CONNECTIONS

On a smooth fibration $\pi: M \to N$, where M, N are smooth manifolds, the set $VM = \ker T\pi$ of the vectors that project onto the null space of TN is an integrable subbundle of TM. An *Ehresmann connection* on $\pi: M \to N$ is the assignment of a distribution HM transversal to VM, $HM \oplus VM = TM$ and the elements of HM are the horizontal vectors. Since $T\pi$ restricted to HM is an isomorphism, it has a fiberwise defined inverse, the horizontal lift: hor: $T_{\pi(z)}N \to T_z M$, $\text{hor}(X) \in H_z M$.

Let $X = X^h + X^v$ be the splitting of a vector in $T_z M$ into its horizontal and vertical component. The projection on VM with respect to the horizontal subspace defines the vector-valued connection one-form

$$\omega: TM \to VM, \qquad \omega(z)(X) = X^v \tag{4}$$

whose kernel is the horizontal distribution. The assignment of an horizontal distribution, of an horizontal lift operator or of a connection one-form are equivalent ways to define a connection on $\pi: M \to N$.

The *curvature* of the connection is the VM-valued two-form obtained by restricting the exterior derivative of ω to the horizontal distribution:

$$\Omega(X, Y) = d\omega(X^h, Y^h) \tag{5}$$

If we extend the vectors X, Y to vector fields $X, Y \in \Gamma(M)$, and we use Cartan's formula $d\omega(X, Y) = X\omega(Y) - Y\omega(X) - \omega([X, Y])$, we get the equivalent expression for the curvature $\Omega(X, Y) = -\omega([X^h, Y^h])$ that shows that the curvature exactly measures the failure of the horizontal distribution to be integrable.

Next we give the local expressions of connection and curvature in a fibered chart. Notice that since every foliation is locally a fibration, the following relations hold locally for a foliation. Let $z = (x, y)$ be a fibered chart on $U \subset M$, $\pi(x, y) = y$. Then the vertical space is

$$V_z U = \ker T_z \pi = \operatorname{span}\left\{\frac{\partial}{\partial x^\alpha}\right\}, \qquad \alpha = 1,..., \dim M - \dim N \tag{6}$$

and the connection one-form ω is

$$\omega = \omega^\alpha \otimes \frac{\partial}{\partial x^\alpha}, \qquad \omega^\alpha = dx^\alpha + A^\alpha_l(z)\, dy^l \tag{7}$$

The $A^\alpha_l(z)$ are the connection's coefficients. The connection's curvature can be expressed as: $X, Y \in \Gamma(U)$

$$\Omega(X, Y) = \Omega^\alpha(X, Y) \otimes \frac{\partial}{\partial x^\alpha} = (\Omega^\alpha_{lk}\, dy^l \otimes dy^k) \otimes \frac{\partial}{\partial x^\alpha} \tag{8}$$

where

$$\Omega^\alpha_{kl} = A^\alpha_{l,k} - A^\alpha_{k,l} - A^\alpha_{l,\beta} A^\beta_k + A^\alpha_{k,\beta} A^\beta_l \tag{9}$$

Now we particularize the above notions to the important class of *principal* fiber bundles, where the fibration is the one defined by the set of orbits of a smooth group action. Suppose that a Lie group G acts freely and properly on the left on M and that, to every $z \in M$, $z \mapsto Gz$ is an immersion, so that $\pi: M \to M/G$ is a principal bundle. On a principal bundle, the group action induces an Ad-equivariant isomorphism σ: $\mathcal{G} \to \Gamma(M, VM)$, $\sigma(\xi) := \xi_M(z)$, where ξ_M is the infinitesimal vector field associated to $\xi \in \mathcal{G}$. A *principal* connection on the bundle is the assignment of a Ehresmann connection compatible with the group action, i.e., satisfying

$$T\Phi_g(H_z M) = H_{gz} M, \qquad \forall z \in M, \quad \forall g \in G \tag{10}$$

Moreover, the related connection one form $\tilde{\omega}$ is G-invariant and it defines an equivariant $\mathscr{G}$-valued connection one form by setting $\omega = \sigma^{-1} \circ \tilde{\omega}$. In a principal bundle, the curvature, defined as above, is Ad-equivariant and, unlike the Ehresmann case, it satisfies the *structure equation*

$$\Omega(X, Y) = \mathrm{d}\omega(X, Y) - [\omega(X), \omega(Y)] \tag{11}$$

Now consider a local trivialization $z = (y, g)$ of the bundle, where g are the coordinates on the fiber isomorphic to G. By equivariance of ω, the local expression of the connection become

$$\omega(y, g)(\dot{y}, \dot{g}) = \mathrm{Ad}_g(\xi + A(y)\,\dot{y}) \tag{12}$$

where $\xi = g^{-1}\dot{g} \in T_e G = \mathscr{G}$ is the left translation to the origin of $\dot{g} \in T_g G$ and A^α_I are the connection's coefficients with respect to a chosen basis $\{e_\alpha\}$ of $\mathscr{G}$. It is important to notice that now the connection's coefficients are represented by functions constant on the fibers.

3. SYSTEMS WITH TRANSVERSE CONSTRAINTS

The aim of this section is to introduce a certain class of constrained systems with non integrable constraints *transversal* to a foliation of the configuration manifold and to discuss the related dynamics. We first split in two separate set the dynamic equations by projecting the vakonomic equations on the tangent space to the leaves and on the constraint distribution respectively. This gives a clearer picture of the geometrical structure of the reaction forces of the constraint, expressed by the right hand side of the equations, in terms of curvature of a connection. Moreover, the evolution of the Lagrange multipliers is seen to be entirely determined by a subset of the dynamic equations, namely those projected on the vertical subspace—see Proposition 4.1 below. Both results help to answer to the question

> *when the vakonomic solution coincides with the nonholonomic one for a suitable choice of the Lagrange multipliers in the initial configuration.*

This is surely possible, for the system at hand, if the reaction forces vanish along both the nonholonomic and vakonomic motion and an instance of the conditions for this to happen are stated in Theorem 3.2. Moreover, Theorem 3.1 below gives a complete answer to the aforementioned question in case of non vanishing reaction forces. The above statements extend the "equivalence" of the theories from the case of integrable constraints, proved in [7, 14], to a certain class of linear non integrable constraint. We first discuss the more general case of a constraint transverse to a foliation and then we rephrase some of the results for a principal bundle fibration.

Let us consider the constrained system $(M, L, \mathscr{A})$, where $\mathscr{A}$ is a non integrable linear constraint, and suppose that

(H.1) the constraint distribution $\mathscr{A}$ admits an integrable distribution $\mathscr{A}^T$ transverse to $\mathscr{A}$

$$\forall z \in M, \qquad \mathscr{A}_z^T \oplus \mathscr{A}_z = T_z M$$

Since $\dim \mathscr{A}^T = \dim M - \dim \mathscr{A}$, (H.1) is very reasonable for a weakly constrained system and it is trivially satisfied in the limiting case $\dim \mathscr{A} = \dim M - 1$. Then the *foliation* defined by the integrable distribution $\mathscr{A}^T$ is the *fibration* $\pi: U \to U/\mathscr{A}^T$, $\pi(x^\alpha, y^l) = y^l$, $\alpha = 1,..., \dim M - \dim \mathscr{A}$, $l = 1,..., \dim \mathscr{A}$, in a open set $U \subset M$. The vertical space of the fibration is $\mathscr{A}^T$ and the constraint distribution $\mathscr{A}$ defines an Ehresmann connection by hypothesis (H.1). Consequently, the horizontal constraint distribution $\mathscr{A}$ is the kernel of a connection one form $\ker \omega = \mathscr{A}$, whose local expression in U is (7). Therefore, the vakonomic lagrangian $\mathscr{L}$ associated to $(M, L, \mathscr{A})$ is

$$\mathscr{L} = \mathscr{L}(z, \dot{z}, \lambda) = L(x, y, \dot{x}, \dot{y}) - \lambda_\alpha(\dot{x}^\alpha + A_l^\alpha(x, y)\, \dot{y}^l)$$

and, by introducing the vertical and horizontal projectors, (3) become

$$\begin{cases} [L] = \dot{\lambda}_\alpha \omega^\alpha + \lambda_\alpha\, d\omega^\alpha(\dot{z}, v(\cdot)) + \lambda_\alpha \Omega^\alpha(\dot{z}, \cdot) \\ \omega(z)\, \dot{z} = 0 \end{cases} \tag{13}$$

The above formula is interesting because it describes the constraint's reactions in terms of the curvature Ω of the connection defined by the linear constraint $\mathscr{A}$. Notice (see [7]) that for an integrable constraint the curvature is vanishing and the two formulations coincide. The local form of the above vakonomic equations is

$$\begin{aligned} [L]_\gamma &= \dot{\lambda}_\gamma - \lambda_\alpha A_{l,\gamma}^\alpha \dot{y}^l \\ [L]_l &= \dot{\lambda}_\alpha A_l^\alpha + \lambda_\alpha A_{l,\beta}^\alpha \dot{x}^\beta + \lambda_\alpha (A_{l,m}^\alpha - A_{m,l}^\alpha)\, \dot{y}^m \\ [\mathscr{L}]_\lambda &= \dot{x}^\beta + A_m^\beta \dot{y}^m = 0 \end{aligned} \tag{14}$$

The above description of linear constraints has a local character. In [7] we show how to make the local fibration into a principal bundle one by adding suitable hypotheses on the constraint $\mathscr{A}$. This construction is then applied to the mechanical example of the vertical rolling disk. In the following we assume that the configuration manifold is a principal bundle and that the linear constraint defines a principal connection on it. In this way we can profit of globally defined geometrical objects; for instance, the kinematically admissible paths are simply the horizontal lift of paths in the base space. The second aim is to develop a vakonomic version of the reduced equations for a nonholonomically constrained system with symmetry. For the nonholonomic approach, these are displayed in [5].

Let $\phi: G\times M\to M$ be a free and proper group action on the manifold M, so that $\pi: M\to M/G$ is a principal bundle; our purpose now is to derive the vakonomic equations for a mechanical system (M, L) subject to *equivariant affine* constraints. These latter are defined as follows: given a principal connection ω on the bundle and an equivariant map $\mu: M\to\mathscr{G}$, the constraint is the affine sub-bundle of TM

$$\mathscr{A}^{\mu} := \{(z, \dot{z})\in TM : \omega(z)\,\dot{z} = \mu(z)\} \tag{15}$$

We refer to $(M, L, \mathscr{A}^{\mu})$ as a mechanical system with equivariant affine constraints and as a mechanical system with *horizontal* constraints in the case $\mu = 0$. The vakonomic lagrangian associated to $(M, L, \mathscr{A}^{\mu})$ is from (2)

$$\mathscr{L}(z, \dot{z}, \lambda) = L(z, \dot{z}) - \langle \lambda, \omega(z)\,\dot{z} - \mu(z)\rangle \tag{16}$$

where $\langle\ ,\ \rangle$ is the pairing between $\mathscr{G}$ and $\mathscr{G}^*$, and the related vakonomic equations of motion are

$$[L] = \langle \dot{\lambda}, \omega\rangle + \langle \lambda, \mathrm{d}\omega(\dot{z}, \cdot) + \mathrm{d}\mu\rangle, \qquad \omega(z)\,\dot{z} - \mu(z) = 0 \tag{17}$$

As before, we gain a deeper insight of the geometrical structure of the above Eq. (17) by projecting it on the vertical and horizontal subspace of TM respectively. Since $VM = \sigma(\mathscr{G})$, and by using the *structure* Eq. (11), we get from (17)

$$\begin{aligned}
\langle [L], \xi_M\rangle &= \langle \dot{\lambda}, \xi\rangle + \langle \lambda, \mathrm{d}\omega(\dot{z}, \xi_M)\rangle + \langle \lambda, \mathrm{d}\mu(\xi_M)\rangle \\
&= \langle \dot{\lambda}, \xi\rangle + \langle \lambda, \Omega(\dot{z}, \xi_M) + [\omega(\dot{z}), \omega(\xi_M)]\rangle + \langle \lambda, \mathrm{L}_{\xi_M}\mu\rangle \\
&= \langle \dot{\lambda}, \xi\rangle + \langle \lambda, [\omega(\dot{z}), \xi]\rangle + \langle \lambda, ad_{\xi}\mu\rangle \\
&= \langle \dot{\lambda}, \xi\rangle + \langle \lambda, ad_{\omega\dot{z}}\xi\rangle - \langle \lambda, ad_{\mu}\xi\rangle \\
&= \langle \dot{\lambda}, \xi\rangle + \langle ad^*_{\omega\dot{z}}\lambda, \xi\rangle - \langle ad^*_{\mu}\lambda, \xi\rangle
\end{aligned}$$

Now, for an horizontal vector $w \in HM$, one has that $d\omega(\dot{z}, w) = \Omega(\dot{z}, w)$ again by structure equation, so (17) gives

$$\langle [L], w \rangle = \langle \lambda, \Omega(\dot{z}, w) \rangle + \langle \lambda, D\mu(w) \rangle$$

since $D\mu(w) = d\mu(\text{hor } w)$ is the covariant derivative in the adjoint bundle. Collecting the results obtained, we can rewrite the vakonomic equations for $(M, L, \mathscr{A}^{\mu})$ as

$$\begin{cases} \langle [L], \xi_M \rangle = \langle \dot{\lambda}, \xi \rangle \\ \langle [L], w \rangle = \langle \lambda, \Omega(\dot{z}, w) \rangle + \langle \lambda, D\mu(w) \rangle \\ \omega(z)\, \dot{z} - \mu(z) = 0 \end{cases} \tag{18}$$

The nonholonomic equations for $(M, L, \mathscr{A}^{\mu})$ are easily derived from (1)

$$\begin{cases} \langle [L], \xi_M \rangle = \langle \nu, \xi \rangle, & \forall \xi \in \mathscr{G} \\ \langle [L], w \rangle = 0, & \forall w \in HM \\ \omega(z)\, \dot{z} - \mu(z) = 0 \end{cases} \tag{19}$$

The general solution to (18) depends on 2 dim M arbitrary constants. If the matrix

$$B := \langle \omega, FL^{-1} \omega \rangle$$

is nonsingular, then in (19) the multipliers ν can be expressed along the motion as functions of the state of the system and time, $\nu = \nu(t, z, \dot{z})$ (see [7]). Hence, the general solution to (19) depends on $2 \dim M - \dim \mathscr{G}$ arbitrary constants and the system of Eqs. (18) and (19) are nonequivalent in the general case.

Nonetheless, a positive answer can be given to the question hinted at in Section 3 in terms of Theorem 3.1 below. To this, we recall some basic facts about mechanical systems with symmetry. We say that (M, L) is a mechanical system with symmetry if the lagrangian L is invariant for the induced action of G on TM, that is $L(z, \dot{z}) = L(\phi_g z, T_z \phi_g \dot{z})$ to every g in G. Given a G-invariant lagrangian with L *hyperregular*, the Legendre transform $FL\colon TM \to T^*M$ is a vector bundle isomorphism and the *momentum map* associated to L is the Ad*-equivariant map $J\colon TM \to \mathscr{G}^*$ defined as

$$\langle J(z, v), \xi \rangle = \langle FL(z, v), \xi_M(z) \rangle := \left\langle \frac{\partial L}{\partial \dot{z}}(z, v), \xi_M(z) \right\rangle \tag{20}$$

The link between the momentum map and the Lagrange bracket relative to a Lagrangian L is afforded by the important formula below that holds along any path $t \mapsto z(t)$ on M:

$$\langle [L], \xi_M(z) \rangle = -\mathrm{d}L \cdot \dot{\xi}_M + \langle \dot{J}, \xi \rangle, \qquad \forall \xi \in \mathscr{G} \tag{21}$$

The term $\mathrm{d}L \cdot \dot{\xi}_M$ is the *infinitesimal variation* of L, and $\dot{\xi}_M$ represents the vector field on TM naturally induced by ξ_M. If L is a G-invariant Lagrangian, $\mathrm{d}L \cdot \dot{\xi}_M \equiv 0$, and along the motion of the unconstrained system (M, L) the momentum map is constant in time. This last statement is known as *Noether' Theorem.*

***Theorem** 3.1. Let $(M, L, \mathscr{A}^\mu)$ be a mechanical system with symmetry and equivariant affine constraints and suppose that the matrix $B = \langle \omega, FL^{-1}\omega \rangle$ is non singular. The following are equivalent:*

(i) the nonholonomic solution $z(t)$ of (19) satisfies the condition

$$\langle J(z, \dot{z}) - J_0 + \lambda_0, \Omega(\dot{z}, w) + \mathrm{d}\mu(w) \rangle = 0 \qquad \forall w \in HM \tag{22}$$

for some $\lambda_0 \in \mathscr{G}^*$;

(ii) the curve $(z(t), \lambda(t))$, constructed with the same $z(t)$, where

$$\lambda(t) = \lambda_0 + \int \nu(t, z, \dot{z})\, dt = J(z, \dot{z}) - J_0 + \lambda_0 \tag{23}$$

is a solution of the vakonomic system (18).

Proof. (i) → (ii). The curve $(z(t), \lambda(t))$ satisfies $(18)_1$ by its very definition. Moreover, if $\lambda(t)$ is given by (23), then the right hand side of $(18)_2$ vanishes by hypothesis (22). Therefore $(z(t), \lambda(t))$ satisfies Eq. $(18)_2$, being equivalent to $(19)_2$.

(ii) → (i). Suppose that $(z(t), \lambda(t))$ satisfies the vakonomic system of Eqs. (18) whereas its component $z(t)$ satisfies (19). The necessity of condition (22) is straightforward by comparing $(18)_2$ and $(19)_2$. This completes the proof. The equivalence $\dot{J} = \nu(t, z, \dot{z})$ for $z(t)$ is straightforward from $(19)_1$ by using (21). Q.E.D

Remark. A version of the above theorem is proved in [19] for a mechanical system (M, L) subject to the nonlinear constraint $f_\alpha(t, z, \dot{z}) = 0$, and the analogous of condition (22) is

$$\lambda_\alpha(t) \left(\frac{d}{dt} \left(\frac{\partial f_\alpha}{\partial \dot{z}} \right) - \frac{\partial f_\alpha}{\partial z} \right) \delta z = 0 \qquad \text{when} \quad \frac{\partial f_\alpha}{\partial \dot{z}} \delta z = 0$$

Notice that this condition is considerably more difficult to prove because it involves at one time the nonholonomic motion $z(t)$ and the λ-component of the vakonomic motion whereas (22) concerns only the nonholonomic solution. Moreover, the same condition is proved in [19] to be necessary and sufficient to apply to the nonholonomic equations the Hamilton–Jacobi method.

Condition (22) is satisfied in a number of interesting cases.

- For a system with *horizontal* and *integrable* constraints, that is with $\mu = 0$, and vanishing curvature Ω,
- If the momentum map $J(z, \dot{z})$ is constant along the nonholonomic motions, by setting $\lambda_0 = 0$.

In more detail, for a class of mechanical systems with symmetry, the momentum map itself provides a connection, the *mechanical connection* of Kummer–Smale that we recall briefly (see also [15] and the references therein). We consider a mechanical system with symmetry (M, g, G) with g the kinetic energy Riemannian metric and the group G acting by isometries. Given a regular value $\mu \in \mathscr{G}^*$ of the momentum map J, we consider the constraint $J^{-1}(\mu) \subset TM$, the isotropy subgroup of G at μ, $G_\mu := \{g \in G : \mathrm{Ad}_g^* \mu = \mu\}$ and the *locked inertia tensor* $\Pi_\mu(z)\colon \mathscr{G}_\mu \to \mathscr{G}_\mu^*$, defined as $\langle \Pi_\mu(z)\xi, \mu\rangle := g(z)(\xi_M(z), \mu_M(z))$. It is known that

$$\alpha\colon TM \to \mathscr{G}_\mu, \qquad \alpha := \Pi_\mu^{-1} \circ J$$

defines a principal connection on the bundle $\pi\colon M \to M/G_\mu$, called the mechanical connection. Moreover, $\tilde{\mu}\colon M \to \mathscr{G}_\mu$, $\tilde{\mu} := \Pi_\mu^{-1}(z)\,\mu$, is an equivariant section of $\mathscr{G}_\mu$ and the affine constraint $J(z, \dot{z}) = \mu$ can be given the equivalent form

$$\mathscr{A}^\mu = \{(z, \dot{z}) \in TM : \alpha(z)\,\dot{z} = \tilde{\mu} \text{ i.e. } J(z, \dot{z}) = \mu\} \tag{24}$$

The system with equivariant affine constraints $(M, g, \mathscr{A}^\mu, G_\mu)$ is a special instance of the system $(M, L, \mathscr{A})$. Anyway, the above result has a scanty value in that the affine constraint (24) is a collection of integrals of motion, therefore it can hardly be seen as a genuine constraint for the system;

- If $\mu = 0$, we have by the Ambrose–Singer Theorem that, along the nonholonomic motion,

  $$\Omega(\dot{z}, w) \in \mathscr{G}_{z_0} := \{\Omega(y)(u, w) : y \in M(z_0),\, u, w \in H_y M\} \leqslant \mathscr{G}$$

  where $\mathscr{G}_{z_0}$ is the Lie algebra of the holonomy group at z_0 and $M(z_0)$ is the accessibility set of the horizontal constraint at z_0. Therefore, condition (22) is satisfied, e.g., if $J(z, \dot{z}) - J_0 + \lambda_0 \in \mathscr{G}^\circ_{z_0}$, the annihilator of $\mathscr{G}_{z_0}$.

We will apply Theorem 3.1 to mechanical systems in Section 5. The request of G-invariance of the lagrangian can be weakened in some cases.

Indeed, for the system $(M, L, \mathscr{A})$, suppose that the Lagrangian is a Riemannian metric g on M and the distribution $\mathscr{A}^\perp$, *orthogonal* to the constraint $\mathscr{A}$, is *integrable*. We call the resulting system $(M, g, \mathscr{A})$ a system with *orthogonal* constraints. Moreover, we suppose the metric to be *bundle-like* with respect to the foliation defined by $\mathscr{A}^\perp$. A bundle-like metric is the completion to TM of a transverse metric $g_{\mathscr{A}}$, which is the pull back of a metric on the local quotient manifold by π, the projection on the local quotient manifold. Equivalently,bundle-like metrics can be defined as follows.

Definition 3.1 [18]. A Riemannian metric g on M is bundle-like for a foliation $\mathscr{F}$ if the normal plane field to $\mathscr{F}$, $\mathscr{A}$, is totally geodesic, that is each geodesic which is tangent to $\mathscr{A}$ at one point remains tangent for its entire length.

Now we are able to prove the following equivalence theorem

***Theorem** 3.2. Let $(M, g, \mathscr{A})$ be a mechanical system with orthogonal constraint and suppose that the kinetic energy metric is bundle-like for the foliation orthogonal to the constraint. Moreover, let the initial velocity (at $t = 0$) be compatible with the constraint. Then*

(i) the trajectory of the resulting nonholonomic motion, viewed as an unparametrized curve, is a geodesic of (M, g) perpendicular to the leaves;

(ii) furthermore, if we set $\lambda = 0$ at $t = 0$, then the corresponding vakonomic solution of (26) coincides with the nonholonomic solution of (25).

Proof. If we set $2L = g$, the nonholonomic and vakonomic equations for $(M, g, \mathscr{A})$ are respectively

$$[L] = \lambda\omega, \qquad \omega(z)\, \dot{z} = 0 \tag{25}$$

$$[L] = \dot{\lambda}\omega + \lambda\, d\omega(\dot{z}, \cdot), \qquad \omega(z)\, \dot{z} = 0 \tag{26}$$

A geodesic of (M, g) satisfies the Lagrange equation $[L] = 0$ for the unconstrained system. At the same time, if its tangent vector at the initial point belongs to the constraint distribution, it is an integral curve of the constraint distribution by Definition 3.1. Hence it is the unique solution to (25) with $\lambda = 0$. So (i) is proved.

If we substitute the solution $z(t)$ of (25) in the left hand side of (26), we get $[L] \equiv 0$; having fixed $\lambda = 0$ for $t = 0$, it follows then that λ is constantly zero along the motion. Therefore the nonholonomic solution satisfies (26) too. This concludes the proof of (ii) and of the theorem. Q.E.D

4. THE REDUCED EQUATIONS OF MOTION

In this section we derive the *reduced* vakonomic dynamics for the system with constraints transverse to a foliation of Section 3 and for the above system with affine equivariant constraints $(M, L, \mathscr{A}^\mu)$ and we contrast the results with the analogous ones for the nonholonomic case.

There is a well-established procedure, developed in [9, 16] for unconstrained systems and in [5] for a nonholonomically constrained mechanical system that enables one to derive the *reduced equations* of motion. Under suitable conditions, one can "drop" the dynamic equation to the quotient (or shape space) of the configuration manifold with respect to the group action, writing a *reduced* equation that contains only the quotient space variables. Once this is solved, the motion on the fiber is recovered by a standard *reconstruction* procedure which utilizes the constraint equation. For system with abelian symmetry, the motion on the fiber is determined by quadratures (see also [15]).

We begin by considering the system $(M, L, \mathscr{A})$ of Section 3. The expressions of the Lagrangian function and of the Ehresmann connection defined by the constraint on a fibered chart are respectively $L(x, y, \dot{x}, \dot{y})$ and $\omega^\alpha = dx^\alpha + A^\alpha_I(z)\, dy^I$, where $z = (x, y)$ are local coordinates on M adapted to the foliation whose leaves have equation $y^I = \text{const}$. By introducing the constraint $\dot{x} + A(z)\, \dot{y} = 0$ in L we form the *constrained Lagrangian* ([5])

$$L_c = L(x, y, -A(z)\, \dot{y}, \dot{y}) \tag{27}$$

Now consider the following function, defined along the vakonomic motions of $(M, L, \mathscr{A})$,

$$\zeta_\alpha(t) := \lambda_\alpha(t) - \left(\frac{\partial L}{\partial \dot{x}^\alpha}\right)^* (y(t), \dot{y}(t), -A(z(t))\, \dot{y}(t)) \tag{28}$$

where the $(\ )^*$ means that the quantity is evaluated for $\dot{x} = -A(z)\, \dot{y}$ *after* partial derivation. By substituting (27) in the left-hand side of the *vakonomic* dynamic Eq. (14) for $(M, L, \mathscr{A})$ of Section 3, and using (28), we find

$$\begin{cases} [L_c]_\alpha = \dot{\zeta}_\alpha - A^\beta_{l,\alpha} \dot{y}^l \zeta_\beta \\ [L_c]\delta y = \langle \zeta, \Omega(\dot{y}, \delta y) \rangle = \left\langle \lambda_\alpha - \left(\dfrac{\partial L}{\partial \dot{x}^\alpha}\right)^*, \Omega^\alpha(\dot{y}, \delta y) \right\rangle \\ \dot{x} + A(y)\, \dot{y} = 0 \end{cases} \tag{29}$$

These can be considered as the *semi-reduced* equations since the right-hand side of (29) depends on x-variables. The local equivalent of the G-invariance for $(M, L, \mathscr{A})$ is that the Lagrangian L and the connection's coefficients A^α_l are *constant* on the leaves of the transverse foliation, that is $L(y, \dot{y}, \dot{x})$ and $A^\alpha_l(y)$. If these conditions are verified for $(M, L, \mathscr{A})$, then $[L_c]_\alpha \equiv 0$ and the ζ_α are *constant* in time along the vakonomic motions. Equation (29) then represents the vakonomic reduced equations. It is interesting to compare this equation with the *nonholonomic reduced* equations for $(M, L, \mathscr{A})$ in [5]:

$$\begin{cases} [L_c]\, \delta y = -\left\langle \left(\dfrac{\partial L}{\partial \dot{x}^\alpha}\right)^*, \Omega^\alpha(\dot{y}, \delta y) \right\rangle \\ \dot{x} + A(y)\, \dot{y} = 0 \end{cases} \tag{30}$$

By comparing (29) and (30), we realize that the difference is represented by the extra term

$$F^{vak}_{gyr} = \langle \lambda_\alpha, \Omega^\alpha(\dot{y}, \cdot) \rangle$$

representing a gyroscopic force that is nonzero in the typical case. Moreover, by setting $\lambda_\alpha(t_0) = (\partial L/\partial \dot{x}^\alpha)^*(t_0)$, one has $\zeta_\alpha \equiv 0$ along the vakonomic motion, so there is a choice of the Lagrange multipliers in the initial configuration that annihilate the nonholonomic reaction force in the quotient space. We recall that these forces arise as a consequence of the reduction procedure.

Vakonomic Momentum Map

The aim of this paragraph is to provide the technical background for the reduced dynamics equations on a trivial bundle to be derived in the next section.

Recall that in the vakonomic model, the variables $c = (z, \lambda)$ are to be treated on the same level. Therefore, to the system $(M, L, \mathscr{A}^\mu)$, it is natural to introduce the extended configuration space $\tilde{M} = M \times \mathscr{G}^*$ for the variables (z, λ) and to define an extension to $\tilde{M}$ of the group action by

$$\tilde{\phi}_g(z, \lambda) = (\phi_g(z), \mathrm{Ad}^*_{g^{-1}} \lambda)$$

where $\langle \mathrm{Ad}^*_g \xi, X \rangle = \langle \xi, \mathrm{Ad}_g X \rangle$. Notice that Ad^* (and hence $\tilde{\phi}$) is a proper action only in special cases, e.g., if G is a compact group. Nonetheless, the vakonomic unconstrained system $(\tilde{M}, \mathscr{L})$ (where $\mathscr{L}$ is the one in (16)) associated to $(M, L, \mathscr{A}^\mu)$ is a mechanical system with symmetry if L is a symmetric Lagrangian. Even if

$$F\mathscr{L} = \frac{\partial \mathscr{L}}{\partial \dot{c}} = (FL - \langle \lambda, \omega \cdot \rangle, 0)$$

fails to define a vector bundle isomorphism, we introduce the related momentum map $\tilde{J}\colon T\tilde{M} \to \mathscr{G}^*$. A straightforward calculation shows that, since $\xi_{\tilde{M}}(c) = (\xi_M(z), -\mathrm{ad}^*_\xi \lambda)$, the ordinary (see (20)) and vakonomic momentum maps are related by

$$\langle \tilde{J}(c, \dot{c}), \xi \rangle = \langle F\mathscr{L}(c, \dot{c}), \xi_{\tilde{M}}(c) \rangle = \langle J(z, \dot{z}), \xi \rangle - \langle \lambda, \xi \rangle. \tag{31}$$

Notice that, due to the definition of $\tilde{\phi}$, $\tilde{J}$ is an Ad^*-equivariant map that can be defined as the difference between the momentum maps on TM and $\mathscr{G}^*$ respectively, since $J_{\mathscr{G}^*}(\lambda)\xi = \langle \lambda, \xi \rangle$.

The system $(\tilde{M}, \mathscr{L})$ being a system with symmetry, the corresponding momentum map $\tilde{J}$ takes constant value along the (vakonomic) motions. Therefore

Proposition 4.1. *Along the motions of the vakonomic system with symmetry $(\tilde{M}, \mathscr{L})$ associated to $(M, L, \mathscr{A}^\mu)$, the time derivative of the momentum map $\tilde{J}$ is zero*:

$$\langle \dot{\tilde{J}}, \xi \rangle = \langle \dot{J}, \xi \rangle - \langle \dot{\lambda}, \xi \rangle = 0, \qquad \forall \xi \in \mathscr{G} \tag{32}$$

Indeed, Eq. (32) is nothing but Eq. $(18)_1$, relative to vertical variations of $\mathscr{L}$, where the Lagrange bracket $[L]$ of the symmetric Lagrangian L has been replaced by its equivalent expression (21).

A result equivalent to Proposition 4.1 is contained in [2] (Theorem 9, p. 84) and a similar one is proved in [20] (Theorem 4.6) for a mechanical system with symmetry acted on by a gyroscopic force field. This latter is used as a control in feedback form to stabilize the system around a position of relative equilibrium.

Our next aim is to rewrite the content of Proposition 4.1 for a trivial bundle $M = B \times G$. We will use the result in the derivation of the *reduced* vakonomic equations for a system with affine equivariant constraints in the next section. As it is well known, the tangent space TG to a group G is isomorphic to $G \times \mathscr{G}$, e.g., by *left translation* to the origin $\xi = g^{-1}\dot{g} = T_g L_{g^{-1}} \dot{g}$; therefore we can identify the tangent space TM with $TB \times G \times \mathscr{G}$ and, if $z = (y, g)$ are coordinates on $B \times G$, induced coordinates on $TB \times \mathscr{G}$ are $(y, \dot{y}, \xi)$. A Lagrangian L which is invariant for the left G-action has the form $L(y, \dot{y}, g^{-1}\dot{g})$ therefore it induces a *reduced* lagrangian function $l = l(y, \dot{y}, \xi)$ on $TB \times \mathscr{G}$. Now we come to the description of the constraint $\mathscr{A}^\mu$ in $B \times G$. Using the local expression (12) of the connection, which holds globally for a trivial bundle, and writing the equivariant section μ of $M \times \mathscr{G}$ as $\hat{\mu}(y, g) = \mathrm{Ad}_g \mu(y)$, the equivariant constraint $\mathscr{A}^\mu$ in (15) become $\mathrm{Ad}_g(g^{-1}\dot{g} + A(y)\,\dot{y} - \mu(y)) = 0$. For the Lagrange multipliers $\lambda \in \mathscr{G}^*$, we perform the transformation

$$\lambda_b = \mathrm{Ad}_g^* \lambda \tag{33}$$

as much as when one passes from *space* to *body* coordinates in the description of rigid body dynamics. Collecting the above definitions, we define the vakonomic *reduced* Lagrangian associated to $(M, L, \mathscr{A}^\mu)$ to be the following function on $TB \times \mathscr{G} \times \mathscr{G}^*$

$$\mathscr{L}_{loc}(y, \dot{y}, \xi, \lambda_b) = l(y, \dot{y}, \xi) - \langle \lambda_b, \xi + A(y)\,\dot{y} - \mu(y) \rangle \tag{34}$$

To rewrite the vakonomic momentum map in this framework, note that by chain rule one has that $\partial L/\partial \dot{g} = (\partial l/\partial \xi)\, T_g L_{g^{-1}}$, therefore for the ordinary (see (20)) momentum map one has

$$\langle J, \zeta \rangle = \left\langle \frac{\partial L}{\partial \dot{z}}, \zeta_M \right\rangle = \left\langle \left(\frac{\partial L}{\partial \dot{y}}, \frac{\partial L}{\partial \dot{g}} \right), (0, TR_g \zeta) \right\rangle = \left\langle \frac{\partial l}{\partial \xi}, \mathrm{Ad}_{g^{-1}} \zeta \right\rangle$$

and consequently the vakonomic momentum map (31) can be expressed as

$$\langle \tilde{J}, \zeta \rangle = \langle J - \lambda, \zeta \rangle = \left\langle \frac{\partial l}{\partial \xi} - \lambda_b, \mathrm{Ad}_{g^{-1}} \zeta \right\rangle = \left\langle \frac{\partial \mathscr{L}_{loc}}{\partial \xi}, \mathrm{Ad}_{g^{-1}} \zeta \right\rangle \tag{35}$$

If we finally set $\tilde{J}_{loc} = \partial l/\partial \xi - \lambda_b$, we get

$$\tilde{J}_{loc} = \mathrm{Ad}_g^* \tilde{J}$$

The evolution of $\tilde{J}_{loc}$ along the vakonomic motions of $(M, L, \mathscr{A}^\mu)$ is quickly computed using Proposition 4.1 and the fact that if $\zeta(t) = \mathrm{Ad}_g^* \zeta$, its time derivative is $\dot{\zeta}(t) = \mathrm{ad}^*(TL_{g^{-1}}\dot{g})\,\zeta(t)$ (see [1], p. 275). We thus rewrite Proposition 4.1 as

Proposition 4.2. *Along the motions of the unconstrained vakonomic system with symmetry associated to $(M, L, \mathscr{A}^\mu)$, the time derivative of the momentum map $\tilde{J}_{loc}$ satisfies*

$$\frac{d}{dt}\left(\frac{\partial l}{\partial \xi} - \lambda_b\right) = \mathrm{ad}^*_\xi\left(\frac{\partial l}{\partial \xi} - \lambda_b\right) \tag{36}$$

Notice that (36) or equivalently (32) can also be obtained by performing a vertical (i.e., along the fibers) variation of the lagrangian $\mathscr{L}_{loc}$ and $\mathscr{L}$ respectively. This is done in [16] for $l(y, \dot{y}, \xi)$; moreover, by setting $\lambda \equiv 0$, Eq. (36) coincides with the *Euler–Poincaré* equations (see [16]). We will complete the comparison with the reduced Euler–Lagrange equations in the next section.

Vakonomic Reduction on a Trivial Bundle

Now we apply the reduction procedure to the vakonomic unconstrained system $(\tilde{M}, \mathscr{L})$ associated to the system with equivariant affine constraints that we studied in the Section 3. See, e.g., [7] or [14] for a detailed derivation of the vak equations using the variational principle. Since we do not suppose the extended configuration space $\tilde{M} = M \times \mathscr{G}^*$ to be a principal bundle for the extended action $\tilde{\phi}$, we will adopt here a somewhat different approach to the derivation of the reduced equations. Consider the reduced lagrangian function defined in (34) for a trivial bundle $B \times G$

$$\mathscr{L}_{loc}(y, \dot{y}, \xi, \lambda_b) = l(y, \dot{y}, \xi) - \langle \lambda_b, \xi + A(y)\,\dot{y} - \mu(y)\rangle$$

The vakonomic equations for $\mathscr{L}_{loc}$ can be obtained from $\delta \int \mathscr{L}_{loc} = 0$ with respect to variations with fixed endpoints performed in the (y, ξ, λ_b) variables. We have already taken into account for the ξ-variables by deriving Eq. (36) from Eq. (32) relative to vertical variations of $\mathscr{L}$. For an alternative derivation, one can rephrase the argument used in [16]. Now, performing a variation in the λ_b-variables gives back the constraint equation $\xi + A(y)\,\dot{y} - \mu(y) = 0$, while for a variation δy of y with fixed endpoints and $\delta \dot{y} = (d/dt)(\delta y)$ one has

$$\delta \int \mathscr{L}_{loc} = \int -[l]\,\delta y + \langle \lambda_b, \mathrm{d}A(\dot{y}, \delta y) + \mathrm{d}\mu(\delta y)\rangle + \langle \dot{\lambda}_b, A\delta y\rangle \tag{37}$$

Collecting the above results, the reduced vakonomic equations display as

$$\begin{cases} \dfrac{d}{dt}\left(\dfrac{\partial l}{\partial \xi} - \lambda_b\right) = \mathrm{ad}^*_\xi \left(\dfrac{\partial l}{\partial \xi} - \lambda_b\right) \\ \dfrac{d}{dt}\left(\dfrac{\partial l}{\partial \dot{y}}\right) - \dfrac{\partial l}{\partial y} = \langle \lambda_b, \mathrm{d}A(\dot{y}, \cdot) + \mathrm{d}\mu \rangle + \langle \dot{\lambda}_b, A \rangle \\ \xi + A(y)\,\dot{y} - \mu(y) = 0 \end{cases} \tag{38}$$

To make a comparison with the nonholonomic reduced equations in [5], we need to carry on a further step. Following [5], we use the constraint equation $\xi = -A\dot{y} + \mu$ to form the *constrained reduced lagrangian*

$$l_c = \mathscr{L}_{loc}(y, \dot{y}, \lambda_b, -A\dot{y} + \mu) = l(y, \dot{y}, -A\dot{y} + \mu) \tag{39}$$

The Lagrange bracket for l and l_c are related through

$$[l_c] = [l] - \frac{d}{dt}\left(\frac{\partial l}{\partial \xi}\right) A - \left\langle \frac{\partial l}{\partial \xi}, \mathrm{d}A(\dot{y}, \cdot) + \mathrm{d}\mu \right\rangle \tag{40}$$

By substituting $\partial l/\partial \xi$ and $[l]$ as given respectively from the first and the second of (38) into Eq. (40), and using the relations $\Omega_{loc} = \mathrm{d}A - [A, A]$ and $D\mu = d\mu - [A, \mu]$, one gets

$$[l_c] = \left\langle \lambda_b - \frac{\partial l}{\partial \xi}, \Omega_{loc}(\dot{y}, \cdot) + D\mu \right\rangle$$

Therefore we can replace the vakonomic reduced Eqs. (38) with

$$\begin{cases} \dfrac{d}{dt}\left(\dfrac{\partial l}{\partial \xi} - \lambda_b\right) = \mathrm{ad}^*_\xi \left(\dfrac{\partial l}{\partial \xi} - \lambda_b\right) \\ [l_c] = \left\langle \lambda_b - \dfrac{\partial l}{\partial \xi}, \Omega_{loc}(\dot{y}, \cdot) + D\mu \right\rangle \\ \xi + A(y)\,\dot{y} - \mu(y) = 0 \end{cases} \tag{41}$$

The nonholonomic reduced equations for the system $(M, L, \mathscr{A}^\mu)$ are computed in [5] as

$$\begin{cases} [l_c] = -\left\langle \dfrac{\partial l}{\partial \xi}, \Omega_{loc}(\dot{y}, \cdot) + D\mu \right\rangle \\ \xi + A(y)\,\dot{y} - \mu(y) = 0 \end{cases} \tag{42}$$

As it is easily seen, nonholonomic Eqs. (42) are simpler than their vakonomic counterpart. This is due, among other things, to the fact that in the vakonomic theory one enlarges the configuration manifold to the (z, λ)-space to look at the evolution of the system and at the reaction forces simultaneously. However, it is possible to get rid of the ξ-variables in (41) and to write a closed set of dynamic equations for the (y, λ)-variables. Moreover, a further reduction could be operated for the λ-variables only by adding more hypotheses on the group action of G on $\mathscr{G}^*$ so as to introduce the bundle $\tilde{\pi}: M \times \mathscr{G}^* \to M \times \mathscr{G}^*/G$.

To carry on the comparison between the two theories it is useful to "drop" condition (22) in Theorem 3.1 to TB. As it is easily seen, condition (22) for the system $(M, L, \mathscr{A}^\mu)$ can be rewritten as

$$\left\langle \left(\frac{\partial l}{\partial \xi}\right)^* - \left(\frac{\partial l}{\partial \xi}\right)^*\Big|_0 + \lambda_b(0), \Omega_{loc}(\dot{y}, \delta y) + D\mu(\delta y) \right\rangle = 0, \qquad \forall \delta y \in TB \quad (43)$$

where the $(\ )^*$ means that the quantity in brackets is evaluated for $\xi = -A\dot{y} + \mu$. Accordingly, from (23)

$$\lambda_b(t) = \lambda_b(0) - \left(\frac{\partial l}{\partial \xi}\right)^*\Big|_0 + \left(\frac{\partial l}{\partial \xi}\right)^* \quad (44)$$

5. MECHANICAL EXAMPLES

In this section we will study in details the vakonomic dynamic for a couple of mechanical systems by applying the reduction procedure and we will check the equivalence of the nonholonomic and vakonomic theory for the systems at hand using Theorem 3.1.

Two-Wheeled Carriage

The following example is widely treated in literature (see, e.g., [8] or [9]). See also [8] for the derivation of the nonholonomic equations and as a general reference. We consider the two-wheeled carriage depicted in Fig. 1. For simplicity' sake, we suppose the center of mass of the carriage to lie on the middle of the wheels' axis. The typical configuration of the system is defined by the position (x, y) of the center of mass G, by the heading angle θ and by the angles ψ_1 and ψ_2 formed by fixed radii on each wheel with a vertical axis. Since the position of the carriage, regardless to the angles ψ_1 and ψ_2, is defined by an element of $SE(2)$, the non abelian

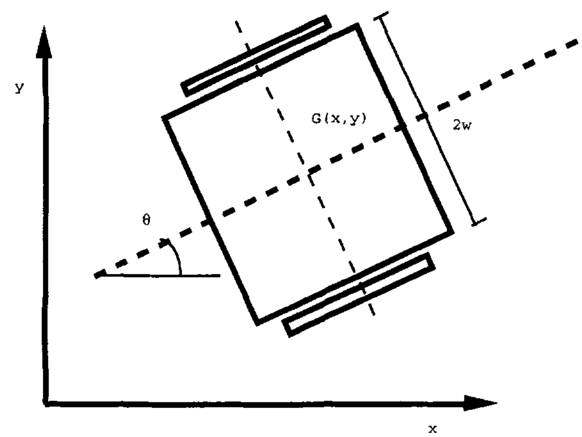

Fig. 1. Two-wheeled carriage.

group of planar rigid motions, the configuration manifold is the trivial bundle $SE(2) \times T^2 \to T^2$, where T^2 is the two-dimensional torus. It is convenient to replace in the following the variables ψ_1 and ψ_2 in the base space with $\phi = (\phi_1, \phi_2)$ where

$$\phi_1 = \tfrac{1}{2}(\psi_1 + \psi_2), \qquad \phi_2 = \tfrac{1}{2}(\psi_1 - \psi_2)$$

Moreover, let m and J be the mass of the system and its inertia and let I, r, $2w$ be the wheels' inertia, radius and mutual distance.

The lagrangian of the system, which reduces to its kinetic energy part, can be written as

$$L = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) + \frac{J}{2}\dot{\theta}^2 + I(\dot{\phi}_1{}^2 + \dot{\phi}_2{}^2)$$

The assumption of pure rolling for each wheel and the constraint of no lateral sliding yields constraints

$$\begin{aligned} \dot{x}\cos\theta + \dot{y}\sin\theta - r\dot{\phi}_1 &= 0 \\ \dot{y}\cos\theta - \dot{x}\sin\theta &= 0 \\ \dot{\theta} - \frac{r}{w}\dot{\phi}_2{}^* &= 0 \end{aligned} \tag{45}$$

Now we recall the group action of $SE(2)$ acting on itself on the left: denote with $g=(p, R)$ the typical element of $SE(2)$ and with $g_1 \circ g_2 = (p_1 + R_1 p_2, R_1 R_2)$ the group multiplication. These are expressed in components by

$$p = \begin{bmatrix} x \\ y \\ \theta \end{bmatrix}, \quad R = \begin{bmatrix} \cos\theta & -\sin\theta & 1 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad g_1 \circ g_2 = \begin{bmatrix} x_1 + x_2 \cos\theta_1 - y_2 \sin\theta_1 \\ y_1 + x_2 \sin\theta_1 + y_2 \cos\theta_1 \\ \theta_1 + \theta_2 \end{bmatrix}$$

The velocity of a rigid motion $\xi \in se(2) \simeq R^3$ is given by

$$\xi = g^{-1}\dot{g}, \qquad \xi = \begin{bmatrix} v_x \\ v_y \\ v_\theta \end{bmatrix} = \begin{bmatrix} \dot{x}\cos\theta + \dot{y}\sin\theta \\ \dot{y}\cos\theta - \dot{x}\sin\theta \\ \dot{\theta} \end{bmatrix}$$

One can check that both the lagrangian and constraints are invariant for the left $SE(2)$-action and that (45) defines a principal connection on the bundle $SE(2) \times T^2 \to T^2$ (see [8]). Indeed, (45) can be rewritten as $\xi + A(\phi)\,\dot{\phi} = 0$, where $A\colon T^2 \to se(2)$ is the local connection one form

$$A = \begin{bmatrix} -r & 0 \\ 0 & 0 \\ 0 & -r/w \end{bmatrix}$$

and ξ is as above. The constraint connection yields a non-vanishing curvature two-form

$$\Omega_{loc} = \mathrm{d}A - [A, A] = 0 - \begin{bmatrix} 0 \\ (r^2/w)(\mathrm{d}\phi_1 \wedge \mathrm{d}\phi_2) \\ 0 \end{bmatrix}$$

Now we apply the reduction procedure developed in the last section. Note that by the $SE(2)$ invariance of the lagrangian, L coincides with the *reduced* lagrangian $l = l(\phi, \dot{\phi}, \xi)$

$$l = \tfrac{1}{2}\mathscr{I}\xi \cdot \xi + I\dot{\phi} \cdot \dot{\phi}$$

where $\mathscr{I} = \mathrm{diag}[m, m, J]$, and dot denote the scalar product. The *constrained reduced* lagrangian $l_c = l(\phi, \dot{\phi}, -A\dot{\phi})$ is

$$l_c = \tfrac{1}{2}\mathscr{I}A\dot{\phi} \cdot A\dot{\phi} + I\dot{\phi} \cdot \dot{\phi} = \tfrac{1}{2}\mathscr{K}\dot{\phi} \cdot \dot{\phi}$$

where $\mathscr{K} = \mathrm{diag}[2I + mr^2, 2I + (r^2/w^2)\,J]$.

As a first step, we compute the nonholonomic reduced equations for the carriage system from (42) with $\mu = 0$. Since $\partial l/\partial \xi$ computed for $\xi = -A\dot{\phi}$ is

$$\left(\frac{\partial l}{\partial \xi}\right)^* = -\mathscr{I} A\dot{\phi} = \begin{pmatrix} mr\dot{\phi}_1 \\ 0 \\ (r/w)\, J\dot{\phi}_2 \end{pmatrix}$$

one has

$$\left\langle \left(\frac{\partial l}{\partial \xi}\right)^*, \Omega_{loc}(\dot{y}, \delta y) \right\rangle \equiv 0$$

Therefore *condition* (43)—see Theorem 3.1—*for the equivalence between the two formulations is satisfied by setting* $\lambda_b(0) = (\partial l/\partial \xi)^*|_0 = -\mathscr{I} A\dot{\phi}(0)$. The nonholonomic reduced Eqs. (42) are simply

$$\mathscr{K}\ddot{\phi} = 0, \qquad \text{i.e.,} \quad \begin{cases} (2I + mr^2)\, \ddot{\phi}_1 = 0 \\ (2I + (r^2/w^2)\, J)\, \ddot{\phi}_2 = 0 \end{cases} \tag{46}$$

Thus in the nonholonomic case, the motion in the base space is given by

$$\phi_1(t) = \dot{\phi}_1(0)t + \phi_1(0), \qquad \phi_2(t) = \dot{\phi}_2(0)t + \phi_2(0) \tag{47}$$

The motion for the whole system is recovered from the above trajectory in the base space by using the constraint Eq. (45) and it yields

$$\begin{aligned} \dot{x}(t) &= r\dot{\phi}_1(0) \cos \theta(t) \\ \dot{y}(t) &= r\dot{\phi}_1(0) \sin \theta(t) \\ \theta(t) &= \theta(0) + \frac{r}{w}\, \dot{\phi}_2(0)\, t \end{aligned} \tag{48}$$

Therefore the center of mass of the system moves along a straight line if $\dot{\phi}_2(0) = \dot{\psi}_1(0) - \dot{\psi}_2(0) = 0$, and it performs a circular orbit if $\dot{\phi}_2(0) \neq 0$ with radius

$$R = \frac{v}{\omega} = \frac{(\dot{x}^2 + \dot{y}^2)^{1/2}}{(r/w)\, \dot{\phi}_2(0)} = w\, \frac{\dot{\phi}_1(0)}{\dot{\phi}_2(0)}$$

Now we compute the vakonomic reduced equations from (41) with $\mu = 0$. These latter read

$$\left\{\begin{array}{l} (2I+mr^2)\,\ddot{\phi}_1=\lambda_{b_2}\,\frac{r^2}{w}\,\dot{\phi}_2 \\ \left(2I+\frac{r^2}{w^2}\,J\right)\ddot{\phi}_2=\lambda_{b_2}\,\frac{r^2}{w}\,\dot{\phi}_1 \\ \dot{\lambda}_{b_1}=mr\ddot{\phi}_1-\frac{r}{w}\,\dot{\phi}_2\lambda_{b_2} \\ \dot{\lambda}_{b_2}=(\lambda_{b_1}-mr\dot{\phi}_1)\,\frac{r}{w}\,\dot{\phi}_2 \\ \dot{\lambda}_{b_3}=\frac{Jr}{w}\,\ddot{\phi}_2+r\dot{\phi}_1\lambda_{b_2} \end{array}\right. \tag{49}$$

According to Theorem 3.1, the nonholonomic solution (47) determines a solution of the vakonomic reduced Eqs. (49) by setting

$$(\phi(t),\lambda_b(t))=\left(\phi(t),\left(\frac{\partial l}{\partial \xi}\right)^*\right)=(\phi(t),\ -\mathscr{I}A\dot{\phi}(t))$$

Next we look at the constraint reaction forces in the nonholonomic theory. Nonholonomic reaction forces are expressed by the right-hand side of the nonholonomic equations [see (1)].

$$[L]=\lambda_\alpha\omega^\alpha, \qquad \omega^\alpha(z)\,\dot{z}=0$$

In this case the above equations read

$$\left\{\begin{array}{l} m\ddot{x}=\lambda_1\cos\theta-\lambda_2\sin\theta \\ m\ddot{y}=\lambda_1\sin\theta+\lambda_2\cos\theta \\ J\ddot{\theta}=\lambda_3 \\ 2I\ddot{\phi}_1=-r\lambda_1 \\ 2I\ddot{\phi}_2=-\frac{r}{w}\,\lambda_3 \end{array}\right. \tag{50}$$

By substituting the trajectory (47) and (48) in (50), one gets

$$\lambda_1\equiv 0, \qquad \lambda_3\equiv 0, \qquad \lambda_2=\frac{mr^2}{w}\,\dot{\phi}_1(0)\,\dot{\phi}_2(0)=m\omega^2R$$

so that the reaction forces are vanishing ($\lambda_i \equiv 0$, $i = 1,..., 3$) only for rectilinear motions ($\dot{\phi}_2(0) = 0$) or "point" motions ($\dot{\phi}_1(0) = 0$), i.e., rotations about a fixed center of mass position.

Sphere Rolling on a Rotating Table

The following example is dealt with in, e.g., [17] or [6]. See also [14] for a detailed analysis of the dynamics from both a theoretical and experimental point of view.

We consider an homogeneous sphere having mass m and inertia $I = \frac{2}{5}ma^2$ and rolling without sliding on a plane that can eventually rotate with constant angular velocity Ω about a vertical axis. The typical configuration is defined by the contact point between the sphere and the plane and by the orientation of a frame joint to the sphere with respect to an inertial frame. Thus the configuration manifold is the trivial bundle $R^2 \times SO(3) \to R^2$ and the lagrangian function, which reduces to its kinetic part, is invariant for the *right* $SO(3)$ action. We make use of the Lie algebra isomorphism $(so(3), [\ ,\ ]) \to (R^3, \times)$ and we take $(r, \dot{r}, \omega)$, where $\omega = \dot{A}A^{-1}$ is the *right* translation to the identity of $\dot{A} \in T_A SO(3)$, as coordinates of the reduced space $TR^2 \times so(3)$ (see Section 3). The reduced lagrangian is

$$l(r, \dot{r}, \omega) = \tfrac{1}{2}m\dot{r} \cdot \dot{r} + \tfrac{1}{2}I\omega \cdot \omega$$

We recall that using the aforementioned Lie algebra isomorphism, the expressions for the adjoint and coadjoint actions on R^3 are

$$\mathrm{ad}_a b = a \times b, \qquad \mathrm{ad}_a^* b = b \times a$$

and the pairing between $so(3)$ and its dual space is simply the R^3 scalar product, which defines a bi-invariant metric on $SO(3)$.

The pure rolling constraint coupled to conservation of the angular momentum about a vertical axis (which holds in the nonholonomic theory in absence of external torque) define a principal connection on the trivial bundle, whose local expression of the form $\dot{g}g^{-1} + A(y)\,\dot{y} = 0$ for a *right* action is

$$\omega + A\dot{r} = 0, \qquad \text{i.e.,} \quad \begin{cases} \omega_x + a^{-1}\dot{y} = 0, \\ \omega_y - a^{-1}\dot{x} = 0, \\ \omega_z = 0. \end{cases} \qquad A = \begin{pmatrix} 0 & 1/a \\ -1/a & 0 \\ 0 & 0 \end{pmatrix}$$

We warn that in the vakonomic framework the vertical component of the angular momentum is not conserved a priori and thus the vakonomic lagrangian contains only the rolling constraint. We will take $\omega_z = const.$ as an a priori constraint because we want to profit of the description of the constraint as an affine equivariant constraint of Section 3.

If the platform is rotating with constant angular velocity Ω, then the above linear constraint become the equivariant affine constraint of the form $\dot{g}g^{-1} + A(y)\,\dot{y} = \mu$

$$\omega + A\dot{r} = \mu, \qquad \mu = \begin{pmatrix} (\Omega/a)\,x \\ (\Omega/a)\,y \\ \omega_z(0) \end{pmatrix}$$

Now we can enquire about the equivalence of the two formulations for this example using condition (42)—see Theorem 3.1—that reads

$$\langle I\omega - I\omega_0 + \lambda_b(0), \mathrm{d}\mu + \omega \times A \rangle = 0$$

It is easy to see that this condition is not satisfied in the general case and that *it holds only if* $\mathrm{d}\mu = 0$ (e.g., if $\Omega = 0$ that is the platform does not rotate) *by setting* $\lambda_b(0) = I\omega_0$. The nonholonomic reduced equations for the ball system can be computed from (42) or directly using Newton's equations as in [14, 6] to give

$$\ddot{x} = -\frac{I\Omega}{I + ma^2}\,\dot{y}$$

$$\ddot{y} = \frac{I\Omega}{I + ma^2}\,\dot{x}$$

According to the above dynamic equations, the contact point between the ball and the plate performs a circular orbit whose radius depends linearly on the plate' angular velocity, Ω.

The reduced vakonomic lagrangian (34) is given in this case by

$$\mathscr{L}_{loc} = l(r, \dot{r}, \omega) - \lambda_b \cdot (\omega + A\dot{r} - \mu)$$

and the reduced vakonomic equations can be computed from (41) as

$$\begin{aligned} a\dot{\lambda}_{b1} &= I(\Omega\dot{x} - \ddot{y}) - \lambda_{b3}\dot{x} + a\lambda_{b2}\omega_z(0) - \Omega\lambda_{b3}\,y \\ a\dot{\lambda}_{b2} &= I(\Omega\dot{y} + \ddot{x}) - \lambda_{b3}\,\dot{y} - a\lambda_{b1}\omega_z(0) + \Omega\lambda_{b3}x \\ a\dot{\lambda}_{b3} &= \lambda_{b1}\dot{x} + \lambda_{b2}\,\dot{y} + \Omega(\lambda_{b1}\,y - \lambda_{b2}x) \\ (ma^2 + I)\,\ddot{x} &= -I\Omega\dot{y} + a\lambda_{b1}\omega_z(0) - \lambda_{b3}(\Omega x - \dot{y}) \\ (ma^2 + I)\,\ddot{y} &= I\Omega\dot{x} + a\lambda_{b2}\omega_z(0) - \lambda_{b3}(\Omega y + \dot{x}) \end{aligned}$$

As it can be easily seen from above, if the platform does not rotate ($\Omega = 0$), the solution to the nonholonomic equations $(\ddot{x}, \ddot{y}) = 0$, provides a solution to the above vakonomic equations by setting $\lambda_b(t) = I\omega(t)$.

We have shown that in the ball example the nonholonomic motion cannot be obtained from the vakonomic scheme by a particular choice of the initial conditions $\lambda(0)$ on the multipliers λ. Moreover, vakonomic trajectories are highly sensitive to the choice of $\lambda(0)$, as it is shown by a numerical treatment of the equations in [14]; therefore, the vakonomic approach is unreliable in this example.

ACKNOWLEDGMENTS

The author wishes to thank Prof. F. Cardin for introducing him to the subject and for valuable discussions on this work and Dr. Barbara Pivetta for her kind help with translations from Russian.

REFERENCES

1. Abraham, R., and Marsden, J. (1978). *Foundations of Mechanics*, 2nd ed., Addison Wesley, Reading, MA.
2. Arnold, V. I., Kozlov, V. V., and Nejshtadt, A. I. (1988). Mathematical aspects of classical and celestial mechanics. In *Dynamic Systems III, Encyclopaedia of Mathematical Sciences*, Springer-Verlag, Berlin/Heidelberg/New York.
3. Bloch, A., Krishnaprasad, P. S., Marsden, J. E., and Ratiu, T. S. (1996). The Euler Poincaré equations and double bracket dissipation. *Comm. Math. Phys.* **175**, 1, 1 42.
4. Bloch, A., and Crouch, P. (1993). *Nonholonomic and Vakonomic Control Systems on Riemannian Manifolds*, Field Inst. Comm. 1.
5. Bloch, A., Krishnaprasad, P. S., Marsden, J. E., and Murray, R. (1996). Nonholonomic mechanical systems with symmetry. *Arch. Rational Mech. Anal.* **136**, 1, 21–99.
6. Burns, J. A. (1981). Ball rolling on a turntable: Analog for charged particle dynamics. *Am. J. Phys.* **49**, 1.
7. Cardin, F., and Favretti, M. (1996). On nonholonomic and vakonomic dynamics of mechanical systems with nonintegrable constraints. *J. Geom. Phys.* **18**, 295 325.
8. Kelly, S., and Murray, R. (1995). Geometric phases and robotic locomotion. *J. Robot. Syst.* **12**, (6), 417–431.
9. Koiller, J. (1992). Reduction of some classical nonholonomic systems with symmetry. *Arc. Rat. Mech. Anal.* **118**, 113 148.
10. Kozlov, V. V. (1982). Dynamics of systems with nonintegrable constraints. I. *Mosc. Univ. Mech. Bull.* **37**, (3)-(4), 27 34.
11. Kozlov, V. V. (1982). Dynamics of systems with nonintegrable constraints. II. *Mosc. Univ. Mech. Bull.* **37**, (4), 70 76.
12. Kozlov, V. V. (1983). Dynamics of systems with nonintegrable constraints. III. *Mosc. Univ. Mech. Bull.* **38**, (3), 102–111.
13. Kozlov, V. V. (1992). The problem of realizing constraints in dynamics, *J. Appl. Math. Mech.* **56**, (4), 594–600.

14. Lewis, D. A., and Murray, R. (1995). Variational principles for constrained systems: Theory and experiments. *Internat. J. Non-Linear Mech.* **30**, VI, 793–815.
15. Marsden, J. E., Montgomery, R., and Ratiu, T. (1990). Reduction, symmetry and phases in mechanics. Memoirs of the AMS No. 436, 88, Providence, RI.
16. Marsden, J. E., and Scheurle, J. (1993). The reduced Euler–Lagrange equations. *Fields Inst. Commun.* **1**, 139–164.
17. Neimark, J., and Fufaev, N. (1972). *Dynamics of Nonholonomic Systems*, Transl. Math. Monographs 33, AMS, Providence, RI.
18. Reinhart, B. (1983). *Differential Geometry of Foliations*, Springer-Verlag, Berlin/Heidelberg/New York/Tokio.
19. Rumiantsev, V. V. (1978). On Hamilton's principle for nonholonomic systems, *Prikl. Math. Mech.* **42**, (3), 387–399.
20. Wang, L. S., and Krishnaprasad, P. S. (1992). Gyroscopic control and stabilization, *J. Nonlin. Sci.* **2**, 367–415.