Journal of Geometry and Physics 38 (2001) 343–365

Symmetries in vakonomic dynamics: applications to optimal control

Sonia Martínez, Jorge Cortés*, Manuel de León

Laboratory of Dynamical Systems, Mechanics and Control, Instituto de Matemáticas y Física Fundamental, Consejo Superior de Investigaciones Científicas, Serrano 123, 28006 Madrid, Spain

Received 29 June 2000; received in revised form 29 November 2000

Abstract

Symmetries in vakonomic dynamics are discussed. Appropriate notions are introduced and their relationship with previous work on symmetries of singular Lagrangian systems is shown. Some Noether-type theorems are obtained. The results are applied to a class of general optimal control problems and to kinematic locomotion systems. © 2001 Elsevier Science B.V. All rights reserved.

MSC: 37J15; 49J15; 49K15; 70F25; 70G65

PACS: 03.20.+i

Subj. Class.: Classical mechanics

Keywords: Constrained variational problem; Symmetry; Optimal control

1. Introduction

The existence of symmetries for a dynamical system is of major theoretical and practical importance. In fact, there are plenty of works devoted to develop methods and algorithms to find symmetries for a given problem or to characterize the different types of symmetries it can admit.

A paradigmatic example of the utility of symmetry properties is the Noether's theorem for Hamiltonian systems, which asserts that if one has a certain type of symmetry (called Noether symmetry), then a conservation law for the equations of motion can be directly obtained. The relevance of this result is obvious in the Marsden–Weinstein theorem via the momentum map, where one can reduce the number of degrees of freedom of the system even preserving the symplectic structure.

* Corresponding author.
E-mail addresses: s.martinez@imaff.cfmac.csic.es (S. Martínez), j.cortes@imaff.cfmac.csic.es (J. Cortés), mdeleon@imaff.cfmac.csic.es (M. de León).

PII: S0393-0440(00)00069-3

In geometric mechanics, there has been a considerable effort on the description of the symmetry properties of general Lagrangian and Hamiltonian systems, even with nonholonomic constraints. A classification of infinitesimal symmetries of a given dynamical system was given in [30,31]. Making use of the constraint algorithm [13,14], one can extend many of the results obtained for regular Lagrangian systems to singular Lagrangians [7,19].

In this paper, we focus our attention on symmetries in vakonomic dynamics. A vakonomic system consists of a Lagrangian $L : TQ \to \mathbb{R}$ on the tangent bundle of an n-dimensional configuration manifold Q and a $(2n-m)$-dimensional submanifold of constraints $M \subseteq TQ$. The point is to extremize the functional defined by the Lagrangian among all the curves $c(t)$ on Q which satisfy the constraints, i.e. $\dot{c}(t) \in M$. This constrained variational problem is the natural setting of many optimization problems encountered in economics, control theory, motion of microorganisms, etc. [16,33,34]. A thorough discussion of the relationship between optimal control problems and constrained variational problems can be found in [4]. We would like to stress that the relevant equations describing the dynamic behaviour of systems subject to general constraints are obtained through Lagrange–d'Alembert principle, which is not a truly variational principle. This gives rise to the so-called nonholonomic mechanics, which has been a field of intensive research in the last years. We use the term "vakonomic dynamics" to refer to the use of typical tools from geometric mechanics (such as the ones described below) in the study of optimization problems subject to constraints, which we feel can bring new insights to these problems [4].

There are several geometric descriptions of the vakonomic problem [6,8,12,20,23]. Some of them are based on the fact that, under certain regularity conditions, the vakonomic equations of motion can be obtained as the Euler–Lagrange equations for an extended Lagrangian $\mathcal{L} : T(Q \times \mathbb{R}^m) \to \mathbb{R}$, $\mathcal{L} = L + \lambda^\alpha \phi_\alpha$, where $\phi_\alpha = 0,\ 1 \leq \alpha \leq m$, describe locally the constraint submanifold M and the λ^α are Lagrange multipliers. The Lagrangian $\mathcal{L}$ is obviously singular and the vakonomic system can be studied as a presymplectic system. This will be the point of view adopted in Section 2.

This approach will allow us to adapt the theory of symmetries developed in the general presymplectic setting to vakonomic dynamics. This is done in the first part of the paper, where the notions of vakonomic symmetry, vakonomic infinitesimal symmetry and vakonomic Noether symmetry are introduced (Section 3). These concepts are developed both in the Lagrangian and the Hamiltonian formalisms, where corresponding versions of the Noether's theorem are obtained (Sections 4 and 5). Section 6 is devoted to the case of a Lie group acting by vakonomic symmetries on a vakonomic system.

The developments of the first part are exploited in the second one, where we have dealt with some applications to control theory (Section 7). In particular, we have considered a general optimal control problem consisting of a set of differential equations $\dot{x}^i = f^i(x(t), u(t))$, where x^i are the states and u^a the control variables, and a cost function $L = L(x, u)$ which must be extremized during the motion. In a geometrical setting, this problem is modelled on an affine bundle $C \to B$, where B is the manifold of states. Then, the controls u^a are seen as the fibres of the affine bundle. We can consider a vakonomic problem whose solutions exactly correspond to the solutions of the general optimal control problem and we can make use of the results obtained in the first part generalizing some of the results stated in [10].

We have also treated another application: an optimal control problem for kinematic locomotion systems [16,25,34]. Such systems (which include, among others, robotic devices and microorganisms) are modelled on a principal G-bundle $Q \to B$ endowed with a principal connection. Q is the space of configurations of the system, B the shape space and G the (Lie group) manifold of all possible positions of the device in its environment. In this case, the controls are precisely the shape velocities, which are the variables the device can affect directly. There is also a cost function to minimize, generally associated to the energy "expenditure" of the manoeuvres the device is making. This cost function is accordingly defined on TB from a Riemannian metric on B. The problem is then the following: given two points in Q, find the optimal controls u which steer the system from one point to the other minimizing the cost function while satisfying the constraints provided by the horizontal distribution of the connection. Again, this can be seen as a vakonomic system and we can apply the results for symmetries. This leads us to obtain Wong's equations [24] via a Poisson reduction in a rather straightforward way.

Finally, we have included in Appendix A the basic definitions concerning lifts of vectors and functions as well as symmetries of presymplectic systems.

2. Vakonomic dynamics

Unlike what happens in nonholonomic mechanics [27], in vakonomic mechanics the equations of motion for systems in the presence of nonholonomic constraints are obtained through the application of a variational principle.

The starting point is an n-dimensional configuration manifold Q, a $(2n-m)$-dimensional constraint submanifold M of TQ, locally defined by the independent equations $\phi_\alpha = 0,\ 1 \leq \alpha \leq m$, and a Lagrangian $L : TQ \to \mathbb{R}$. If (q^A) are coordinates in Q with $(q^A, \dot{q}^A)$ the induced coordinates in TQ, then we write $L = L(q^A, \dot{q}^A)$. In general, M will be a subbundle of TQ over Q. For example, in the following sections we will treat the case of a vector subbundle of TQ, $M \equiv D$, defined by a distribution D on Q, or the case of an affine subbundle M modelled on the vector subbundle D of TQ with an additional vector field γ on Q.

Now, according to the theory of the calculus of variations, we extremize the functional

$$\mathcal{J}(c(t)) = \int_0^1 L(c(t), \dot{c}(t))\, \mathrm{d}t$$

defined by L on the set of twice piecewise differentiable curves $c(t)$ joining $c(0) = q_0$ and $c(1) = q_1$, and satisfying the constraints $\dot{c}(t) \in M_{c(t)}\ \forall t$.

We denote the space of such curves by $\tilde{C}(q_0, q_1)$ and will assume that it is a non-empty manifold. A curve in $\tilde{C}(q_0, q_1)$, $c_s : (-\epsilon, \epsilon) \subseteq \mathbb{R} \to \tilde{C}(q_0, q_1)$, is a function such that $c_s(t)$ is a curve in Q joining q_0 and q_1 for all s. A curve c_s will be a variation of c if $c_0(t) = c(t)\ \forall t$. The tangent space of $\tilde{C}(q_0, q_1)$ at $c(t)$ consists of the infinitesimal variations of c, i.e. given $s \mapsto c_s(t),\ c_0 = c$, then

$$X : [0, 1] \to TQ, \qquad X(t) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} c_s(t)$$

is an infinitesimal variation of c. We will assume that there are enough variations and non-trivial infinitesimal variations for each $c \in \tilde{C}(q_0, q_1)$ (see [1,26] for a discussion of the contrary situation or abnormal case).

Now, we set up the equation

$$\mathrm{d}\mathcal{J}_c(X) = 0 \quad \forall X \in T_c\tilde{C}(q_0, q_1),$$

and use the Lagrange multipliers theorem in an infinite-dimensional context to state (see [1,2,22]) that c is an admissible motion if and only if there exist m functions $\lambda^1, \dots, \lambda^m,\ \lambda^\alpha : [0,1] \to \mathbb{R}$ such that

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

and $\phi_\alpha(q^A, \dot{q}^A) = 0$, $1 \le \alpha \le m$. From (1), we deduce that a curve $c = (q^A(t))$ in $\tilde{C}^2(q_0, q_1)$ is a solution of the vakonomic equations if and only if there exist local functions $\lambda^1, \dots, \lambda^m$ on $\mathbb{R}$ such that $\bar{c}(t) = (q^A(t), \lambda^\alpha(t))$ is an extremal for the extended Lagrangian

$$\mathcal{L} : T(Q \times \mathbb{R}^m) \to \mathbb{R}, \quad \mathcal{L} = L + \lambda^\alpha \phi_\alpha,$$

i.e. it satisfies the Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}^A}\right) - \frac{\partial \mathcal{L}}{\partial q^A} = 0, \quad 1 \le A \le n,$$
$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial \mathcal{L}}{\partial \dot{\lambda}^\alpha}\right) - \frac{\partial \mathcal{L}}{\partial \lambda^\alpha} \equiv \phi_\alpha(q^A, \dot{q}^A) = 0, \quad 1 \le \alpha \le m$$

(see [1,22] for details).

From the extended Lagrangian $\mathcal{L}$ we can construct the system $(TP, \omega_\mathcal{L}, \mathrm{d}E_\mathcal{L})$, where $\omega_\mathcal{L} = -\mathrm{d}\theta_\mathcal{L}$ is the Poincaré–Cartan 2-form, $\theta_\mathcal{L} = S^*(\mathrm{d}\mathcal{L})$ the Poincaré–Cartan 1-form, and $S = (\partial/\partial \dot{q}^A) \otimes \mathrm{d}q^A + (\partial/\partial \dot{\lambda}^\alpha) \otimes \mathrm{d}\lambda^\alpha$ the canonical almost tangent structure on TP. $E_\mathcal{L} = \Delta\mathcal{L} - \mathcal{L}$ is the energy associated with $\mathcal{L}$, which is defined using the Liouville vector field $\Delta = \dot{q}^A(\partial/\partial \dot{q}^A) + \dot{\lambda}^\alpha(\partial/\partial \dot{\lambda}^\alpha)$. We will assume that $(TP, \omega_\mathcal{L}, \mathrm{d}E_\mathcal{L})$ is presymplectic, i.e. $\omega_\mathcal{L}$ has constant rank.

Within this geometrical framework, we can pose the equation

$$i_\Gamma \omega_\mathcal{L} = \mathrm{d}E_\mathcal{L}, \tag{2}$$

which codifies the vakonomic equations (1). In [23], this point of view for vakonomic dynamics was developed for a natural Lagrangian L (Lagrangian equal to kinetic minus potential energy) and linear constraints. Γ will be a second-order differential equation (SODE) to be found on TP whose integral curves $(q^A(t), \lambda^\alpha(t))$ are the vakonomic solutions $(q^A(t))$ together with the corresponding Lagrange multipliers $(\lambda^\alpha(t))$.

Eq. (2) will not have in general a global well-defined solution on TP. Applying the Gotay–Nester algorithm [13,14] for presymplectic systems, we generate a sequence of submanifolds as follows (this is valid for general presymplectic systems). First put $P_1 = TP$. Then, consider the set

$$P_2 = \{x \in P_1 | \exists Z_x \in T_x P_1 \text{ solution of (2)}\}.$$

Assume that P_2 is a submanifold of P_1. It may happen that the obtained solutions are not tangent to P_1. Then, we restrict P_2 to the submanifold

$$P_3 = \{x \in P_2 | \exists Z_x \in T_x P_2 \text{ solution of} (i_Z \omega_{\mathcal{L}} = \mathrm{d}E_{\mathcal{L}})_{|P_2}\}.$$

Proceeding further, we construct a sequence

$$\cdots \hookrightarrow P_k \hookrightarrow \cdots \hookrightarrow P_3 \hookrightarrow P_2 \hookrightarrow P_1.$$

Alternatively, the constraint submanifolds can be described by

$$P_k = \{x \in P_1 | \mathrm{d}E_{\mathcal{L}}(x)(v) = 0 \ \ \forall v \in T_x P_{k-1}^{\perp}\},$$

where

$$T_x P_{k-1}^{\perp} = \{v \in T_x P_1 | \omega_{\mathcal{L}}(x)(u, v) = 0 \ \ \forall u \in T_x P_{k-1}\}.$$

We say that P_2 is the secondary constraint submanifold, P_3 the tertiary constraint submanifold, and so on.

In the most favourable case, the algorithm will stabilize at some step k and a final constraint submanifold $P_k = P_f$ will exist where there is a well-defined vector field $\Gamma \in TP_f$ such that

$$(i_\Gamma \omega_{\mathcal{L}} = \mathrm{d}E_{\mathcal{L}})_{|P_f}. \tag{3}$$

This solution is not necessarily unique and we will usually have a set of solutions $\mathfrak{X}^{\omega_{\mathcal{L}}}(P_f)$, where

$$\mathfrak{X}^{\omega_{\mathcal{L}}}(P_f) = \{\bar{\Gamma} \in TP_f | (i_{\bar{\Gamma}} \omega_{\mathcal{L}} = \mathrm{d}E_{\mathcal{L}})_{|P_f}\}.$$

See Appendix A for other notations and basic definitions that will be used along the paper.

Remark 1. In [8], an alternative geometric description of vakonomic dynamics in the extended phase space $T^*Q \times_Q M$ was described. This formulation was used to compare the solutions of vakonomic dynamics with the solutions of nonholonomic mechanics for nonholonomic Lagrangian systems.

3. Symmetries

In this section, we study the general symmetries of a vakonomic system (L, M) on TQ and their relationship with the symmetries of $\mathcal{L}$, an extended Lagrangian of the form $\mathcal{L} = L + \lambda^\alpha \phi_\alpha$, where $\{\phi_\alpha | 1 \leq \alpha \leq m\}$ is a global basis of functions defining the submanifold of constraints M.

We will consider that M is an affine subbundle of TQ modelled on the vector subbundle $D \subseteq TQ$, $\dim M = \dim D = 2n - m$ with an additional vector field $\gamma : Q \to TQ$. We say that a vector X_q is in M_q if and only if $X_q - \gamma_q \in D_q$. In other words, if the annihilator D^0

of D is spanned by $\{\omega_\alpha(q) = \mu_\alpha(q)\,\mathrm{d}q^A | 1 \leq \alpha \leq m\}$ and $\omega_\alpha(\gamma_q) = -h_\alpha(q),\ 1 \leq \alpha \leq m$, then

$$X_q \in M_q \Leftrightarrow \omega_\alpha(X_q - \gamma_q) = \mu_{\alpha A} X^A + h_\alpha = 0, \quad 1 \leq \alpha \leq m,$$

i.e. the constraint functions defining M are

$$\phi_\alpha(q, \dot{q}) = \mu_{\alpha A}(q)\dot{q}^A + h_\alpha(q), \quad 1 \leq \alpha \leq m.$$

In the sequel, $\pi_1 : Q \times \mathbb{R}^m \to Q$ and $\pi_2 : Q \times \mathbb{R}^m \to \mathbb{R}^m$ will denote the projections onto each factor of $Q \times \mathbb{R}^m$.

3.1. Vakonomic symmetries

Definition 2 (Arnold [1]). A vakonomic symmetry for (L, M) will be a diffeomorphism $s : Q \to Q$ such that Ts leaves M and $L_{|M}$ invariant, i.e. $Ts(M) = M$ and $(L \circ Ts)_{|M} = L_{|M}$.

In this way, we assure that the constrained variational problem is preserved by s and so will be its solutions.

The condition $(L \circ Ts)_{|M} = L_{|M}$ is equivalent to say that there exist m local functions $\{\lambda_0^\alpha : TQ \to \mathbb{R} | 1 \leq \alpha \leq m\}$ such that $L \circ Ts - L = \lambda_0^\alpha \phi_\alpha$, while the condition $Ts(M) = M$ means that the transformation $\{\phi_\alpha \circ Ts = \bar{\phi}_\alpha | 1 \leq \alpha \leq m\}$ gives rise to new independent constraint functions defining M.

In fact, if $D \subseteq TQ$ is the distribution modelling M, we have

1. Since $\gamma \in M$, then $T_q s(\gamma_q) \in M_{s(q)}$, or equivalently, $T_q s(\gamma_q) - \gamma_{s(q)} \in D_q$.
2. Let X_q be a vector in D_q. Then, $X_q + \gamma_q \in M_q$ and $Ts(X_q) + Ts(\gamma_q) \in M_{s(q)}$. But again, this means $Ts(X_q) + Ts(\gamma_q) - \gamma_{s(q)} \in D_{s(q)}$. By (1) we deduce that $Ts(X_q) \in D_{s(q)}$.

That is, D is invariant by Ts and so is D^0. Thus, a basis $\{\omega_\alpha\}_{\alpha=1}^m$ of D^0 is transformed into a new one $T^*s(\omega_\alpha) = \bar{\omega}_\alpha$. Then, there exists a non-singular matrix-valued function on Q, $\Lambda_\alpha^\beta(s) : Q \to GL(m, \mathbb{R})$ such that $\bar{\omega}_\alpha = \Lambda_\alpha^\beta(s)\omega_\beta$. In other words, if $\phi_\alpha = \mu_{\alpha A}\dot{q}^A + h_\alpha$, $\bar{\phi}_\alpha = \bar{\mu}_{\alpha A}\dot{q}^A + \bar{h}_\alpha$, $1 \leq \alpha \leq m$, are local expressions for the constraint functions corresponding to these basis, we get,

$$\bar{\omega}_\alpha(Y - \gamma) = \bar{\phi}_\alpha(Y), \qquad \bar{\omega}_\alpha(Y - \gamma) = \Lambda_\alpha^\beta(s)\omega_\beta(Y - \gamma) = \Lambda_\alpha^\beta(s)\phi_\beta(Y)$$

for a given $Y \in TQ$, i.e. $\bar{\phi}_\alpha = \Lambda_\alpha^\beta(s)\phi_\beta$, or equivalently, if $\bar{\Lambda}_\alpha^\beta(s)$ denotes the entries of the inverse matrix of $(\Lambda_\alpha^\beta(s))$, we have $\bar{\Lambda}_\alpha^\beta(s)\bar{\phi}_\beta = \phi_\alpha$.

When $L \circ Ts = L$, we can extend the diffeomorphism s to $P = Q \times \mathbb{R}^m$ as

$$\begin{array}{llll} \bar{s} : & P & \to & P \\ & (q^A, \lambda^\alpha) & \mapsto & (s^A(q), \bar{\Lambda}_\beta^\alpha(s)(q)\lambda^\beta) \end{array}$$

so that $T\bar{s}$ leaves $\mathcal{L}$ invariant. Indeed,

$$\mathcal{L} \circ T\bar{s} = L + \bar{\Lambda}_\beta^\alpha \lambda^\beta \bar{\phi}_\alpha = L + \lambda^\beta \phi_\beta = \mathcal{L}.$$

This systematic procedure allows us to translate all the vakonomic symmetries s into symmetries $\bar{s}$ of the singular Lagrangian $\mathcal{L}$ and vice versa, we can recover them just by projecting $\bar{s}$ to Q.

3.2. Vakonomic infinitesimal symmetries

Definition 3. A vakonomic infinitesimal symmetry (from now on VIS) for (L, M) is a vector field X on Q such that its complete lift $X^c \in \mathfrak{X}(TQ)$ is tangent to M and satisfies $X^c(L)_{|M} = X^c_{|M}(L_{|M}) = 0$.

In other words, X is a VIS if and only if its flow, $\{s_t : Q \to Q\}$, consists of vakonomic symmetries for (L, M).

For simplicity, we will consider those X such that $X^c(L) = 0$. Then, from a VIS $X \in \mathfrak{X}(Q)$, one can obtain an infinitesimal symmetry of $\mathcal{L}$, $\bar{X} \in \mathfrak{X}(P)$. Indeed, since $X^c(L) = 0$ and $X^c(\phi_\alpha)_{|M} = 0\ \ \forall 1 \leq \alpha \leq m$, the flow of X, $\{s_t\}$ verifies for all $-\epsilon < t < \epsilon$,

$$L \circ Ts_t = L, \qquad \phi_\alpha \circ Ts_t = \bar{\phi}_{\alpha t} = \Lambda^\alpha_\beta(t)\phi_\alpha.$$

We can then define the one-parameter group

$$\begin{array}{rccc} \bar{s}_t : & P & \to & P \\ & (q, \lambda) & \mapsto & (s_t(q), \Lambda^\alpha_\beta(-t)(q)\lambda^\beta) \end{array}$$

and take the vector field whose flow is given by $\{\bar{s}_t\}$ (its infinitesimal generator), $\bar{X} \equiv X + Y_{\mathcal{L}}$, where

$$Y_{\mathcal{L}} = \left(\frac{\mathrm{d}}{\mathrm{d}t}_{|t=0} \Lambda^\alpha_\beta(-t)(q)\right)\lambda^\beta \frac{\partial}{\partial \lambda^\alpha}.$$

Since $\mathcal{L} \circ T\bar{s}_t = \mathcal{L}$ for all $-\epsilon < t < \epsilon$, it is immediate that $\bar{X}^c(\mathcal{L}) = 0$.

Conversely, given an infinitesimal symmetry of $\mathcal{L}$, $\bar{X} = X^A(q)(\partial/\partial q^A) + f^\alpha_\beta(q)\lambda^\beta (\partial/\partial\lambda^\alpha)$, we have

$$\bar{X}^c(\mathcal{L}) = \bar{X}^c(L) + \lambda^\alpha(f^\beta_\alpha \phi_\beta + \bar{X}^c(\phi_\alpha)) = 0.$$

Since this is valid for every λ^α, we obtain

$$\bar{X}^c(L) = 0, \qquad \bar{X}^c(\phi_\alpha) = -f^\beta_\alpha \phi_\beta.$$

That is, $\bar{X}$ projects onto a vector field on Q, $X = X^A(q)(\partial/\partial q^A)$, which is a VIS for (L, M). For this reason, we will focus our attention on infinitesimal symmetries of $\mathcal{L}$ given by $\bar{X} = X + \lambda^\beta f^\alpha_\beta(q)(\partial/\partial\lambda^\alpha)$, where (f^α_β) is a matrix-valued function on Q, $(f^\alpha_\beta) : Q \to \mathfrak{gl}(m, \mathbb{R})$. We will call to this type of symmetry a VIS for (L, M) on P.

Definition 4. A vakonomic Noether symmetry (VNS) for (L, M) will be a vector field X on Q such that $X^c_{|M} \in \mathfrak{X}(M)$ and $X^c(L)_{|M} = F^c_{|M}$ for some associated function $F : Q \to \mathbb{R}$.

Observe that, although the flow of a Noether symmetry preserves M, it does not consist of vakonomic symmetries in the sense of Definition 2. Its role will be explained in the next section.

In case $X^c(L) = F^c$ on the whole of TQ, the above-defined extension $\bar{X} = X + Y_{\mathcal{L}}$ gives rise to a Noether symmetry of $\mathcal{L}$, i.e. $\bar{X}^c(\mathcal{L}) = \pi_1^*(F^c) \equiv F^c$.

Conversely, let $\bar{X} = X^A(q)(\partial/\partial q^A) + f_\beta^\alpha(q)\lambda^\beta(\partial/\partial\lambda^\alpha) = X + f_\beta^\alpha(q)\lambda^\beta(\partial/\partial\lambda^\alpha)$ be a Noether symmetry for $\mathcal{L}$, say,

$$\bar{X}^c(\mathcal{L}) = \bar{X}^c(L) + \lambda^\alpha(f_\alpha^\beta\phi_\beta + \bar{X}^c(\phi_\alpha)) = \bar{F}^c \tag{4}$$

for some $\bar{F} : P \to \mathbb{R}$. Since $\partial\bar{F}/\partial\lambda^\alpha = \partial\bar{X}^c(\mathcal{L})/\partial\dot{\lambda}^\alpha = 0$, equating $\lambda^\alpha = 0$, we have

$$\bar{X}^c(L) = \bar{F}^c,$$

and being (4) valid for all λ^α, we also have

$$\bar{X}^c(\phi_\alpha) = -f_\alpha^\beta\phi_\beta.$$

Thus, $\bar{F}$ must be the pullback of a function $F : Q \to \mathbb{R}$, and $\bar{X}$ projects to X, a VNS for (L, M) with associated function F. These type of symmetries $\bar{X}$ will be referred as VNS for (L, M) on P.

Example 5 (Closed von Neumann model). In economics, the variational calculus is an indispensable tool when dealing with typical optimization problems. The following example was taken from [8,32,33]. The n capital goods $K_1, \dots, K_n$ and the respective capital formations $\dot{K}_1, \dots, \dot{K}_n$ can be considered as coordinates $(K_1, \dots, K_n, \dot{K}_1, \dots, \dot{K}_n)$ in $T\mathbb{R}^n$.

Given the Lagrangian $L : T\mathbb{R}^n \to \mathbb{R}$, $L(K_1, \dots, K_n, \dot{K}_1, \dots, \dot{K}_n) = \dot{K}_n$ and the constraint function

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

with $\sum_{i=1}^n \alpha_i = 1$, which defines the submanifold $M = \{\phi \equiv 0\}$ of $T\mathbb{R}^n$, the von Neumann problem consists of maximizing

$$\int_0^T \dot{K}_n\,\mathrm{d}t \quad \text{subject to } \phi \equiv 0$$

for some $T \geq 0$ and appropriate initial conditions.

Alternatively, we can formulate the von Neumann problem in terms of the extended Lagrangian $\mathcal{L}(K_1, \dots, K_n, \lambda, \dots, \dot{K}_n, \dot{\lambda}) = \dot{K}_n + \lambda\phi$.

Let X be a vector field on $\mathbb{R}^n$, $X = \sum_{j=1}^n X_j(K_1, \dots, K_n)\partial/\partial K_j$. Then,

$$X^c(L) = \sum_{i=1}^n \dot{K}_i \frac{\partial X_n}{\partial K_i} = (X_n)^c,$$

$$X^c(\phi) = \sum_{i=1}^n \left\{ \alpha_i K_1^{\alpha_1}\cdots K_{i-1}^{\alpha_{i-1}} X_i K_i^{\alpha_i - 1} K_{i+1}^{\alpha_{i+1}} \cdots K_n^{\alpha_n} \right.$$
$$\left. -(\dot{K}_1^2 + \cdots + \dot{K}_n^2)^{-1/2} \sum_{j=1}^n \dot{K}_i \dot{K}_j \frac{\partial X_i}{\partial K_j} \right\}.$$

If $X = C\sum_{i=1}^{n} K_i(\partial/\partial K_i)$, with C a non-zero constant, we have

$$X^c(\phi) = C\left(\sum_{i=1}^{n}\alpha_i\right) K_1^{\alpha_1}\cdots K_n^{\alpha_n} - C(\dot{K}_1^2 + \cdots + \dot{K}_n^2)^{1/2} = C\phi,$$

or equivalently, $X^c(\phi)_{|M} = 0$. Thus, $X = C\sum_{i=1}^{n} K_i(\partial/\partial K_i)$ is a VNS for (L, M) with associated function $X_n = CK_n$. We have shown above that X gives rise to a Noether symmetry $\bar{X}$ of the extended Lagrangian $\mathcal{L}$. In fact, we have

$$\bar{X} = C\left(\sum_{i=1}^{n} K_i\frac{\partial}{\partial K_i} - \lambda\frac{\partial}{\partial\lambda}\right).$$

3.3. Symmetries given by the action of a Lie group

Finally, let $\Phi : G \times Q \to Q$ be a free and proper left action of a Lie group G on the configuration space Q. Denote by Φ_g the diffeomorphism of Q, $q \mapsto \Phi(g, q)$ for each $g \in G$. The group G will be a group of vakonomic symmetries for (L, M), if each Φ_g is a vakonomic symmetry, i.e. if the lifted action $T\Phi : G \times TQ \to TQ$ satisfies $L_{|M} \circ T\Phi_g = L_{|M}$ and $T\Phi_g(M) = M\, \forall g \in G$.

We can make use of the procedure described before to extend a symmetry from Q to $P = Q \times \mathbb{R}^m$. Given a fixed Lagrangian, $\mathcal{L} = L + \lambda^\alpha\phi_\alpha$, let us assume that $L \circ T\Phi_g = L$ for all $g \in G$. Then we define the new action

$$\begin{array}{cccc} \Psi : & G \times P & \to & P \\ & (g, (q, \lambda)) & \mapsto & (\Phi_g(q), \bar{\Lambda}^\alpha_\beta(g)(q)\lambda^\beta) \end{array}$$

It is easy to check that this is indeed a free action and, when G is compact, one can assure that it is also proper.

4. Constants of the motion

One is commonly interested in studying the symmetry properties of a dynamical problem because this can yield, via e.g. a Noether's theorem,e.g. information about conservation laws or reduction of the number of degrees of freedom. In the following three sections we shall explore this topic. Some of the work developed in [19] for symmetries of singular Lagrangian systems will be helpful in the context we have exposed for vakonomic mechanics. We refer to Appendix A for a review of several definitions of symmetries that will be used in the sequel.

Lemma 6. *Let (N, ω, α) be a presymplectic system and $\phi : N \to N$ a diffeomorphism such that*

$$\phi^*\omega = \omega, \qquad \phi^*\alpha = \alpha.$$

Consider $\cdots N_k \hookrightarrow \cdots \hookrightarrow N_2 \hookrightarrow N_1$ the sequence of constraint submanifolds obtained applying the Gotay–Nester algorithm. Then, ϕ restricts to diffeomorphisms $\phi_k : N_k \to N_k\, \forall k$.

Proof. See [19]. □

Now, we are in a position to prove the following proposition.

Proposition 7 (Noether's theorem). *Assume that the sequence of submanifolds obtained through the application of the Gotay–Nester algorithm stabilizes at some step* $k_f \equiv f$. *Let* $\bar{X} \in \mathfrak{X}(P)$ *be a VNS for* (L, M) *with associated function* $F : P \to \mathbb{R}$. *Then,*

1. $\bar{X}^c_{|P_k} \in \mathfrak{X}(P_k)\, \forall 1 \leq k \leq k_f$ *and* $\bar{X}^c_{|P_f}$ *is a dynamical symmetry of* $\mathfrak{X}^{\omega_\mathcal{L}}(P_f)$.
2. $(F^v - i_{\bar{X}^c}\theta_\mathcal{L})_{|P_f} : P_f \to \mathbb{R}$ *is a constant of the motion for* $\mathfrak{X}^{\omega_\mathcal{L}}(P_f)$.

Proof. Let

$$\bar{X} = X^A(q)\frac{\partial}{\partial q^A} + g^\alpha_\beta(q)\lambda^\beta\frac{\partial}{\partial\lambda^\alpha},$$

be the local expression of $\bar{X}$. Since

$$\bar{X}^c(\mathcal{L}) = X^A\frac{\partial\mathcal{L}}{\partial q^A} + \lambda^\beta g^\alpha_\beta\frac{\partial\mathcal{L}}{\partial\lambda^\alpha} + \dot{q}^B\frac{\partial X^A}{\partial q^B}\frac{\partial\mathcal{L}}{\partial\dot{q}^A} = F^c,$$

and F is the pullback of a function on Q, then

$$\begin{aligned}\bar{X}^c(E_\mathcal{L}) &= \bar{X}^c\left(\frac{\partial\mathcal{L}}{\partial\dot{q}^B}\dot{q}^B\right) - \bar{X}^c(\mathcal{L}) = \bar{X}^c\left(\frac{\partial\mathcal{L}}{\partial\dot{q}^B}\right)\dot{q}^B + \frac{\partial\mathcal{L}}{\partial\dot{q}^B}\frac{\partial X^B}{\partial q^C}\dot{q}^C - F^c \\ &= \left(X^A\frac{\partial^2\mathcal{L}}{\partial\dot{q}^B\partial q^A} + \lambda^\beta g^\alpha_\beta\frac{\partial^2\mathcal{L}}{\partial\dot{q}^B\partial\lambda^\alpha} + \dot{q}^C\frac{\partial X^A}{\partial q^C}\frac{\partial^2\mathcal{L}}{\partial\dot{q}^A\partial\dot{q}^B} + \frac{\partial\mathcal{L}}{\partial\dot{q}^A}\frac{\partial X^A}{\partial q^B}\right)\dot{q}^B - F^c \\ &= \frac{\partial}{\partial\dot{q}^B}(\bar{X}^c(\mathcal{L}))\dot{q}^B - F^c = 0,\end{aligned}$$

and similarly,

$$\begin{aligned}L_{\bar{X}^c}\theta_\mathcal{L} &= L_{\bar{X}^c}\left(\frac{\partial\mathcal{L}}{\partial\dot{q}^B}\mathrm{d}q^B\right) = \bar{X}^c\left(\frac{\partial\mathcal{L}}{\partial\dot{q}^B}\right)\mathrm{d}q^B + \frac{\partial\mathcal{L}}{\partial\dot{q}^A}\frac{\partial X^A}{\partial q^B}\mathrm{d}q^B \\ &= \left(\frac{\partial}{\partial\dot{q}^B}(\bar{X}^c(\mathcal{L}))\right)\mathrm{d}q^B = \mathrm{d}F^v.\end{aligned}$$

In particular, these computations imply

$$i_{\bar{X}^c}\omega_\mathcal{L} = \mathrm{d}(i_{\bar{X}^c}\theta_\mathcal{L}) - L_{\bar{X}^c}\theta_\mathcal{L} = \mathrm{d}(i_{\bar{X}^c}\theta_\mathcal{L} - F^v), \tag{5}$$

and

$$L_{\bar{X}^c}\omega_\mathcal{L} = i_{\bar{X}^c}\,\mathrm{d}\omega_\mathcal{L} + \mathrm{d}i_{\bar{X}^c}\omega_\mathcal{L} = \mathrm{dd}(i_{\bar{X}^c}\theta_\mathcal{L} - F^v) = 0.$$

Therefore, the presymplectic structure of $(P_1, \omega_\mathcal{L}, \mathrm{d}E_\mathcal{L})$ is invariant along the flow of $\bar{X}^c$, $\{T\bar{s}_t : P_1 \to P_1\}$,

$$(T\bar{s}_t)^*\omega_\mathcal{L} = \omega_\mathcal{L}, \qquad (T\bar{s}_t)^*(E_\mathcal{L}) = E_\mathcal{L}.$$

By Lemma 6, the flow $\{T\bar{s}_t\}$ restricts to each P_k and $\bar{X}^c_{|P_k} \in \mathfrak{X}(P_k)\,\forall 1 \leq k \leq k_f$. In particular, $[\bar{X}^c_{|P_f}, \Gamma] \in \mathfrak{X}(P_f)$ for every $\Gamma \in \mathfrak{X}^{\omega_{\mathcal{L}}}(P_f)$, and

$$i_{[\bar{X}^c_{|P_f},\Gamma]}\omega_{\mathcal{L}|P_f} = L_{\bar{X}^c_{|P_f}}(i_\Gamma\omega_{\mathcal{L}}) - i_\Gamma(L_{\bar{X}^c_{|P_f}}\omega_{\mathcal{L}}) = 0.$$

Thus, $\bar{X}^c_{|P_f}$ is a dynamical symmetry of $\mathfrak{X}^{\omega_{\mathcal{L}}}(P_f)$ (see Appendix A).

To prove (2) take $\Gamma \in \mathfrak{X}^{\omega_{\mathcal{L}}}(P_f)$. Using (5), we get

$$\begin{aligned}\Gamma(i_{\bar{X}^c}\theta_{\mathcal{L}} - F^v)_{|P_f} &= \mathrm{d}(i_{\bar{X}^c}\theta_{\mathcal{L}} - F^v)_{|P_f}(\Gamma) = (i_{\bar{X}^c}\omega_{\mathcal{L}})_{|P_f}(\Gamma)\\ &= -(i_\Gamma\omega_{\mathcal{L}})_{|P_f}(\bar{X}^c) = -\mathrm{d}E_{\mathcal{L}}(\bar{X}^c)_{|P_f} = 0.\end{aligned}$$

Therefore, $(i_{\bar{X}^c}\theta_{\mathcal{L}} - F^v)_{|P_f}$ is constant along the integral curves of $\Gamma \in \mathfrak{X}^{\omega_{\mathcal{L}}}(P_f)$. □

Example 8 (Closed von Neumann model, revisited). As a consequence of Proposition 7, we are able to find a constant of the motion for the von Neumann problem in a systematic way. We have that $\bar{X} = C(\sum_{i=1}^n K_i(\partial/\partial K_i) - \lambda(\partial/\partial\lambda))$ is a VNS for (L, M) on $\mathbb{R}^n \times \mathbb{R}$ with Noether function CK_n. Consequently, we obtain the conservation law

$$\begin{aligned}(CK_n)^v - i_{\bar{X}^c}\theta_{\mathcal{L}} &= (CK_n)^v - i_{\bar{X}^c}\left(\frac{\partial\mathcal{L}}{\partial\dot{K}_j}\mathrm{d}K_j\right) = C\left\{K_n - K_n - \sum_{j=1}^n \lambda K_j\frac{\partial\phi}{\partial\dot{K}_j}\right\}\\ &= \frac{C\lambda}{\sqrt{\dot{K}_1^2 + \cdots + \dot{K}_n^2}}\left(\sum_{j=1}^n K_j\dot{K}_j\right)\end{aligned}$$

on the final submanifold of constraints.

5. Relationship with the Hamiltonian formulation and the SODE's problem

If the extended Lagrangian $\mathcal{L}$ is almost regular, then the vakonomic problem admits an equivalent formulation which is Hamiltonian. In this case, as the constraint functions are linear or affine, it can be proven that $\mathcal{L}$ is almost regular if and only if L is almost regular (see [23]).

Let $\mathcal{F}L : TP \to T^*P$ be the Legendre mapping of $\mathcal{L}$. If $(q^A, \lambda^\alpha, p_A, p_\alpha)$ are local coordinates in T^*P, then the Legendre mapping is locally written as

$$\mathcal{F}L(q^A, \lambda^\alpha, \dot{q}^A, \dot{\lambda}^\alpha) = \left(q^A, \lambda^\alpha, \frac{\partial\mathcal{L}}{\partial\dot{q}^A}, \frac{\partial\mathcal{L}}{\partial\dot{\lambda}^\alpha}\right) = \left(q^A, \lambda^\alpha, \frac{\partial\mathcal{L}}{\partial\dot{q}^A}, 0\right).$$

We say that $\mathcal{L}$ is almost regular if $M_1 = \mathcal{F}L(TP)$ is a submanifold of T^*P, $j_1 : M_1 \hookrightarrow T^*P$, and $\mathcal{F}L : TP \to M_1$ is a submersion whose fibres are connected. When this holds true, it can be assured that $E_{\mathcal{L}}$ is constant along the fibres of $\mathcal{F}L$ and a Hamiltonian $h_1 : M_1 \to \mathbb{R}$ can be defined implicitly as $h_1 \circ \mathcal{F}L = E_{\mathcal{L}}$. Taking $\omega_1 = j_1^*(\omega_P)$ the pullback to M_1 of the canonical symplectic form ω_P of T^*P, we obtain a presymplectic system (M_1, ω_1, h_1). The equations of motion are then

$$i_\Upsilon \omega_1 = \mathrm{d}h_1. \tag{6}$$

To solve it we apply the Gotay–Nester algorithm and get the sequence of submanifolds

$$\cdots \hookrightarrow M_k \hookrightarrow \cdots \hookrightarrow M_3 \hookrightarrow M_2 \hookrightarrow M_1.$$

The Gotay–Nester equivalence theorem [14] relates this sequence with the former one $\{P_k\}_{k\geq 1}$ of $(TP, \omega_\mathcal{L}, E_\mathcal{L})$. Denote by P_f, M_f the final submanifolds of constraints (if they exist). Then, the theorem asserts

1. $\mathcal{F}L_{|P_k} \equiv \mathcal{F}L_k : P_k \to M_k$ is a fibration and M_k is diffeomorphic to $P_k/\mathrm{Ker}(\mathcal{F}L_k)\, \forall k$.
2. If the sequence $\{P_k\}_{k\geq 1}$ terminates at step k_f so will $\{M_k\}_{k\geq 1}$ and the solutions of the systems are equivalent in the following sense. Given $\Gamma \in \mathfrak{X}^{\omega_\mathcal{L}}(P_f)$ which is $\mathcal{F}L_f$-projectable, then $T\mathcal{F}L_f(\Gamma) = \Upsilon$ is a solution of
$$(i_\Upsilon \omega_1 = \mathrm{d}h_1)_{|M_f}. \tag{7}$$
On the other hand, if Υ is a solution of (7) and $\Gamma \in \mathfrak{X}(P_f)$ projects by $T\mathcal{F}L_f$ onto Υ, then Γ is a solution of (3).

Thus, solving (7) we will obtain a set $\mathfrak{X}^{\omega_1}(M_f)$ of vector fields such that their integral curves, $(q^A(t), \lambda^\alpha(t), p_A(t), 0)$, give the vakonomic solutions $(q^A(t), \lambda^\alpha(t))$.

Now, we study how the vakonomic symmetries can be seen as symmetries of (M_1, ω_1, h_1).

Proposition 9 (Noether's theorem). *Let $\bar{X} : P \to TP$ a VNS for (L, M) with associated function $F : P \to \mathbb{R}$. Then,*

1. *$\bar{X}^c_{|P_k}$ is $\mathcal{F}L_k$-projectable onto $\bar{X}^{c*}_{|M_k} + ((\partial F/\partial q^B)(\partial/\partial p_B))_{|M_k} \in \mathfrak{X}(M_k)\, \forall k \geq 1$.*
2. *$\bar{X}^{c*}_{|M_f} + ((\partial F/\partial q^B)(\partial/\partial p_B))_{|M_f}$ is a Cartan symmetry for $\mathfrak{X}^{\omega_1}(M_f)$ and $\iota\bar{X}(\theta)_{|M_f} - F^{v*}_{|M_f}$ is a constant of the motion.*

Proof. We extend a result of [19] for infinitesimal symmetries. We consider here the more general case of Noether symmetries.

It is easy to show that $\bar{X}$ is $\mathcal{F}L_1$-projectable and $T\mathcal{F}L_1(\bar{X}^c) = Y \circ \mathcal{F}L_1$ with $Y \in \mathfrak{X}(M_1)$. Let us see what the expression for $T\mathcal{F}L_1(\bar{X}^c)$ is in local coordinates. On one hand, we have

$$\bar{X}^c(\mathcal{L}) = X^A(q)\frac{\partial \mathcal{L}}{\partial q^A} + \lambda^\beta g^\alpha_\beta(q)\frac{\partial \mathcal{L}}{\partial \lambda^\alpha} + \dot{q}^C \frac{\partial X^A}{\partial q^C}\frac{\partial \mathcal{L}}{\partial \dot{q}^A} = \frac{\partial F}{\partial q^B}\dot{q}^B, \tag{8}$$

while

$$T\mathcal{F}L_1(\bar{X}^c) = X^A \frac{\partial}{\partial q^A} + \lambda^\beta g^\alpha_\beta \frac{\partial}{\partial \lambda^\alpha} + \left(X^A \frac{\partial^2 \mathcal{L}}{\partial q^A \partial \dot{q}^B} + \lambda^\beta g^\alpha_\beta \frac{\partial^2 \mathcal{L}}{\partial \lambda^\alpha \partial \dot{q}^B} + \dot{q}^C \frac{\partial X^A}{\partial q^C}\frac{\partial^2 \mathcal{L}}{\partial \dot{q}^A \partial \dot{q}^B} \right) \frac{\partial}{\partial p_B}.$$

Taking the derivative $\partial/\partial\dot{q}^B$ in (8), and substituting into the last expression we get

$$T\mathcal{F}L_1(\bar{X}^c) = X^A\frac{\partial}{\partial q^A} + \lambda^\beta g^\alpha_\beta\frac{\partial}{\partial\lambda^\alpha} + \left(\frac{\partial F}{\partial q^B} - p_A\frac{\partial X^A}{\partial q^B}\right)\frac{\partial}{\partial p_B},$$

which is just $\bar{X}^{c*}_{|M_1} + ((\partial F/\partial q^A)(\partial/\partial p_A))_{|M_1}$. Now, it is clear that

$$T\mathcal{F}L_k(\bar{X}^c_{|P_k}) = \bar{X}^{c*}_{|M_k} + \left(\frac{\partial F}{\partial q^A}\frac{\partial}{\partial p_A}\right)_{|M_k} = Y_{|M_k} \quad \text{with } Y_{|M_k} \in \mathfrak{X}(M_k)\,\forall k \geq 1.$$

Finally, (2) can be proven using similar arguments as in Proposition 7. Firstly, if Z is a vector field on M_1 and U a vector field on TP projecting onto it by $\mathcal{F}L$, then for any $z \in M_1$ and arbitrary $x \in \mathcal{F}L^{-1}(z)$,

$$\begin{aligned} i_{T\mathcal{F}L(\bar{X}^c)}\omega_1(z)(Z_z) &= \omega_1(z)(T\mathcal{F}L(\bar{X}^c_x), T\mathcal{F}L(U_x)) = (\mathcal{F}L^*\omega_1)(x)(\bar{X}^c_x, U_x) \\ &= \omega_{\mathcal{L}}(x)(\bar{X}^c_x, U_x) = \mathrm{d}(i_{\bar{X}^c}\theta_{\mathcal{L}} - F^v)(x)(U_x) \\ &= \mathrm{d}(\mathcal{F}L^*(\iota\bar{X}(\theta) - F^{v*}))(U_x) = \mathrm{d}(\iota\bar{X}(\theta) - F^{v*})(T\mathcal{F}L(U_x)) \\ &= \mathrm{d}(\iota\bar{X}(\theta) - F^{v*})(Z_z). \end{aligned}$$

Secondly, h_1 is invariant by $T\mathcal{F}L(\bar{X}^c)$ due to

$$L_{T\mathcal{F}L(\bar{X}^c)}(h_1) = L_{\bar{X}^c}\mathcal{F}L^*(h_1) = L_{\bar{X}^c}(E_{\mathcal{L}}) = 0.$$

Therefore, $(\bar{X}^{c*} + (\partial F/\partial q^A)(\partial/\partial p_A))_{|M_f}$ is a Cartan symmetry for $\mathfrak{X}^{\omega_1}(M_f)$ with $\iota\bar{X}(\theta) - F^{v*}$ the associated constant of the motion. □

It is possible to find a submanifold $\mathcal{S} \subseteq P_f$ on which there exists a tangent solution $\Gamma_{\mathcal{L}} \in T\mathcal{S}$, satisfying the SODE condition [14]. Let Υ be a vector field on M_f satisfying (7) and $\Gamma \in \mathfrak{X}(P_f)$ a vector field which projects onto Υ. Now define the mapping

$$\begin{array}{llll} \sigma: & M_f & \to & P_f \\ & y & \mapsto & T\tau_P(\Gamma(x)) \end{array}$$

where $\tau_P : TP \to P$ is the canonical projection and $\mathcal{F}L_f(x) = y$. Observe that σ is well defined as it does not depend on the choice of $x \in \mathcal{F}L_f^{-1}(y)$, because Γ is $\mathcal{F}L$-projectable. In fact, σ is a section of $\mathcal{F}L_f$, $\mathcal{F}L_f \circ \sigma = i\mathrm{d}_{|M_f}$ and its image $\sigma(M_f) = \mathcal{S}$ is a submanifold of P_f. The vector field $\Gamma_{\mathcal{L}} \circ \sigma = T\sigma(\Upsilon)$ satisfies

$$(i_{\Gamma_{\mathcal{L}}}\omega_{\mathcal{L}} = \mathrm{d}E_{\mathcal{L}})_{|\mathcal{S}},$$

and the SODE condition, $S(\Gamma_{\mathcal{L}}) = \Delta$.

Locally,

$$\Gamma_{\mathcal{L}} = \dot{q}^A\frac{\partial}{\partial q^A} + \dot{\lambda}^\alpha\frac{\partial}{\partial\lambda^\alpha} + C^A\frac{\partial}{\partial\dot{q}^A} + D^\alpha\frac{\partial}{\partial\dot{\lambda}^\alpha}$$

for certain functions $C^A = C^A(q, \lambda, \dot{q}, \dot{\lambda})$ and $D^\alpha = D^\alpha(q, \lambda, \dot{q}, \dot{\lambda})$.

Notice that $\mathcal{S}$ and M_f are diffeomorphic by σ and $\mathcal{F}L_{f|\mathcal{S}} = \sigma^{-1}$ and that the dynamics on them are equivalent. Then, we have a complete equivalence between symmetries and constants of the motion on both the Lagrangian and the Hamiltonian sides via σ and $\mathcal{F}L_{f|\mathcal{S}}$.

6. Constants of the motion given by the action of a Lie group

We particularize now the results given in Sections 4 and 5 to the case of a Lie group G which acts on Q, $\Phi : G \times Q \to Q$, freely and properly. We will make use of it in the applications that follow.

As we have seen, if $\Phi : G \times Q \to Q$ verifies $L \circ T\Phi_g = L$ and $T\Phi_g(M) = M \, \forall g \in G$, then we can build an action on P, $\Psi : G \times P \to P$ as $\Psi_g(q, \lambda) = (\Phi_g(q), \bar{\Lambda}^\alpha_\beta(g)(q)\lambda^\beta)$ such that its complete lift to TP, $\Psi^T : G \times TP \to TP$ satisfies $\mathcal{L} \circ \Psi^T_g = \mathcal{L} \, \forall g \in G$. This implies that Ψ^T restricts to well-defined actions $\Psi^T_{|P_k} \equiv \Psi^T_k : G \times P_k \to P_k \, \forall k$.

Let ξ be an element in $\mathfrak{g}$, the Lie algebra of G. Denote by ξ_P (respectively ξ_{P_k}, ξ_Q) the vector field generated by the flow $\Psi_{\exp(t\xi)}$ (respectively $(\Psi^T_k)_{\exp(t\xi)}$, $\Phi_{\exp(t\xi)}$). Then, as a consequence of Proposition 7 we have that ξ_P is a VIS for (L, M) on P, ξ_{P_f} is a dynamical symmetry for $\mathfrak{X}^{\omega_\mathcal{L}}(P_f)$ and

$$\begin{array}{lllllll} J_f: & P_f & \to & \mathfrak{g}^* & & & \\ & x & \mapsto & J_f(x): & \mathfrak{g} & \to & \mathbb{R} \\ & & & & \xi & \mapsto & J_f(\xi)(x) = i_{\xi_{P_f}}\theta_\mathcal{L}(x) \end{array}$$

is a momentum map for the presymplectic system $(P_f, \omega_{P_f}, \mathrm{d}E_{\mathcal{L}|P_f})$. We will call it the vakonomic momentum map [1,12]. Therefore, we have that $J_f(\xi) : P_f \to \mathbb{R}$, $x \mapsto J_f(\xi)(x) = J_f(x)(\xi)$ is a constant of the motion.

If $\xi_Q(q) = \xi^A_Q(q)(\partial/\partial q^A)$ and $\xi_P(q, \lambda) = \xi^A_Q(q)(\partial/\partial q^A) + \xi^\alpha_\beta(q)\lambda^\beta(\partial/\partial\lambda^\alpha)$, then, given $x \in P_f$, we have,

$$J_\xi(x) = i_{\xi_{P_k}}\theta_\mathcal{L}(x) = \frac{\partial\mathcal{L}}{\partial\dot{q}^A}(x)\xi^A_Q(x) = \left(\frac{\partial L}{\partial\dot{q}^A} + \lambda^\alpha\frac{\partial\phi_\alpha}{\partial\dot{q}^A}\right)(x)\xi^A_Q(x).$$

Using this momentum map, one can perform a presymplectic reduction as developed in [11].

On the other hand, the action $\Psi : G \times P \to P$ can be lifted to T^*P, $\Psi^{T*} : G \times T^*P \to T^*P$ as follows. Let $\alpha_{(q,\lambda)}$ be a 1-form in $T^*_{(q,\lambda)}P$. Then $\Psi^{T*}_g(\alpha_{(q,\lambda)}) \in T^*_{\Psi_g(q,\lambda)}P$ will be such that

$$\Psi^{T*}_g(\alpha_{(q,\lambda)})(v) = \alpha_{(q,\lambda)}(\Psi^T_{g^{-1}}(v))$$

for every $v \in T_{\Psi_g(q,\lambda)}P$. In coordinates Ψ^{T*}_g reads as

$$\Psi^{T*}_g(q^A, \lambda^\alpha, p_A, p_\alpha) = \left(\Phi^A_g(q), \Lambda^\alpha_\beta(g)(q)\lambda^\beta, p_B\frac{\partial\Phi^B_{g^{-1}}}{\partial q^A} + p_\alpha\lambda^\beta\frac{\partial\Lambda^\alpha_\beta(g^{-1})}{\partial q^A}, p_\beta\Lambda^\beta_\alpha(g^{-1})(q)\right).$$

This action restricts to M_1 and it is the $\mathcal{F}L_1$-projection of Ψ^T. To check this, observe that, since $\mathcal{L} \circ \Psi^T = \mathcal{L} \;\; \forall g \in G$, after a straightforward computation, we have

$$\frac{\partial\mathcal{L}}{\partial\dot{q}^A}(x) = \frac{\partial\mathcal{L}}{\partial\dot{q}^B}(\Psi^T_g(x))\frac{\partial\Phi^B_g}{\partial q^A}(x)$$

for $x \in P_1$. Then,

$$\mathcal{F}L(x) = \mathcal{F}L(q^A, \lambda^\alpha, \dot{q}^A, \dot{\lambda}^\alpha) = \left(q^A, \lambda^\alpha, \frac{\partial \mathcal{L}}{\partial \dot{q}^B}(\Psi_g^T(x)) \frac{\partial \Phi_g^B}{\partial q^A}(x), 0\right),$$

$$\Psi_g^{T*}(\mathcal{F}L(x)) = \left(\Phi_g^B(q), \Lambda_\beta^\alpha(g)(q)\lambda^\beta, \frac{\partial \mathcal{L}}{\partial \dot{q}^A}(\Psi_g^T(x)), 0\right).$$

But

$$\mathcal{F}L(\Psi_g^T(x)) = \left(\Phi_g^B(q), \Lambda_\beta^\alpha(q)\lambda^\alpha, \frac{\partial \mathcal{L}}{\partial \dot{q}^A}(\Psi_g^T(x)), 0\right).$$

That is, $\Psi_g^{T*}(\mathcal{F}L(x)) = \mathcal{F}L(\Psi_g^T(x)) \in M_1$ for all $x \in P_1$, so for all $g \in G$ the following diagrams

$$\begin{array}{ccc} P_k & \xrightarrow{\mathcal{F}L} & M_k \\ (\Psi_k^T)_g \downarrow & & \downarrow (\Psi_k^{T*})_g \equiv (\Psi^{T*}_{|M_k})_g \\ P_k & \xrightarrow{\mathcal{F}L} & M_k \end{array}$$

are commutative. Therefore, the actions Ψ_k^T and Ψ_k^{T*} are equivalent for all k, through the Legendre transformation.

If ξ_{M_k} denotes the vector field generated by $(\Psi_k^{T*})_{\exp(t\xi)}$, $\xi \in \mathfrak{g}$, then, by Proposition 9, we have that $\mathcal{F}L_*(\xi_{P_k}) = \xi_{M_k}\ \forall k$ and

$$\begin{array}{lllllll} \tilde{J}_f : & M_k & \to & \mathfrak{g}^* & & & \\ & z & \mapsto & J_f(z) : & \mathfrak{g} & \to & \mathbb{R} \\ & & & & \xi & \mapsto & \tilde{J}_f(\xi)(z) = \iota\xi_P(z) \end{array}$$

is a momentum map associated to Ψ_f^{T*} such that $\mathcal{F}L^* \tilde{J}_f = J_f$, which gives the constants of the motion $\tilde{J}_f(\xi) : P_f \to \mathbb{R}$, $z \mapsto \tilde{J}_f(\xi)(z)$. Locally, $\tilde{J}_f$ simply reads as $\tilde{J}_f(\xi)(z) = p_A(z)\xi_Q^A(z)$.

7. Applications to control theory

In this section, we describe optimal control problems in terms of vakonomic mechanics and apply the theory of symmetries developed in the preceding sections.

7.1. General optimal control problems

A general optimal control problem consists of a set of differential equations

$$\dot{x}^i = f^i(x(t), u(t)), \quad 1 \leq i \leq n, \tag{9}$$

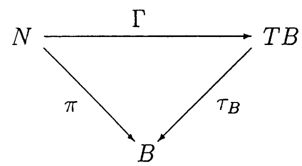
where x^i denote the states and u the control variables, and a cost function $L(x, u)$. Given initial and final states x_0, x_f, the objective is to find a C^2-piecewise smooth curve $c(t) =$

$(x(t), u(t))$ such that $x(t_0) = x_0$, $x(t_f) = x_f$, satisfying the control equations (9) and minimizing the functional

$$\mathcal{J}(c) = \int_{t_0}^{t_f} L(x(t), u(t))\,\mathrm{d}t.$$

In a global description, one assumes an affine bundle structure $\pi : N \to B$, where B is the configuration manifold with local coordinates x^i and N the bundle of controls, with local coordinates (x^i, u^a).

The ordinary differential equations (9) on B depending on the parameters u can be seen as a vector field Γ along the projection map π, i.e. Γ is a smooth map $\Gamma : N \to TB$ such that the diagram



is commutative. This vector field is locally written as $\Gamma = f^i(x, u)(\partial/\partial x^i)$.

In the following, similarly to [15], we show how this kind of problems admit a formulation in terms of vakonomic dynamics. Consider the cost function $L : N \to \mathbb{R}$ and its pullback $\tau_N^* L$ to *TN*. Let us define the set

$$M = \{v \in TN | \pi_*(v) = \Gamma(\tau_N(v))\},$$

and assume that it is a submanifold of *TN*. Locally, this submanifold is defined by the conditions $\dot{x}^i = f^i(x, u)$, $1 \leq i \leq n$, which are just the differential equations (9). Then, to solve the vakonomic problem with Lagrangian $\tau_N^* L : TN \to \mathbb{R}$ and constraint submanifold $M \subset TN$ is equivalent to solve the original general optimal control problem. Moreover, one can make use of the already developed theory of the dynamics of vakonomic systems in the singular Lagrangian framework and of the different types of symmetries associated to such systems, to analyse general control problems.

Remark 10. An alternative way of rephrasing the general optimal control problem in terms of a constrained variational problem is considered in [3,4]. Assuming that Eq. (9) determines u as a function of $(x, \dot{x})$, one can pose the vakonomic problem with Lagrangian $L = L(x, u(x, \dot{x}))$ and constraints $\dot{x} - f(x, u(x, \dot{x}))$ on *TB*. In particular, it can be shown that the condition found in [3] to be able to generalize the Legendre transformation arises naturally in the vakonomic setting as the *compatibility condition* between the Lagrangian and the constraints [8,23], provided L is regular.

If one performs the Gotay–Nester algorithm with the extended Lagrangian $\mathcal{L} = L + \lambda_i(\dot{x}^i - f^i(x, u))$, one finds that the second constraint submanifold P_2 is the final constraint manifold if and only if the matrix

$$W_{ab} = \frac{\partial^2 L}{\partial u^a \partial u^b} - \lambda_i \frac{\partial^2 f^i}{\partial u^a \partial u^b}$$

is invertible, which is exactly the characterization found in [10] for the so-called regular optimal control problem.

On the other hand, one can easily state a version of the Noether's theorem for general optimal control problems.

Proposition 11 (Noether's theorem). *Consider a regular optimal control problem. Let* $X \in \mathfrak{X}(N)$ *be a vakonomic Noether symmetry (VNS) for* $(\tau_N^* L, M)$ *with associated function* $F : N \to \mathbb{R}$. *Then* $F^v - i_{\bar{X}^c}\theta_{\mathcal{L}} : P_2 \to \mathbb{R}$ *is a constant of the motion along any optimal trajectory.*

Locally, if $X = X^i(x,u)(\partial/\partial x^i) + X^a(x,u)(\partial/\partial u^a)$, then $\bar{X} = X + g^i_j(x,u)\lambda_i(\partial/\partial\lambda_j)$ for some $(g^i_j) : N \to \mathfrak{gl}(n,\mathbb{R})$, and the constant of the motion reads locally as

$$F^v - i_{\bar{X}^c}\theta_{\mathcal{L}} = F^v - \frac{\partial\mathcal{L}}{\partial\dot{x}^i}X^i - \frac{\partial\mathcal{L}}{\partial\dot{u}^a}X^a = F^v - \sum_{i=1}^{n}\lambda_i X^i.$$

This result is a corollary of Proposition 7 when applied to general optimal control problems. This theorem generalizes the results obtained in [10].

7.2. Optimal control problems for kinematic locomotion systems

In the following, we shall focus our attention on a more concrete type of optimal control problems associated to kinematic locomotion systems.

Robotic locomotion can be described via a trivial principal bundle (Q, B, π, G) equipped with a connection $\mathcal{A}$. Examples of locomotion systems described in this framework include legged robots, snake-like robots and wheeled mobile robots [16]. Even the motion of paramecia in fluids at very low Reynolds number can be understood by calculating the geometric phase with respect to a certain connection, determined by the underlying fluid dynamics [34]. We remark here that there is also another important type of problems which do not fall into this category, the so-called *dynamic* locomotion systems. A well-known example of this kind of problems is the snakeboard [5,29]. A treatment of the optimal control problem of such systems using Lagrangian reduction has been developed in [17] and has also been considered within a vakonomic perspective in [9].

Basically, $B = Q/G$ represents the shape space of the robot, G the manifold consisting of all possible positions and orientations of the robot in its environment and Q the system's space of configurations. The rigid motions of the system are given by a free and proper left action of G on Q, $\Phi : G \times Q \to Q$ (left multiplication on Q). The relevant kinematics of the locomotion system is modelled in terms of a connection $\mathcal{A} : TQ \to \mathfrak{g}^*$. Indeed, the horizontal subspace of $\mathcal{A}$ will be the set of velocities for which the constraints on the system (usually of non-slipping-type) are satisfied.

Now, a closed path in the shape space B induces a net motion in the group variables, which is nothing but the holonomy of the connection $\mathcal{A}$ associated to the concrete path.

Denote by (r^a) the local coordinates in the quotient B and by (r^a, g^α) the fibred coordinates on Q such that the surjective submersion π reads as $\pi(r^a, g^\alpha) = (r^a)$. Then, if $\tilde{g}$ is a Riemannian metric on B, define the cost function C of the problem

$$C(r, \dot{r}) = \tfrac{1}{2}\tilde{g}_{ab}\dot{r}^a\dot{r}^b,$$

where $\tilde{g}_{ab}$ are the components of the metric on B in the local chart (r^a). Now, consider the following control problem.

Strong optimal control problem for robotic locomotion. Given two points q_0, q_1 in Q, find the optimal controls $u(\cdot)$ which steer the system from q_0 to q_1 and minimize $\int_0^1 C(r, u)\,\mathrm{d}t$ subject to the constraints $\dot{r} = u,\ g^{-1}\dot{g} = -\mathcal{A}_{\mathrm{loc}}(r)u$.

Observe that the statement of the optimal control problem is equivalent to the vakonomic problem on Q corresponding to the Lagrangian function $L : TQ \to \mathbb{R}$ defined as $L(v_q) = \frac{1}{2}\tilde{g}(\pi_* v_q, \pi_* v_q)$, or, in fibre coordinates

$$L(r, g, \dot{r}, \dot{g}) = \tfrac{1}{2}\tilde{g}_{ab}(r)\dot{r}^a\dot{r}^b,$$

and the constraint submanifold $M = H \subseteq TQ$, the horizontal subspace of the connection $\mathcal{A}$. That is, $M = \{v_q | \mathcal{A}(v_q) = 0\} = \{(r, g, \dot{r}, \dot{g}) | \dot{g} + g\mathcal{A}_{\mathrm{loc}}(r)\dot{r} = 0\}$. In what follows, we fix a basis $\{e_1, \dots, e_m\}$ of the Lie algebra $\mathfrak{g}$. Then we have $\mathcal{A}(v_q) = \mathcal{A}^\alpha(v_q)e_\alpha$, where $\mathcal{A}^\alpha$ are functions on TQ defining M globally. Alternatively, we can write

$$\mathcal{A}_{\mathrm{loc}}(r)\dot{r} = \mathcal{A}_a^\alpha(r)\dot{r}^a e_\alpha, \qquad TL_g e_\alpha = M_\alpha^\beta(g)\frac{\partial}{\partial g^\beta},$$

and consider the constraint functions $\phi^\alpha = \dot{g}^\alpha + \mathcal{A}_a^\beta(r)\dot{r}^a M_\beta^\alpha(g),\ 1 \le \alpha \le m$. These functions define M locally (because of the choice of coordinates). However, we will use them in our description, for reasons that will be more clear later.

Now, we apply the theory explained in former sections to the extended Lagrangian $\mathcal{L} = L + \lambda_\alpha\phi^\alpha$ defined on $TP = T(Q \times \mathbb{R}^m)$, and see what happens. By the equivalence theorem, we can perform the constraint algorithm either in the Lagrangian context or in the Hamiltonian one, provided that our Lagrangian $\mathcal{L}$ is almost regular.

In our case, the local expression of the Legendre transformation is

$$(r^a, g^\alpha, \lambda_\alpha, \dot{r}^a, \dot{g}^\alpha, \dot{\lambda}_\alpha) \overset{\mathcal{F}L}{\mapsto} (r^a, g^\alpha, \lambda_\alpha, \tilde{g}_{ab}\dot{r}^b + \lambda_\alpha M_\beta^\alpha(g)\mathcal{A}_a^\beta(r), \lambda_\alpha, 0).$$

Observe that

$$z = (r^a, g^\alpha, \lambda_\alpha, p_a, p_\alpha, p^\alpha) \in \mathcal{F}L(TP) \Leftrightarrow p^\alpha = 0,\ p_\alpha = \lambda_\alpha \quad \forall\alpha.$$

The right implication is clear, while the left one is equivalent to say that

$$\mathcal{F}L^{-1}(z) = \{(r^a, g^\alpha, \lambda_\alpha, \tilde{g}^{ab}(p_b - \lambda_\alpha M_\beta^\alpha\mathcal{A}_b^\beta), \dot{g}^\alpha, \dot{\lambda}_\alpha) | \dot{g}^\alpha, \dot{\lambda}_\alpha \in \mathbb{R}^m\},$$

which can be identified with $\mathbb{R}^{2m}$. Then, by the rank theorem, $M_1 = \mathcal{F}L(TP)$ is the manifold with atlas given by the local charts $\{(r^a, g^\alpha, \lambda_\alpha, p_a)\}$, embedded into T^*P as

$$\begin{array}{cccc} j: & M_1 & \hookrightarrow & T^*P \\ & (r^a, g^\alpha, \lambda_\alpha, p_a) & \mapsto & (r^a, g^\alpha, \lambda_\alpha, p_a, \lambda_\alpha, 0) \end{array}$$

Besides, $\mathcal{F}L^{-1}(z) \equiv \mathbb{R}^{2m}$ are connected submanifolds of $P_1\ \forall z \in M_1$. Hence, $\mathcal{L}$ is almost regular and so we can perform the constraint algorithm on M_1.

Notice that the dimension of M_1 is just $2n$ and that we can regard the Lagrange multipliers λ^α as the generalized momenta corresponding to g^α. Now, consider the pullback of the canonical 1-form ω_P of T^*P to M_1. We obtain

$$\omega_1 = j^*\omega_P = \mathrm{d}r^a \wedge \mathrm{d}p_a + \mathrm{d}g^\alpha \wedge \mathrm{d}\lambda_\alpha,$$

which is clearly symplectic. (M_1, ω_1, h_1) is then a symplectic system, where $h_1 : M_1 \to \mathbb{R}$ is the push forward by $\mathcal{F}L$ of the energy $E_{\mathcal{L}}$. The local explicit expression for h_1 is

$$h_1 = \tfrac{1}{2}\tilde{g}^{cd}(r)(p_c - \mathcal{A}_c^\alpha(r)M_\alpha^\beta(g)\lambda_\beta)(p_d - \mathcal{A}_d^\alpha(r)M_\alpha^\beta(g)\lambda_\beta).$$

Therefore, the Gotay–Nester algorithm for this system will stop at the first step and we have a well-defined solution of the vakonomic problem, and of the optimal control problem, on M_1.

Remark 12. In [25], the author proves a nice theorem which essentially asserts that a curve $q(t)$ in Q is a solution of the optimal control problem if and only if there is a curve $z(t)$ in T^*Q satisfying $\tau_Q(z(t)) = q(t)$, which is a solution curve of a Hamilton differential equation with Hamiltonian H_0. To define H_0, we first take $\mathbf{h} : \pi^*(T(Q/G)) \to TQ$ the horizontal lift operator associated to the connection $\mathcal{A}$ and consider the dual operator

$$\begin{array}{rccc} \mathbf{h}^* : & T^*Q & \to & \pi^*(T(Q/G))^* \\ & p_q & \mapsto & \mathbf{h}_q^*(p_q) \in T^*_{\pi(q)}(Q/G) \end{array}$$

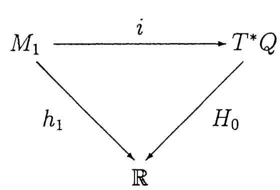
Next, using the vector bundle isomorphisms induced by $\tilde{g}$, $\flat : T(Q/G) \to T^*(Q/G)$ and $\sharp : T^*(Q/G) \to T(Q/G)$, where $\flat(X)(Y) = \tilde{g}(X, Y)$ and $\sharp = \flat^{-1}$, define H_0 as

$$H_0(q, p) = \tfrac{1}{2}\tilde{g}(\sharp\mathbf{h}_q^* p, \sharp\mathbf{h}_q^* p).$$

The relation between this result and what we have obtained here is the following. We can embed the manifold M_1 into T^*Q simply as

$$\begin{array}{rccc} i : & M_1 & \to & T^*Q \\ & (r^a, g^\alpha, \lambda_\alpha, p_a) & \mapsto & (r^a, g^\alpha, p_a, \lambda_\alpha) \end{array}$$

regarding the λ_α as the generalized momenta corresponding to g^α. Now, some easy computations prove that the next diagram



is commutative. The point we want to stress here is that the formulation of our problem in vakonomic terms and the use, via the extended Lagrangian, of tools of singular Lagrangian theory has led us to the same results of [18,25] in a straightforward way.

The action of the Lie group G on TQ leaving invariant the Lagrangian L and the constraint submanifold M can be lifted to an action on TP leaving invariant the extended Lagrangian $\mathcal{L}$. Take $g \in G$ and consider the diffeomorphism $\Phi_g : Q \to Q$. Since $L = L(r, \dot{r})$, we have that $L \circ T\Phi_g = L$. After some computations it can be verified that

$$\phi^\alpha \circ T\Phi_h = \Lambda^\alpha_\gamma(g, h)\phi^\gamma,$$

where $T_g L_h(\partial/\partial g^\gamma)_g = \Lambda^\beta_\gamma(g, h)(\partial/\partial g^\beta)_{hg}$. In addition, as $L_h = L_{hg} \circ L_{g^{-1}}$, we get $\Lambda^\alpha_\gamma(g, h) = M^\alpha_\beta(hg)\bar{M}^\beta_\gamma(g)$, where $\bar{M}(g)$ denotes the inverse matrix of $M(g)$.

Consequently, we have an action of the Lie group on the manifold M_1 given by

$$\begin{array}{ccc} G \times M_1 & \to & M_1 \\ (h, (r, g, \lambda, p)) & \mapsto & (r, hg, \bar{\Lambda}(g, h^{-1})\lambda, p) \end{array}$$

It is easy to see that this action is free and proper, and leaves invariant the symplectic form ω_1 and the Hamiltonian h_1.

Now, we perform a Poisson reduction on (M_1, ω_1, h_1). We choose local coordinates (r^a, μ_α, p_a) on M_1/G, just taking the representative $(r^a, e, \mu_\alpha, p_a)$ of each equivalence class. Then, the equations of motion on M_1/G are

$$\begin{aligned} \dot{r}^a &= \tilde{g}^{ad}(p_d - \mathcal{A}^\alpha_d(r)\mu_\alpha), \\ \dot{p}_a &= \tilde{g}^{cd}\frac{\partial \mathcal{A}^\alpha_c}{\partial r^a}\mu_\alpha(p_d - \mathcal{A}^\alpha_d(r)\mu_\alpha) - \frac{1}{2}\frac{\partial \tilde{g}^{cd}}{\partial r^a}(p_c - \mathcal{A}^\alpha_c(r)\mu_\alpha)(p_d - \mathcal{A}^\alpha_d(r)\mu_\alpha), \\ \dot{\mu}_\alpha &= \dot{r}^c \mathcal{A}^\beta_c c^\gamma_{\alpha\beta}\mu_\gamma, \end{aligned} \tag{10}$$

where $c^\gamma_{\alpha\beta}$ are the structural constants of the Lie algebra $\mathfrak{g}$. Through the change of coordinates $\tilde{p}_a = p_a - \mathcal{A}^\alpha_a(r)\mu_\alpha$, the equations of motion (10) take the simpler form

$$\dot{r}^a = \tilde{g}^{ad}\tilde{p}_d, \qquad \dot{\tilde{p}}_a = \mu_\alpha \mathcal{B}^\alpha_{ca}\dot{r}^c - \frac{1}{2}\frac{\partial \tilde{g}^{cd}}{\partial r^a}\tilde{p}_c\tilde{p}_d, \qquad \dot{\mu}_\alpha = \dot{r}^c \mathcal{A}^\beta_c c^\gamma_{\alpha\beta}\mu_\gamma, \tag{11}$$

where $\mathcal{B}^\alpha_{ca} = (\partial \mathcal{A}^\alpha_c/\partial r^a) - (\partial \mathcal{A}^\alpha_a/\partial r^c) - c^\alpha_{\beta\gamma}\mathcal{A}^\beta_a\mathcal{A}^\gamma_c$ are the local coefficients of the curvature of the connection $\mathcal{A}$. Eqs. (11) are precisely Wong's equations [17,24,25]. Here it is the reason why in the beginning of this section we chose the constraint functions ϕ^α and not the $\mathcal{A}^\alpha$. Although both formulations are clearly equivalent, the derivation of Wong's equations is more straightforward with the choice done.

If we perform a symplectic reduction on the symplectic manifold (M_1, ω_1), we indeed obtain reduced symplectic manifolds. A standard result in the theory of Hamiltonian systems with symmetry (see Theorem 6.48 in [28]) states that the reduced symplectic manifolds obtained by using the momentum map can be seen as submanifolds of M_1/G; more precisely, they constitute the canonical symplectic foliation of the Poisson structure.

Acknowledgements

This work was partially supported by grant DGICYT (Spain) PB97-1257. S. Martínez and J. Cortés wish to thank the Spanish Ministerio de Educación y Cultura for FPI and FPU grants, respectively. We would like to thank F. Cantrijn for helpful conversations and critical reading of preliminary versions of this paper, as well as the referee for pointing out us references [3,4].

Appendix A

We briefly review in this section the basic notions of vertical and complete lifts of vector fields and functions. We refer to [21,35] for a comprehensive treatment of the subject.

Let (y^a) be local coordinates of a manifold N and let $(y^a, \dot{y}^a)$ (respectively (y^a, p_a)) be the induced fibred coordinates in TN (respectively T^*N). Consider a vector field $Y \in \mathfrak{X}(N)$ with local expression $Y = Y^a(\partial/\partial y^a)$ and let $F : N \to \mathbb{R}$ be a function on N.

A 1-form β in N can be regarded as a function in TN. We denote it as $\iota\beta$. If β is locally written as $\beta = \beta_a \, \mathrm{d}y^a$, then $\iota\beta$ reads as $\iota\beta = \beta_a y^a$. Similarly, ιY is the function given as

$$\begin{array}{rccc} \iota Y: & T^*N & \to & \mathbb{R} \\ & (y^a, p_a) & \mapsto & \iota Y(y^a, p_a) = p_a Y^a \end{array}$$

The *complete lift* of F to TN is another function $F^c : TN \to \mathbb{R}$, defined as $F^c = \iota(\mathrm{d}F)$. Locally, $F^c = (\partial F/\partial y^a)\dot{y}^a$.

The *complete lift* of Y to TN is the unique vector field $Y^c \in \mathfrak{X}(TN)$ such that $\forall F \in C^\infty(N)$, $Y^c(F^c) = (YF)^c$. The local expression for Y is

$$Y^c = Y^a \frac{\partial}{\partial y^a} + \dot{y}^b \frac{\partial Y^a}{\partial y^b} \frac{\partial}{\partial \dot{y}^a}.$$

The *complete lift* of Y to T^*N, $Y^{*c} \in \mathfrak{X}(T^*N)$, is the Hamiltonian vector field associated to the function ιY, i.e.

$$i_{Y^{*c}}\omega = \mathrm{d}(\iota Y),$$

where ω is the canonical 2-form on T^*N, defined as $\omega = -\mathrm{d}\theta$ from the Liouville 1-form $\theta = p_a \, \mathrm{d}y^a$. Locally, we obtain

$$Y^{*c} = Y^a \frac{\partial}{\partial y^a} - p_b \frac{\partial Y^a}{\partial y^b} \frac{\partial}{\partial p_a}.$$

The vertical lift of F to TN is its pullback to TN by the canonical projection $\tau_N : TN \to N$. We denote it as $F^v = \tau_N^* F$. On the other hand, *the vertical lift of F to* T^*N is the pullback $F^{*v} = \pi_N^* F$ by $\pi_N : T^*N \to N$.

Finally, we recall some different types of symmetries associated to a presymplectic system (N, ω, α).

In first place, consider the presymplectic equation

$$i_Z\omega = \alpha, \tag{12}$$

and the sequence of submanifolds that results of applying the Gotay–Nester algorithm

$$\cdots \hookrightarrow N_k \hookrightarrow \cdots \hookrightarrow N_2 \hookrightarrow N_1.$$

We use the following notation

$$\mathfrak{X}^\omega(N) = \{Z \in \mathfrak{X}(N) | Z \text{ is a solution of} (12)\},$$
$$\mathfrak{X}^\omega(N_k) = \{Z \in \mathfrak{X}(N_k) | Z \text{ is a solution of} (i_Z\omega = \alpha)_{|N_k}\}.$$

A *dynamical symmetry* of $\mathfrak{X}^\omega(N)$ (respectively of $\mathfrak{X}^\omega(N_k)$) is a vector field $Y \in \mathfrak{X}(N)$ (respectively $Y \in \mathfrak{X}(N_k)$) such that

$$[Y, Z] \in \operatorname{Ker}\omega \quad (\text{respectively } [Y, Z] \in \operatorname{Ker}\omega \cap TN_k)$$

for all $Z \in \mathfrak{X}^\omega(N)$ (respectively $Z \in \mathfrak{X}^\omega(N_k)$). In this way, we assure that the flow of Y preserves solutions, i.e. the integral curves of Z are transformed into solutions of the system (see [19]).

A *Cartan symmetry* of (N, ω, α) will be a $Y \in \mathfrak{X}(N)$ such that

1. $i_Y\omega = \mathrm{d}F$ for some $F : N \to \mathbb{R}$.
2. $i_Y\alpha = 0$.

References

[1] V.I. Arnold, Dynamical Systems, Vol. III, Springer, New York, 1988.

[2] A.M. Bloch, P.E. Crouch, Nonholonomic and vakonomic control systems on Riemannian manifolds, in: M.J. Enos (Ed.), Fields Institute Communications, Vol. 1, 1993, pp. 25–52.

[3] A.M. Bloch, P.E. Crouch, Reduction of Euler–Lagrange problems for constrained variational problems and relation with optimal control problems, in: Proceedings of the IEEE Conference on Decision & Control, Lake Buena Vista, USA, 1994, pp. 2584–2590.

[4] A.M. Bloch, P.E. Crouch, Optimal control, optimization and analytical mechanics, in: J. Baillieul, J.C. Willems (Eds.), Mathematical Control Theory, Springer, New York, 1999, pp. 268–321.

[5] A.M. Bloch, P.S. Krishnaprasad, J.E. Marsden, R.M. Murray, Nonholonomic mechanical systems with symmetry, Arch. Rat. Mech. Anal. 136 (1996) 21–99.

[6] F. Cardin, M. Favretti, On nonholonomic and vakonomic dynamics of mechanical systems with nonintegrable constraints, J. Geom. Phys. 18 (4) (1996) 95–325.

[7] J.F. Cariñena, M. Rañada, Noether's theorem for singular Lagrangians, Lett. Math. Phys. 15 (1988) 305–311.

[8] J. Cortés, M. de León, D. Martín de Diego, S. Martínez, Geometric description of vakonomic and nonholonomic dynamics, Comparison of solutions, Preprint, IMAFF-CSIC.

[9] J. Cortés, S. Martínez, Optimal control for nonholonomic systems with symmetry, in: Proceedings of the IEEE Conference on Decision & Control, Sydney, Australia, 2000, in press.

[10] M. Delgado-Téllez, A. Ibort, Some geometrical methods in optimal control theory, Preprint, 2000.

[11] A. Echeverría-Enríquez, M.C. Muñóz-Lecanda, N. Román-Roy, Reduction of presymplectic manifolds with symmetry, Rev. Math. Phys. 11 (1999) 1209–1247.

[12] M. Favretti, Equivalence of dynamics for nonholonomic systems with transverse constraints, J. Dyn. Diff. Eqs. 10 (4) (1998) 511–536.

[13] M.J. Gotay, J.M. Nester, Presymplectic Lagrangian systems I: the constraint algorithm and the equivalence theorem, Ann. Inst. H. Poincaré A30 (1979) 129–142.

[14] M.J. Gotay, J.M. Nester, Presymplectic Lagrangian systems II: the second-order differential equation problem, Ann. Inst. H. Poincaré A32 (1980) 1–13.
[15] X. Gràcia, J. Marín-Solano, M. Muñoz-Lecanda, Variational principles in mechanics: geometric aspects, in: Proceedings of the VIIth Fall Workshop, Geometry and Physics, Valencia, September 1998.
[16] S.D. Kelly, R.M. Murray, Geometric phases and robotic locomotion, J. Robotic Syst. 12 (6) (1995) 417–431.
[17] W.S. Koon, J.E. Marsden, Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction, SIAM J. Control Opt. 35 (1997) 901–929.
[18] P.S. Krishnaprasad, R. Yang, W. Dayawansa, Control problems on principal bundles and nonholonomic mechanics, in: Proceedings of the IEEE Conference on Decision & Control, Brighton, England, 1991, pp. 1133–1138.
[19] M. de León, D. Martín de Diego, Symmetries and constants of the motion for singular Lagrangian systems, Int. J. Theoret. Phys. 35 (5) (1996) 975–1011.
[20] M. de León, J.C. Marrero, D. Martín de Diego, Vakonomic mechanics versus nonholonomic mechanics: a unified geometrical approach, J. Geom. Phys. 35 (2–3) (2000) 126–144.
[21] M. de León, P.R. Rodrigues, Methods of Differential Geometry in Analytical Mechanics, Mathematics Studies, Vol. 158, North-Holland, Amsterdam, 1989.
[22] A.D. Lewis, R.M. Murray, Variational principles for constrained systems: theory and experiment, Int. J. Nonlinear Mech. 30 (6) (1995) 793–815.
[23] 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.
[24] R. Montgomery, Canonical formulations of a particle in a Yang–Mills field, Lett. Math. Phys. 8 (1984) 59–67.
[25] R. Montgomery, Isoholonomic problems and some applications, Commun. Math. Phys. 128 (1990) 565–592.
[26] R. Montgomery, Abnormal minimizers, SIAM J. Control Opt. 32 (1994) 1605–1620.
[27] J. Neimark, N. Fufaev, Dynamics of Nonholonomic Systems, Transactions of Mathematical Monographs, Vol. 33, American Mathematical Society, Providence, RI, 1972.
[28] P.J. Olver, Applications of Lie groups to Differential Equations, GTM No. 107, 2nd Edition, Springer, Berlin, 1993.
[29] J.P. Ostrowski, Geometric perspectives on the mechanics and control of undulatory locomotion, Ph.D. Thesis, California Institute of Technology, 1995.
[30] G. Prince, Toward a classification of dynamical systems in classical mechanics, Bull. Aust. Math. Soc. 27 (1983) 53–71.
[31] G. Prince, A complete classification of dynamical symmetries in classical mechanics, Bull. Aust. Math. Soc. 32 (1985) 299–308.
[32] P.A. Samuelson, Law of conservation of the capital–output ratio. Proceedings of the National Academy of Sciences, Appl. Math. Sci. 67 (1970) 1477–1479.
[33] R. Sato, R.V. Ramachandran, Conservation Laws and Symmetry: Applications to Economics and Finance, Kluwer Academic Publishers, Boston, MA, 1990.
[34] A. Shapere, F. Wilczek, Geometry of self-propulsion at low Reynolds number, J. Fluid. Mech. 198 (1989) 557–585.
[35] K. Yano, S. Ishihara, Tangent and Cotangent Bundles, Marcel Dekker, New York, 1973.