Lagrangian reduction of generalized nonholonomic systems

Hernán Cendra\textsuperscript{a,b}, Sebastián Ferraro\textsuperscript{a,c}
Sergio Grillo\textsuperscript{b,d}

\textsuperscript{a} Departamento de Matemática, Universidad Nacional del Sur, 8000 Bahía Blanca, Argentina
\textsuperscript{b} CONICET, Argentina
\textsuperscript{c} ICMat—Instituto de Ciencias Matemáticas (Consejo Superior de Investigaciones Científicas, Universidad Autónoma de Madrid, Universidad Complutense de Madrid, Universidad Carlos III de Madrid), Madrid, Spain
\textsuperscript{d} Centro Atómico Bariloche and Instituto Balseiro, 8400 S.C. de Bariloche, Argentina

Abstract

In this paper we study the Lagrangian reduction of generalized nonholonomic systems (GNHS) with symmetry. We restrict ourselves to those GNHS, defined on a configuration space $Q$, with kinematic constraints given by a general submanifold $C_K \subset TQ$, and variational constraints given by a distribution $C_V$ on $Q$. We develop a reduction procedure that is similar to that of nonholonomic systems satisfying d'Alembert's principle, i.e. with $C_K$ a distribution and $C_V = C_K$. Special care is taken in identifying the geometrical structures and mappings involved. We illustrate the general theory with an example.

1 Introduction

There is a class of mechanical systems with constraints, which includes simplified models of pneumatic tires [1, 2, 3, 4] and certain servomechanisms [5, 6], that do not satisfy d'Alembert's principle nor its nonlinear generalization, Chetaev's principle [7, 8, 9] (see also [10]). Roughly speaking, this means (for linear constraints) that the constraint forces are not necessarily orthogonal to the admissible velocities. Such systems have been studied systematically in [1] and [5] from the Lagrangian and Hamiltonian sides, respectively, and then compared in [11] via the Legendre transformation. In the last reference, they have been called \textit{generalized nonholonomic systems} (GNHS). In order to write down their equations of motion, it is not enough to provide the kinematic constraints
alone, but also the subspace where the constraint forces live, i.e. the space of possible values of the constraint forces. Thus, GNHSs are given by triples $(L, C_K, F_V)$, consisting of a Lagrangian function $L: TQ \to \mathbb{R}$ on a configuration space $Q$, a submanifold $C_K \subset TQ$ that defines the kinematic constraints (the set of admissible velocities), and a codistribution $F_V \subset T^*Q$ where the constraint forces take their values.\(^1\) For linear constraints, that is to say, when $C_K$ is a distribution on $Q$, and when d’Alembert’s principle holds, constraint forces are given by the annihilator of $C_K$, i.e., $F_V = C_K^\circ$: constraint forces are orthogonal to admissible velocities. GNHSs can be described within the variational realm with variations restricted to $C_V = F_V^\circ \subset TQ$, the distribution of the so-called variational constraints. Note that, when kinematic constraints are linear and d’Alembert’s principle is valid, then $C_V = C_K$. In this setting, GNHSs are defined by the equivalent data $(L, C_K, C_V)$. The elements of $C_V$ can be interpreted as the virtual displacements; thus, we have a generalization of the Principle of Virtual Work: no work is done by the constraint forces along directions defined by the virtual displacements. This is a generalization in the sense that, in general, virtual displacements do not coincide with admissible velocities, as in the d’Alembert case.

In this paper we shall consider GNHSs with symmetry, and our main goal will be to develop a theory of Lagrangian reduction for them. We shall formulate a reduced variational principle, and write down the corresponding reduced equations, namely the Generalized Lagrange–d’Alembert–Poincaré Equations for GNHSs. We follow similar steps to those of Ref. [14], the nonholonomic counterpart of the usual Lagrangian reduction (see [15] and references therein).\(^2\) Nevertheless, in order to deal with this more general context, we have decided to make some of these steps more explicit. This has forced us to introduce a rather different notation.

The paper is organized as follows. In Section 2, we first recall some well-known facts about Lagrangian mechanics in the terminology we shall use along the entire work; then, we define the notion of GNHS, present some relevant examples, and construct the local and global versions of the related Generalized Lagrange–d’Alembert–Poincaré Equations. In Section 3, we define the notion of a $G$-invariant GNHS, with $G$ a Lie group acting on the configuration space of the system, we build up the associated reduced data, and then, in terms of a generalized version of the so-called nonholonomic connection (see [14]), we construct the reduced variations and the corresponding reduced variational principle. Finally, we write down the Generalized Lagrange–d’Alembert–Poincaré Equations for these systems, and show that they are equivalent to the original (unreduced) equations of motion. In the last section, as an illustrative example, we consider the stabilization problem of a ball rolling without sliding on a moving platform.

\(^1\)Actually, the systems studied in [1, 5, 11] are more general than those described here. For instance, the kinematic constraints considered in [1] involve higher order derivatives, i.e., they are given by a submanifold of some higher order tangent bundle $T^{(k)}Q$, $k \in \mathbb{N}$ [12, 13].

\(^2\)For nonholonomic systems with symmetry (satisfying d’Alembert’s principle) studied from the Hamiltonian side, see for example [16]–[22].
**Basic notation and terminology.** We shall use a notation close to that of Ref. [23]. Let $Q$ be an $n$-dimensional real manifold. Consider a chart $(U, \varphi)$ of $Q$, with $\varphi: U \to \mathbb{R}^n$. Given $q \in U$, we write

$$\varphi(q) = (q^1, \ldots, q^n).$$

Sometimes, we shall also use $q$ to denote the $n$-tuple $(q^1, \ldots, q^n)$, depending on the context. Given a curve $\gamma: [t_1, t_2] \to Q$ whose range intersects $U$, we write

$$\varphi(\gamma(t)) = (q^1(t), \ldots, q^n(t)) \equiv q(t),$$

in the open interval where $\varphi \circ \gamma$ is defined. The *velocity* of $\gamma$ is given by the application $\gamma'(t_1, t_2) \to TQ$ defined by

$$\gamma'(t) = \frac{d}{dt} \gamma(t) = \gamma_* (d/dt|_{t}) \in T_{\gamma(t)}Q.$$

We shall consider for the tangent bundle of $Q$ the induced coordinate charts $(TU, \varphi_*)$, where $\varphi_*: TU \to TR^n$ is the differential of $\varphi$. Given $q \in U$ and $X \in T_qQ$, we use the notation

$$\varphi_*(X) = (q^1, \ldots, q^n; q^1_1, \ldots, q^n_n) \quad \text{or} \quad (q^1, \ldots, q^n; \delta q^1, \ldots, \delta q^n).$$

We will write, for short, $\varphi_*(X) = (q, \dot{q})$ or $(q, \delta q)$. Given a curve $\gamma$ as above, we write its velocity as

$$\varphi_*(\gamma'(t)) = (q^1(t), \ldots, q^n(t); \dot{q}^1(t), \ldots, \dot{q}^n(t)) \equiv (q(t), \dot{q}(t)).$$

### 2 Generalized nonholonomic mechanics

In this section we present the definition of generalized nonholonomic systems (GNHS), from the Lagrangian point of view, as given in [11]. In order to do that, we first write the equations of motion of a (standard) Lagrangian system in terms of infinitesimal variations. After that, with the definition of GNHS at hand, we give some examples of such systems, write down their equations of motion locally and, by introducing a connection on the configuration space, we construct the corresponding global version.

#### 2.1 Infinitesimal variations

Consider a Lagrangian system defined on a configuration space $Q$ and with Lagrangian function $L: TQ \to \mathbb{R}$. For short, we shall refer to it as the pair $(Q, L)$. Let us recall how the trajectories of $(Q, L)$ are usually described in the variational formalism.

Let $\gamma: [t_1, t_2] \to Q$ be a curve in $Q$, and $\gamma': (t_1, t_2) \to TQ$ its velocity.

- The action $S$ evaluated at $\gamma$ is the integral

$$S[\gamma] = \int_{t_1}^{t_2} L(\gamma'(t)) \, dt.$$
• Deformations of $\gamma$ are smooth mappings

$$\Delta \gamma: [t_1, t_2] \times (-\varepsilon, \varepsilon) \rightarrow Q$$

such that

a. $\Delta \gamma(t_1, \lambda) = \gamma(t_1)$ and $\Delta \gamma(t_2, \lambda) = \gamma(t_2)$ for all $\lambda \in (-\varepsilon, \varepsilon)$;

b. $\Delta \gamma(t, 0) = \gamma(t)$ for all $t \in [t_1, t_2]$.

Note that for each fixed $\lambda$ we have a curve on $Q$ given by

$$\Delta \gamma: [t_1, t_2] \rightarrow Q : t \mapsto \Delta \gamma(t, \lambda).$$

Clearly we have $\Delta \gamma_0 = \gamma$ and $\Delta \gamma_\lambda(t_i) = \gamma(t_i)$ for $i = 1, 2$, and for all $\lambda$.

• By the very definition of a Lagrangian system, the curve $\gamma$ is a trajectory of $(Q, L)$ if for all deformations $\Delta \gamma$ of $\gamma$ we have

$$\frac{dS[\Delta \gamma_\lambda]}{d\lambda} \bigg|_{\lambda=0} = \lim_{\lambda \rightarrow 0} \frac{S[\Delta \gamma_\lambda] - S[\gamma]}{\lambda} = 0. \quad (2)$$

We want to write the equation above in terms of infinitesimal variations of $\gamma$, i.e., in terms of curves $\delta \gamma: [t_1, t_2] \rightarrow TQ$ satisfying

a. $\delta \gamma(t) \in T_{\gamma(t)}Q$, $\forall t \in [t_1, t_2]$;

b. $\delta \gamma(t_1), \delta \gamma(t_2)$ belongs to the zero section of $TQ$.

The velocity of $\delta \gamma$ will be denoted by $\delta \gamma'$, where

$$\delta \gamma': \ (t_1, t_2) \rightarrow TTQ : t \mapsto \delta \gamma'(t) = \delta \gamma_* \left( \frac{d}{dt} \big|_{t_1} \right). \quad (3)$$

Note that $\delta \gamma'(t) \in T_{\delta \gamma(t)}TQ$.

Remark 1. Using the canonical projection $\tau_Q: TQ \rightarrow Q$, condition a can be rewritten as $\tau_Q \circ \delta \gamma = \gamma$. Considering also $\tau_{TQ}: TTQ \rightarrow TQ$, we have

$$\tau_{TQ} \circ \delta \gamma' = \delta \gamma' \quad \text{and} \quad (\tau_Q)_* \circ \delta \gamma' = \gamma', \quad (4)$$

where $(\tau_Q)_*: TTQ \rightarrow TQ$ is the differential of $\tau_Q$.

It is easy to see that any infinitesimal variation $\delta \gamma$ is given by a deformation $\Delta \gamma$, and every deformation gives rise to a variation, by the formula

$$\delta \gamma(t) = \Delta \gamma_* \left( \frac{\partial}{\partial \lambda} \big|_{t, 0} \right). \quad (5)$$

In fact, fixing a Riemannian structure on $Q$, for each infinitesimal variation $\delta \gamma$ we have a related deformation $\Delta \gamma(t, \lambda) = \exp_{\gamma(t)}(\lambda \delta \gamma(t))$ satisfying (5), where $\exp_q: T_qQ \rightarrow Q$ is the exponential map defining the geodesic curves.

Using this relationship, we have the following result.
Theorem 1. A curve $\gamma: [t_1, t_2] \rightarrow Q$ is a trajectory of the Lagrangian system $(Q, L)$, i.e., it satisfies (2), if and only if for every infinitesimal variation $\delta\gamma$ of $\gamma$ we have

$$\int_{t_1}^{t_2} \langle dL(\gamma'(t)), \kappa(\delta\gamma'(t)) \rangle \, dt = 0. \quad (6)$$

Before proving this, let us recall the definition of the canonical involution $\kappa: TTQ \rightarrow TTQ$ (see, for instance, Ref. [25]), that appears in Eq. (6), along with some of its properties. For an intrinsic definition, consider an application

$$\Gamma: (-\varepsilon, \varepsilon) \times (-\varepsilon, \varepsilon) \rightarrow Q : (t, \lambda) \mapsto \Gamma(t, \lambda), \quad (7)$$

and interpret the quantities

$$\Gamma_* \left( \frac{\partial}{\partial \lambda} \big|_{(t, 0)} \right) \quad \text{and} \quad \Gamma_* \left( \frac{\partial}{\partial t} \big|_{(0, \lambda)} \right)$$

as the values of applications

$$(-\varepsilon, \varepsilon) \rightarrow TQ : t \mapsto \Gamma_* \left( \frac{\partial}{\partial \lambda} \big|_{(t, 0)} \right)$$

and

$$(-\varepsilon, \varepsilon) \rightarrow TQ : \lambda \mapsto \Gamma_* \left( \frac{\partial}{\partial t} \big|_{(0, \lambda)} \right).$$

Then, every element $V \in TTQ$ can be written as the second derivative

$$V = \left[ \Gamma_* \left( \frac{\partial}{\partial \lambda} \big|_{(t, 0)} \right) \right]_* \left( \frac{d}{dt} \big|_{0} \right)$$

of some $\Gamma$ as above. In these terms, $\kappa$ is defined by

$$\kappa(V) = \left[ \Gamma_* \left( \frac{\partial}{\partial t} \big|_{(0, \lambda)} \right) \right]_* \left( \frac{d}{d\lambda} \big|_{0} \right). \quad (8)$$

In local terms, given a chart $(U, \varphi)$ of $Q$, with induced chart $(TU, \varphi_{**})$ on $TTQ$, we have

$$[\varphi_{**} \circ \kappa \circ \varphi_{**}^{-1}] (q, \dot{q}, \delta q, \delta \dot{q}) = (q, \delta q, \dot{q}, \delta \dot{q}), \quad (9)$$

using a natural notation. Note that $\kappa$ satisfies

$$\kappa \circ \kappa = \text{id}_{TTQ} \quad \text{and} \quad (\tau_Q)_{**} \circ \kappa = (\tau_Q)_*,$$ 

and accordingly

$$(\tau_Q)_* \circ \kappa = \tau_{TTQ}. \quad (11)$$

Proof. [Proof of Theorem 1] For a given curve $\gamma$, consider a deformation $\Delta \gamma$, its related family of curves $\Delta \gamma_\lambda$, and their corresponding velocities

$$\Delta \gamma_\lambda: (t_1, t_2) \rightarrow TQ,$$
given by
\[ \Delta \gamma'(t) = (\Delta \gamma_\lambda)_* \left( \frac{d}{dt}|_{t_1} \right) = \Delta \gamma_* \left( \frac{\partial}{\partial t}|_{(t, \lambda)} \right). \quad (12) \]
Consider also the related infinitesimal variation
\[ \delta \gamma(t) = \Delta \gamma_* \left( \frac{\partial}{\partial \lambda}|_{(t, 0)} \right) \]
[recall Eq. (5)]. Clearly, in order to prove the theorem, since
\[ \left. \frac{dS \left[ \Delta \gamma_\lambda \right]}{d\lambda} \right|_{\lambda=0} = \int_{t_1}^{t_2} \left. d \left( L \left( \Delta \gamma'_\lambda(t) \right) \right) \right|_{\lambda=0} dt, \]
it is enough to show that the equality
\[ \left. \frac{d}{d\lambda} L \left( \Delta \gamma'_\lambda(t) \right) \right|_{\lambda=0} = \langle dL (\gamma'(t)), \kappa (\delta \gamma'(t)) \rangle \quad (14) \]
holds. On one hand,
\[ \left. \frac{d}{d\lambda} L \left( \Delta \gamma'_\lambda(t) \right) \right|_{\lambda=0} = \langle dL (\gamma'(t)), (\Delta \gamma_\lambda'_\lambda(t))_* \left( \frac{d}{d\lambda}|_{0} \right) \rangle \]
where we have used the equality \( \Delta \gamma'_0 (t) = \gamma'(t) \), and the differential \( (\Delta \gamma_\lambda'_\lambda(t))_* \) has been constructed by regarding \( \Delta \gamma'_\lambda(t) \) as a function of \( \lambda \) only [see (12) for the second identity]. On the other hand, using the definition of \( \kappa \) [see Eq. (8)], and Eqs. (3) and (13),
\[
\left. \left[ \Delta \gamma_* \left( \frac{\partial}{\partial t}|_{(t, \lambda)} \right) \right] \right|_{*} \left( \frac{d}{d\lambda}|_{0} \right) = \kappa \left( \left[ \Delta \gamma_* \left( \frac{\partial}{\partial t}|_{(t, 0)} \right) \right] \right|_{*} \left( \frac{d}{d\lambda}|_{0} \right) \]
\[ = \kappa (\delta \gamma_* (d/dt)|_{0}) = \kappa (\delta \gamma'(t)), \quad (15) \]
finishing our proof. \( \square \)

2.2 Definition and examples
Motivated by mechanical systems such as rubber wheels and certain servomechanisms [1, 5], where d’Alembert’s principle is typically violated, we have defined in [11] a class of dynamical systems that include the latter and encode, in our opinion, their main features. We recall that definition below.

**Definition 2.** Given a manifold \( Q \), let us consider the triples \( (L, C_K, C_V) \) with
\[ L: TQ \to \mathbb{R}, \quad C_K, C_V \subset TQ, \]
where \( C_K \) is a submanifold of \( TQ \) and \( C_V \) is a distribution on \( Q \). We shall refer to them as **generalized nonholonomic systems (GNHS)**, with Lagrangian function \( L \), **kinematic constraints** \( C_K \) and **variational constraints** \( C_V \).
The elements of \( C_V \) will be called virtual displacements. We shall say that \( \gamma: [t_1, t_2] \rightarrow Q \) is a trajectory of \((L, C_K, C_V)\) if \( \gamma'(t) \in C_K \), and for all infinitesimal variations \( \delta \gamma \) such that \( \delta \gamma(t) \in C_V \) we have

\[
\int_{t_1}^{t_2} \left( dL(\gamma'(t)) \cdot \kappa(\delta \gamma'(t)) \right) dt = 0.
\]

(16)

Remark 2. Triples as presented above actually constitute a subclass of the systems defined in [11] (identified there as the \( l = 0 \) subclass). Here, we choose the above definition since in this work we are going to focus only on such systems. The rest of the systems appearing in [11] (the \( l = 1 \) subclass) will be studied in a forthcoming paper, within a more general setting.

Constraint forces and virtual work. The annihilator of \( C_V \), namely \( F_V \equiv C_V \circ \) \( \subset \) \( T^*Q \), is interpreted (see [11]) as the space of values of constraint forces. Then, if \( f \in F_V \) and \( v \in C_V \), both over the same point of \( Q \), we have that \( \langle f, v \rangle = 0 \). This enables us to say that constraint forces do not do work along the directions of virtual displacements. Thus, we have a generalized version of the Principle of Virtual Work. The generalization is in the sense that, unless d’Alembert’s principle holds, virtual displacements have nothing to do with allowed velocities.

Note that GNHSs can be also described in terms of triples \((L, C_K, F_V)\).

Affine constraints and d’Alembert’s principle. When \( C_K \) is a distribution on \( Q \) (resp. an affine subbundle of \( TQ \)) we say that the system has linear (resp. affine) kinematic constraints. In these cases we can write \( C_K = C_K^{\text{vec}} + \nu \), where \( C_K^{\text{vec}} \subset TQ \) is a distribution and \( \nu \subset TQ \) is a section. Such submanifolds can be locally described by a set of equations for \( \dot{q} \)

\[
\omega^a(q) \dot{q}^i = \gamma^a(q), \quad 1 \leq a \leq k,
\]

for each \( q \in Q \). (In this work we adhere to the usual convention of summing over repeated indices.) Of course, linear constraints correspond to the case \( \gamma^a = 0 \), or equivalently, to the case \( \nu = 0 \). Variational constraints \( C_V \) will always be locally described by equations of the form

\[
v^b(q) \delta q^i = 0, \quad 1 \leq b \leq l.
\]

D’Alembert’s principle holds when \( C_V = C_K^{\text{vec}} \), i.e., when the functions \( v^b \) coincide with the functions \( \omega^a \); or equivalently, when \( F_V = (C_K^{\text{vec}})^\circ \), which means that constraint forces do not do work along standard virtual displacements—this is the standard Principle of Virtual Work.

It is worth mentioning that, in general, kinematic and variational constraints are given independently, and one should not attempt to derive, for instance, variational from kinematic constraints by a universal procedure (as one does when d’Alembert’s principle holds). This is the case of the Greidanus and Rocard models of a pneumatic tire, studied in Ref. [1], and the next examples inspired in automatic control problems.
The inverted pendulum. The controlled inverted pendulum on a cart, with control strategies given in Refs. [5, 6, 11], can be described as a GNHS on $Q = \mathbb{R} \times S^1$. The Lagrangian function $L$ is simple,\(^3\) and corresponds to the system without control forces. The constraint submanifold $C_K \subset TQ$ is an affine subbundle (whose precise form depends on the control strategy), and $C_V \subset TQ = T\mathbb{R} \times TS^1$ is given by the distribution

$$C_V = O_\mathbb{R} \times TS^1,$$

where $O_\mathbb{R}$ is the zero section of $T\mathbb{R}$. It is then clear that the latter does not depend at all on the form of $C_K$. In particular, d'Alembert's principle does not hold in general for these mechanical systems.

Underactuated system by virtual constraints. Systems subject to the so-called *virtual constraints* (see Ref. [6]) are triples $(L, C_K, C_V)$ such that $C_K$ and $C_V$ are 1-dimensional distributions, as explained in [11]. The aim of the virtual constraints method is to add forces, called *actuators*, to a given Lagrangian system $(Q, L)$, in order to generate asymptotically stable orbits that are not present in the original system. The idea is to design such actuators as the constraint forces of an appropriate set of constraints. In [6] a total of $n-1$ linear constraints (which are actually holonomic constraints) are imposed on a system with $n$ degrees of freedom. These might be implemented by forces acting on $n-1$ selected independent directions (i.e., $n-1$ independent actuators). In other words, $C_K$ is a 1-dimensional distribution and $F_V$ is an $(n-1)$-dimensional codistribution. Therefore its annihilator $C_V$ is also a 1-dimensional distribution.

Optimal control of the plate-ball system. Consider a ball rolling along an arbitrarily prescribed curve on a horizontal plate, without spinning or slipping. As a consequence of the complete process of rolling from the initial to the final point of the curve, the ball undergoes a certain rotation. The optimal control problem consists in finding the shortest curve joining two given points on the plate and inducing a given rotation of the ball.

One can define a GNHS whose trajectories are directly related to the solutions of the optimal control problem (see [26, 27]). The manifold $Q$ is $\text{SO}(3) \times \mathbb{R}^2 \times M$, where $\text{SO}(3) \times \mathbb{R}^2$ is the configuration space of the ball and $M$ is (a suitable subset of) the space of symmetric bilinear forms $b$ on $\text{so}(3)$. There is a natural action of the group $\text{SO}(3)$ on $M$. It can be shown that if $(R(t), y(t), b(t))$ is a trajectory of the GNHS, then $y(t)$ is a solution of the optimal control problem. It is important to remark that $R(t)$ does not represent the actual evolution of the orientation of the rolling ball.

\(^3\)Recall that $L$ is *simple* if there exists a Riemannian metric $\phi: TQ \times TQ \to \mathbb{R}$ and a function $V: Q \to \mathbb{R}$ such that

$$L = \frac{1}{2} Q_\phi - V \circ \tau_Q,$$

where $Q_\phi$ is the quadratic form related to $\phi$. The function $V$ is called the *potential energy* and $T = \frac{1}{2} Q_\phi$ the *kinetic energy.*
The Lagrangian function is the norm squared of the velocity vector $\dot{x}$ of the contact point on the plate, plus a term involving $b$ and a principal connection $A_0$ on the principal bundle $SO(3) \times \mathbb{R}^2 \to \mathbb{R}^2$ that encodes the rolling constraints. One can identify $so(3) \equiv \mathbb{R}$ (see [27]), and in this way each $b$ can be seen as a $3 \times 3$ matrix. The kinematic constraints $C_K$ are then defined by the equation $\dot{b} = [A_0(R, y, \dot{R}, \dot{y}), b]$ (matrix commutator). In turn, $C_V = TSO(3) \times T\mathbb{R}^2 \times \mathbb{R}_M$. In this case, both $C_K$ and $C_V$ are distributions.

This GNHS is $SO(3)$-invariant. It is worth mentioning that the reduction procedure performed in [26, 27] exploits the principal bundle structure of $SO(3) \times \mathbb{R}^2$ instead of $SO(3) \times \mathbb{R}^2 \times M$, and so it differs slightly from the one presented in this work.

### 2.3 Generalized Lagrange–d’Alembert–Poincaré Equations

**Local version.** It is easy to show that a curve $\gamma$ is a trajectory of $(L, C_K, C_V)$ if and only if for every local chart $(U, \varphi)$ of $Q$ the curve $q(t) = (\varphi \circ \gamma)(t)$ satisfies (on the open interval where the composition is defined)

$$(q(t), \dot{q}(t)) \in \varphi_*(C_K \cap TU)$$

and

$$\left( \frac{d}{dt} \left( \frac{\partial (L \circ \varphi^{-1})}{\partial \dot{q}} (q(t), \dot{q}(t)) \right) - \frac{\partial (L \circ \varphi^{-1})}{\partial q} (q(t), \dot{q}(t)) \right) \delta q^i(t) = 0, \quad (17)$$

for all curves $(q(t), \delta q(t))$ such that

$$(q(t), \delta q(t)) \in \varphi_*(C_V \cap TU).$$

Equation (17) can be called Generalized Lagrange–d’Alembert–Poincaré Equations for our GNHS. Through the choice of a connection on $Q$, we can write a global version of it.

**Vector bundles and related affine connections.** Let $M$ be a manifold and let us consider a vector bundle $\Pi: U \to M$ with projection $\Pi$. Fix an affine connection

$$\nabla: \mathfrak{X}(M) \times \Gamma(U) \to \Gamma(U).$$

This induces a vector bundle isomorphism

$$\beta: TU \to U \oplus TM \oplus U,$$

where $\oplus$ denotes the Whitney sum, defined as follows. Let $V \in TU$, and write $X = \tau_{U}(V)$ and $Y = \Pi_{\ast}(V)$, where $\tau_{U}: TU \to U$ is the canonical projection and $\Pi_{\ast}: TU \to TM$ is the differential of $\Pi$. Consider a curve $W: (-\varepsilon, \varepsilon) \to U$, passing through $X$ at $0$, with tangent vector $V$ at $X$. Thus,

- $W(0) = X$,
\[\begin{align*}
\bullet (\Pi \circ W)_* (d/dt|_0) &= Y, \\
\bullet W_*(d/dt|_0) &= V.
\end{align*}\]

Now define
\[\beta (V) = X \oplus Y \oplus \nabla Y W.\] (18)

In other words,
\[\beta (V) = \tau_U (V) \oplus \Pi_* (V) \oplus \nabla_{\Pi_* (V)} W.\]

For writing the inverse
\[\beta^{-1}: U \oplus TM \oplus U \rightarrow T\mathcal{U},\]
given \(X \oplus Y \oplus Z\) in \(U \oplus TM \oplus U\), let us fix \(W: (-\varepsilon, \varepsilon) \rightarrow \mathcal{U}\) such that
\[\bullet W(0) = X;\]
\[\bullet (\Pi \circ W)_* (d/dt|_0) = Y;\]
\[\bullet \nabla Y W = Z;\]
and define
\[\beta^{-1} (X \oplus Y \oplus Z) = W_* (d/dt|_0).\]

It can be shown that the definitions of \(\beta\) and \(\beta^{-1}\) are independent of \(W\).

For each fixed vector \(X \in \mathcal{U}\), we have applications
\[\begin{align*}
\beta_X: T_X\mathcal{U} &\rightarrow TM \oplus U : V \mapsto Y \oplus \nabla Y W \\
\beta^{-1}_X: TM \oplus U &\rightarrow T_X\mathcal{U} : Y \oplus Z \mapsto W_* (d/dt|_0),
\end{align*}\]
and their transposes
\[\begin{align*}
\beta^*_X: T^* M \oplus U^* &\rightarrow T^*_X \mathcal{U}, \\
\beta^{-1}_X: T^*_X U &\rightarrow T^* M \oplus U^*.
\end{align*}\]

**Definition 3.** The fiber derivative of \(f: \mathcal{U} \rightarrow \mathbb{R}\) is given by an application \(Ff: \mathcal{U} \rightarrow U^*\) such that
\[\langle Ff (X), Z \rangle = \frac{df (X + s Z)}{ds} \bigg|_{s=0},\] (19)
and the base derivative \(Bf: \mathcal{U} \rightarrow T^* M\) by
\[\langle Bf (X), Y \rangle = \frac{df (W(s))}{ds} \bigg|_{s=0},\] (20)

where \(W\) is a curve such that
\[W(0) = X, \quad (\Pi \circ W)_* (d/dt|_0) = Y \quad \text{and} \quad \nabla_Y W = 0.\]

Note that \(Ff\) is independent of \(\nabla\), but \(Bf\) is not. The next result follows immediately.

**Lemma 4.** Given \(f: \mathcal{U} \rightarrow \mathbb{R}\) and \(X \in \mathcal{U}\) we have
\[\beta^{-1}_X (df (X)) = Bf (X) \oplus Ff (X).\] (21)
Global equations. Given an affine connection $\nabla$ on $Q$, we can define a linear bundle isomorphism $\beta: TTQ \rightarrow TQ \oplus TQ \oplus TQ$ as above, by taking $U = TQ$ and $\Pi = \tau_Q$. Then, for the velocity $\delta \gamma'$ of a variation we have

$$\beta(\delta \gamma'(t)) = \delta \gamma(t) \oplus \gamma'(t) \oplus \nabla_{\gamma} \delta \gamma(t)$$

$$= \delta \gamma(t) \oplus \gamma'(t) \oplus \frac{D}{Dt} \delta \gamma(t).$$

(22)

Assuming from now on that $\nabla$ is torsion-free, it can be easily proved that $\kappa_{\beta} = \beta \circ \kappa \circ \beta^{-1}: TQ \oplus TQ \oplus TQ \rightarrow TQ \oplus TQ \oplus TQ$ is given by

$$\kappa_{\beta}(X \oplus Y \oplus Z) = Y \oplus X \oplus Z.$$  (23)

Now, we are in a position to build up a global expression for (17).

Theorem 5. A curve $\gamma: [t_1, t_2] \rightarrow Q$ is a trajectory of the triple $(L, C_K, C_V)$ if and only if $\gamma'(t) \in C_K$ and

$$\left< \frac{D}{Dt} F_L(\gamma'(t)) + B_L(\gamma'(t)), \delta \gamma(t) \right> = 0, \quad \forall t \in (t_1, t_2),$$

(24)

for all variations $\delta \gamma$ in $C_V$. This equation can also be written as

$$\frac{D}{Dt} F_L(\gamma'(t)) + B_L(\gamma'(t)) \in F_V|_{\gamma(t)}, \quad \forall t \in (t_1, t_2),$$

(25)

recalling that $F_V|_{\gamma(t)} = \left(C_V|_{\gamma(t)}\right)^{\circ}$.

Proof. From (6), introducing the identity map $\beta^{-1} \circ \beta$ into the second argument of the pairing, we have

$$\int_{t_1}^{t_2} \left< dL(\gamma'(t)), \frac{D}{Dt} \delta \gamma(t) \right> dt = 0,$$

for all $\delta \gamma$ such that $\delta \gamma(t) \in C_V$. Then, transposing $\beta^{-1}_{\gamma(t)}$ and using (21), (22) and (23),

$$\left< dL(\gamma'(t)), \beta^{-1}_{\gamma(t)} \circ \kappa_{\beta} \circ \beta(\delta \gamma'(t)) \right> = \left< \beta^{-1}_{\gamma(t)}(dL(\gamma'(t))), \delta \gamma(t) \oplus \frac{D}{Dt} \delta \gamma(t) \right>$$

$$= \left< B_L(\gamma'(t)) \oplus F_L(\gamma'(t)), \delta \gamma(t) \oplus \frac{D}{Dt} \delta \gamma(t) \right>$$

$$= \left< B_L(\gamma'(t)), \delta \gamma(t) \right> + \left< F_L(\gamma'(t)), \frac{D}{Dt} \delta \gamma(t) \right>.$$
On the other hand, by definition of the covariant derivative in $T^*Q$, 
\[
\frac{d}{dt} \langle FL(\gamma'), \delta \gamma \rangle = \langle FL(\gamma'), D_t \delta \gamma \rangle + \langle D_t FL(\gamma'), \delta \gamma \rangle,
\]
and, since $\delta \gamma(t_1)$ and $\delta \gamma(t_2)$ belong to the zero section of $TQ$,
\[
\int_{t_1}^{t_2} \frac{d}{dt} \langle FL(\gamma'), \delta \gamma \rangle \, dt = \langle FL(\gamma'(t_1)), \delta \gamma(t_1) \rangle - \langle FL(\gamma'(t_2)), \delta \gamma(t_2) \rangle = 0.
\]
As a consequence, 
\[
\int_{t_1}^{t_2} \left( -\frac{D}{Dt} FL(\gamma'(t)) + BL(\gamma'(t)), \delta \gamma(t) \right) dt = 0
\]
for all variations $\delta \gamma$ in $C_V$, which clearly implies Eqs. (24) and (25).

3 Lagrangian reduction of GNHSs

Consider a GNHS $(L, C_K, C_V)$. Suppose that a Lie group $G$ acts on $Q$, with action $\rho: G \times Q \to Q$, in such a way that the triple $(L, C_K, C_V)$ is $G$-invariant, that is:

a. $L \circ (\rho_g)_* = L$,

b. $(\rho_g)_*(C_K) \subset C_K$ and $(\rho_g)_*(C_V) \subset C_V$.

for all $g \in G$, with $\rho_g: Q \to Q: q \mapsto \rho(g, q)$.

From the canonical projection $p: TQ \to TQ/G$, related to the lifted action $\tilde{\rho}: G \times TQ \to TQ: (g,v) \mapsto (\rho_g)_*(v)$,

we can define the reduced Lagrangian $l: TQ/G \to \mathbb{R}$ by the formula

\[
l \circ p = L,
\]

and the reduced constraints

\[
\mathcal{C}_K = p(C_K) = C_K/G \quad \text{and} \quad \mathcal{C}_V = p(C_V) = C_V/G.
\]

We shall assume that $TQ/G$ is a manifold and $p$ a submersion.

The aim of this section is to write the equations of motion of $(L, C_K, C_V)$ in terms of the reduced data $l$, $\mathcal{C}_K$ and $\mathcal{C}_V$. Moreover, by introducing an appropriate principal connection, we are going to separate the reduced virtual displacements $\mathcal{C}_V$ into horizontal and vertical components, and construct in this way the Horizontal and Vertical Generalized Lagrange–d’Alembert–Poincaré Equations as in [14].
3.1 Generalized nonholonomic connection

From now on, we will write $\mathcal{X} = Q/G$, and assume that the canonical projection $\pi: Q \to \mathcal{X}$ is a principal fiber bundle with structure group $G$. Let us denote the vertical distribution by $\mathcal{V}$, that is, $\mathcal{V} = \ker(\pi_\ast) \subset TQ$.

Now, we shall construct a principal connection related to $(L, C_K, C_V)$ and the group $G$. We shall proceed in several steps.

1. Fix a $G$-invariant metric on $Q$. For instance, if $L$ is simple, (that is, if it is of the form kinetic minus potential energy, as explained above) we can use the metric given by its kinetic term.

2. Consider the intersection
   \[ S = C_V \cap \mathcal{V}, \quad (28) \]
   and write
   \[ C_V = T \oplus S \quad \text{and} \quad \mathcal{V} = S \oplus U, \quad (29) \]
   where $T$ and $U$ are the orthogonal complements of $S$ in $C_V$ and $\mathcal{V}$, respectively.

3. Consider the orthogonal complement of $C_V + \mathcal{V}$ in $TQ$. Let us call it $R$.

4. Define the principal connection form $A: TQ \to \mathfrak{g}$, which we shall call the generalized nonholonomic connection, with horizontal distribution $H = R \oplus T$. Observe that
   \[ T = C_V \cap H. \quad (30) \]

Summing up, we have the Whitney sum
\[ TQ = H \oplus \mathcal{V}, \quad \text{with} \quad H = R \oplus T \quad \text{and} \quad \mathcal{V} = S \oplus U. \quad (31) \]

**Remark 3.** Let us note that connection $A$ has been built up from the variational constraints $C_V$ only. The kinematic constraints $C_K$ were not involved. Also note that some sums in (31) are not necessarily orthogonal.

All the (generalized) distributions $R$, $S$, $T$ and $U$ are $G$-invariant. Therefore we can write
\[ TQ/G = R/G \oplus T/G \oplus S/G \oplus U/G, \]
and in particular [see the first parts of (27) and (29)]
\[ C_V = T/G \oplus S/G. \quad (32) \]
The identification $\alpha_A$. Let $\mathfrak{g}$ be the Lie algebra of $G$ and $\tilde{\mathfrak{g}} = Q \times_G \mathfrak{g}$ the associated bundle over $\mathcal{X} = Q/G$. The elements of $\tilde{\mathfrak{g}}$ will be denoted as equivalence classes $[q, \eta]$, with $q \in Q$ and $\eta \in \mathfrak{g}$.

Given a principal connection $A$, we can construct a fiber bundle isomorphism

$$\alpha_A: TQ/G \to T\mathcal{X} \oplus \tilde{\mathfrak{g}}$$

(see Ref. [15]), such that, for all $q \in Q$ and $v_q \in T_qQ$,

$$[v_q] \mapsto \pi^* (v_q) \oplus [q, A(v_q)].$$

Here, $[v_q] = p(v_q) \in TQ/G$. Denoting by $a$ the application

$$a: TQ \to \tilde{\mathfrak{g}}: v_q \mapsto [q, A(v_q)],$$

we have

$$\alpha_A \circ p(v) = \pi^* (v) \oplus a(v), \quad \forall v \in TQ.$$  \hfill (34)

When there is no risk of confusion, we shall identify the fiber bundles $TQ/G$ and $T\mathcal{X} \oplus \tilde{\mathfrak{g}}$ via the application $\alpha_A$. Of course, the connection involved will be the generalized nonholonomic connection, defined above.

Decomposition of reduced virtual displacements. In terms of the identification $\alpha_A$, we have

$$\mathcal{H}/G = \alpha_A (\mathcal{H}/G) = \pi^* (\mathcal{H}) = T\mathcal{X}$$

and

$$\mathcal{V}/G = \alpha_A (\mathcal{V}/G) = a(\mathcal{V}) = \tilde{\mathfrak{g}}.$$ 

Defining the subbundles $\tilde{s}$ and $\tilde{u}$ of $\tilde{\mathfrak{g}}$ by

$$\tilde{s} = \alpha_A (S/G) \quad \text{and} \quad \tilde{u} = \alpha_A (U/G),$$

we have $\tilde{\mathfrak{g}} = \tilde{s} \oplus \tilde{u}$. In the same way, defining

$$\tau = \alpha_A (R/G) \quad \text{and} \quad t = \alpha_A (T/G),$$

it follows that $T\mathcal{X} = \tau \oplus t$. With all this, we can write the following proposition.

**Proposition 6.** $\mathfrak{C}_V$ can be decomposed as

$$\mathfrak{C}_V = \mathfrak{C}^{hor}_V \oplus \mathfrak{C}^{ver}_V,$$

where [see Eq. (32)]

$$\mathfrak{C}^{hor}_V = \pi^* (C_V) = T\mathcal{X} \cap \mathfrak{C}_V = t$$  \hfill (35)

and

$$\mathfrak{C}^{ver}_V = a (C_V) = \tilde{\mathfrak{g}} \cap \mathfrak{C}_V = \tilde{s}.$$  \hfill (36)
3.2 The application $\beta$ for the bundle $TX \oplus \tilde{g}$

In order to write down the Generalized Lagrange–d'Alembert–Poincaré Equations we need the application $\beta$ for $TX \oplus \tilde{g}$. Let us build it up and study some of its properties.

Consider an affine connection $\nabla_X$ on $\mathcal{X}$ and the usual one $\nabla^A$ on $\tilde{g}$. Defining $\nabla = \nabla_X \oplus \nabla^A$ on $TX \oplus \tilde{g}$, we can write the linear bundle isomorphism

$$\beta: (TX \oplus \tilde{g}) \to (TX \oplus \tilde{g}) \oplus T\mathcal{X} \oplus (TX \oplus \tilde{g})$$

as that given in Eq. (18). In this case $\Pi$: $TX \oplus \tilde{g} \to \mathcal{X}$ is given by the canonical projections $\tilde{\pi}: \tilde{g} \to \mathcal{X}$ and $\tau_\mathcal{X}: T\mathcal{X} \to \mathcal{X}$ as

$$\Pi(\nu \oplus \eta) = \tilde{\pi}(\eta) = \tau_\mathcal{X}(\nu).$$

We shall write

$$\tilde{\tau} \equiv \tau_{TX \oplus \tilde{g}}: T(TX \oplus \tilde{g}) \to T\mathcal{X} \oplus \tilde{g}. \quad (37)$$

For later convenience, let us remark that we have the identities

$$\Pi \circ \alpha_A \circ p = \pi \circ \tau_Q \quad \text{and} \quad \tilde{\tau} \circ \alpha_A \circ p_\ast = \alpha_A \circ p \circ \tau_{TQ}. \quad (38)$$

In these terms $\beta$ is given, for all $\Xi \in T(TX \oplus \tilde{g})$, by

$$\beta(\Xi) = \tilde{\tau}(\Xi) \oplus \Pi_\ast(\Xi) \oplus \nabla_{\Pi_\ast(\Xi)} \varpi,$$

with $\varpi: (-\varepsilon, \varepsilon) \to TX \oplus \tilde{g}$ satisfying

$$\varpi(0) = \tilde{\tau}(\Xi) \quad \text{and} \quad \varpi_\ast(\frac{d}{dt}\big|_0) = \Xi.$$

Note that, for a curve $\varpi: (-\varepsilon, \varepsilon) \to TX \oplus \tilde{g}$ [compare with (22)],

$$\beta(\varpi'(t)) = \varpi(t) \oplus \dot{\varpi}'(t) \oplus \nabla_{\dot{\varpi}'} \varpi(t) = \varpi(t) \oplus \dot{\varpi}'(t) \oplus \frac{D}{Dt} \varpi(t), \quad (39)$$

where $\dot{\varpi} = \Pi \circ \varpi$.

For each $\mu \in TX \oplus \tilde{g}$ we have an application

$$\beta_\mu: T_\mu(TX \oplus \tilde{g}) \to TX \oplus (TX \oplus \tilde{g}) : \Xi \mapsto \Pi_\ast(\Xi) \oplus \nabla_{\Pi_\ast(\Xi)} \varpi$$

for all $\Xi$ such that $\tilde{\tau}(\Xi) = \mu$, and its corresponding transpose

$$\beta^\ast_\mu: T^\ast \mathcal{X} \oplus (T^\ast \mathcal{X} \oplus \tilde{g}^\ast) \to T^\ast_\mu(TX \oplus \tilde{g}^\ast).$$

Besides, for any function $f: TX \oplus \tilde{g} \to \mathbb{R}$ we have the identity

$$\beta_\mu^{-1}(df(\mu)) = Bf(\mu) \oplus F_X f(\mu) \oplus F_\tilde{g} f(\mu), \quad (40)$$

where

$$F_X f \oplus F_\tilde{g} f = \mathcal{B} f: TX \oplus \tilde{g} \to T^\ast \mathcal{X} \oplus \tilde{g}^\ast$$

is the fiber derivative of $f$, and

$$\mathcal{B} f: TX \oplus \tilde{g} \to T^\ast \mathcal{X}$$

is its base derivative.

---

4 Here $\alpha_A$ means the differential of $\alpha_A$. Do not confuse with the notation of [28], where $\alpha_A$ is a vector bundle isomorphism between $T^\ast Q/G$ and $T^\ast X \oplus \tilde{g}^\ast$, obtained from $\alpha_A$ by duality.
3.3 Generalized Lagrange–d’Alembert–Poincaré Equations

As we have said at the beginning of this section, we want to find the equations of motion of a GNHS \((L, C_K, C_V)\) in terms of its reduced data \(l, C_K\) and \(C_V\).

By definition, a curve \(\gamma: [t_1, t_2] \rightarrow Q\) is a trajectory of that system if

\[
\int_{t_1}^{t_2} \langle dL(\gamma'(t)), \kappa(\delta\gamma'(t)) \rangle \, dt = 0
\]

for all \(\delta\gamma\) such that \(\delta\gamma(t) \in CV.\) Since \(L = l \circ p [\text{see (26)}]\), then

\[
\langle dL(\gamma'(t)), \kappa(\delta\gamma'(t)) \rangle = \langle (p^* (dl(\gamma'(t))), \kappa(\delta\gamma'(t))) \rangle = \langle dl(\gamma'(t)), p_* \circ \kappa(\delta\gamma'(t)) \rangle.
\]

Introducing the identity map \(\alpha_A^{-1} \circ \beta^{-1} \circ \beta \circ \alpha_A\) into the second argument of the pairing and using the transpose of \(\beta\) and \(\alpha_A\), it follows that

\[
\langle dL(\gamma'(t)), \kappa(\delta\gamma'(t)) \rangle = \langle (\beta_{\mu(t)}^{-1} \circ \alpha_A^{-1} (dl(\mu(t))), (\alpha_A \circ p)_* \circ \kappa(\delta\gamma'(t))) \rangle,
\]

where \(\mu(t) = \alpha_A \circ p (\gamma'(t))\). Using the identification of \(TQ/G\) and \(TX \oplus \tilde{g}\), we denote the composition

\[
l \circ \alpha_A: TX \oplus \tilde{g} \rightarrow \mathbb{R},
\]

simply as \(l\). Then we have

\[
\langle dL(\gamma'(t)), \kappa(\delta\gamma'(t)) \rangle = \langle \beta_{\mu(t)}^{-1} (dl(\mu(t))), (\alpha_A \circ p)_* \circ \kappa(\delta\gamma'(t)) \rangle.
\]

Reduced variations. Let us calculate \(\beta((\alpha_A \circ p)_* \circ \kappa(\delta\gamma'(t)))\). From now on, we shall omit \(t\) in some expressions, just for brevity.

**Proposition 7.** For any curve \(\gamma: [t_1, t_2] \rightarrow Q\) and any variation \(\delta\gamma\), we have

\[
\Pi_* ((\alpha_A \circ p)_* \circ \kappa(\delta\gamma)) = \pi_* (\delta\gamma)
\]

and

\[
\tilde{\tau} ((\alpha_A \circ p)_* \circ \kappa(\delta\gamma)) = \pi_* (\gamma') \oplus a(\gamma').
\]

**Proof.** It is enough to use Eq. (38) to show that

\[
\Pi_* ((\alpha_A \circ p)_* \circ \kappa(\delta\gamma')) = \pi_* ([\tau_{\pi}]_* \circ \kappa(\delta\gamma'))
\]

and

\[
\tilde{\tau} ((\alpha_A \circ p)_* \circ \kappa(\delta\gamma')) = \alpha_A \circ p [\tau_{\pi} \circ \kappa(\delta\gamma')];
\]

and then use Eqs. (4), (10) and (11) [also taking into account (34)] to achieve the desired result. \(\square\)
Theorem 8. For any curve $\gamma: [t_1, t_2] \to Q$ and any variation $\delta\gamma$,

$$\beta[(\alpha_A \circ p)_{\ast} \circ \kappa(\delta\gamma')] = \pi_{\ast}(\gamma') \oplus a(\gamma') \oplus \pi_{\ast}(\delta\gamma) \oplus \frac{D}{D t}[\pi_{\ast}(\delta\gamma)] \oplus \left(\frac{D}{D t}[a(\delta\gamma)] + [a(\gamma'), a(\delta\gamma)] - b(\gamma', \delta\gamma)\right)$$

where $b$ is the map

$$b: TQ \times_Q TQ \to \tilde{\mathfrak{g}} : (v_q, w_q) \mapsto [q, B(v_q, w_q)]$$

and $B: TQ \times_Q TQ \to \mathfrak{g}$ is the curvature of the connection $A$.

Proof. The first three terms follow immediately from the proposition above. For the last two, let us write

$$\kappa(\delta\gamma'(t)) \in T\gamma'(t)TQ$$

for each $t \in (t_1, t_2)$, in terms of a suitable deformation $\Delta\gamma: [t_1, t_2] \times (-\epsilon, \epsilon) \to Q$,

$$\kappa(\delta\gamma'(t)) = \Delta\gamma_{\ast}\left(\frac{\partial}{\partial t}\big|_{(t, \lambda)}\right)_{\ast}(d/d\lambda|_0).$$

Then

$$\beta[(\alpha_A \circ p)_{\ast} \circ \kappa(\delta\gamma'(t))] = \beta(\varpi_t(0)) = \varpi_t(0) \oplus \pi_{\ast}(\delta\gamma(t)) \oplus \frac{D}{D \lambda}\varpi_t(0),$$

where $\varpi_t = \alpha_A \circ p \left(\Delta\gamma_{\ast}\left(\frac{\partial}{\partial t}\big|_{(t, \lambda)}\right)\right): (-\epsilon, \epsilon) \to TQ \oplus \tilde{\mathfrak{g}}$

is a family of curves. (Let us keep in mind that the prime on $\varpi_t$ means derivative w.r.t. $\lambda$.) From Eq. (39) for $\varpi_t$,

$$\beta[(\alpha_A \circ p)_{\ast} \circ \kappa(\delta\gamma'(t))] = \beta(\varpi_t(0)) = \varpi_t(0) \oplus \pi_{\ast}(\delta\gamma(t)) \oplus \frac{D}{D \lambda}\varpi_t(0),$$

where $\pi_{\ast} = \Pi \circ \varpi_t$. We already know that

$$\varpi_t(0) = \pi_{\ast}(\gamma'(t)) \oplus a(\gamma'(t)) \quad \text{and} \quad \pi_{\ast}(\delta\gamma(t)),$$

so we just must show that

$$\frac{D}{D \lambda}\varpi_t(0) = \frac{D}{D t}\pi_{\ast}(\delta\gamma(t)) \oplus \left(\frac{D}{D t}[a(\delta\gamma(t))] + [a(\gamma'(t)), a(\delta\gamma(t))] - b(\gamma'(t), \delta\gamma(t))\right).$$

Since

$$\varpi_t = \pi_{\ast}\left[\Delta\gamma_{\ast}\left(\frac{\partial}{\partial t}\big|_{(t, \lambda)}\right)\right] \oplus a\left(\Delta\gamma_{\ast}\left(\frac{\partial}{\partial t}\big|_{(t, \lambda)}\right)\right),$$
and we are considering a connection of the form $\nabla_X \oplus \nabla^A$, then

$$\frac{D}{D\lambda} \pi_t(0) = \frac{D}{D\lambda} \left|_0 \left[ (\pi \circ \Delta)_* \left( \frac{\partial}{\partial t}_{|t,\cdot} \right) \right] \right| \oplus \frac{D}{D\lambda} \left|_0 \left[ \Delta_{\pi*} \left( \frac{\partial}{\partial \lambda}_{|t,0} \right) \right] \right|,$$  \hspace{1cm} (45)

It is easy to show that

$$\frac{D}{D\lambda} \left|_0 \left[ (\pi \circ \Delta)_* \left( \frac{\partial}{\partial t}_{|t,\cdot} \right) \right] \right| = \frac{D}{Dt} \left| \left[ \pi \circ \Delta \right] \left( \frac{\partial}{\partial \lambda}_{|t,0} \right) \right|,$$

which implies that the first term of (45) is exactly

$$\frac{D}{Dt} \pi_* (\delta \gamma (t)).$$

On the other hand, since

$$\left[ \pi \circ \Delta \right] \left( \frac{\partial}{\partial t}_{|t,\cdot} \right) = \left[ \pi \circ \Delta \right] \left( \frac{\partial}{\partial \lambda}_{|t,0} \right),$$

we can show from Lemma 2.3.4 of [15] that the second term of (45) is equal to

$$\left[ \gamma(t), [A(\gamma'(t)), A(\delta \gamma(t))] \right] + \frac{D}{D\lambda} \left|_0 \left[ A \left( \Delta_{\pi*} \left( \frac{\partial}{\partial t}_{|t,\cdot} \right) \right) \right] \right|.$$  \hspace{1cm} (46)

Furthermore, using Lemmas 3.1.4 and 3.1.5 it follows that

$$\left[ \gamma(t), \frac{d}{d\lambda} A \left( \Delta_{\pi*} \left( \frac{\partial}{\partial t}_{|t,\cdot} \right) \right) \right] = \frac{D}{Dt} \left[ \gamma(t), A(\delta \gamma(t)) \right] - \left[ \gamma(t), B(\gamma'(t), \delta \gamma(t)) \right].$$  \hspace{1cm} (47)

Finally, combining (46) and (47) we have

$$\frac{D}{D\lambda} \left|_0 \left[ A \left( \Delta_{\pi*} \left( \frac{\partial}{\partial t}_{|t,\cdot} \right) \right) \right] \right| = \frac{D}{Dt} \left[ \gamma(t), A(\delta \gamma(t)) \right] - \left[ \gamma(t), B(\gamma'(t), \delta \gamma(t)) \right].$$

that is to say

$$\frac{D}{D\lambda} \left|_0 \left[ A \left( \Delta_{\pi*} \left( \frac{\partial}{\partial t}_{|t,\cdot} \right) \right) \right] \right| = \frac{D}{Dt} a(\delta \gamma(t)) + [a(\gamma'(t)), a(\delta \gamma(t))] - b(\gamma'(t), \delta \gamma(t)).$$

what completes the proof of our theorem.

Let us translate the result above into the terminology of [15]. For a curve $\gamma$, and a variation $\delta \gamma$, let us write

$$\pi \circ \gamma = x, \quad \alpha_A \circ p \circ \gamma' = \mu = \dot{x} \oplus \pi \quad \text{and} \quad \alpha_A \circ p \circ \delta \gamma = \delta x \oplus \eta,$$

where

$$\dot{x} = \pi_* \circ \gamma', \quad \pi = a \circ \gamma', \quad \delta x = \pi_* \circ \delta \gamma \quad \text{and} \quad \eta = a \circ \delta \gamma.$$
It is easy to see that \( b(\gamma', \delta \gamma) \) depends only on \( \dot{x} \) and \( \delta x \), so we can define
\[
\tilde{B}: TX \times TX \to \tilde{g} : \ (\pi_*(u), \pi_*(v)) \mapsto b(u, v).
\]
With all that,\[
\beta [(\alpha_A \circ p)_* \circ \kappa (\delta \gamma')] = \dot{x} \oplus \bar{v} \oplus \delta x \oplus D \frac{D}{Dt} \delta x \oplus \left( \frac{D}{Dt} \bar{\eta} + [\bar{v}, \bar{\eta}] - \tilde{B}(\dot{x}, \delta x) \right).
\]

**Reduced equations.** Returning to (42), we have that \( \langle dL(\gamma'), \kappa (\delta \gamma') \rangle \) is equal to\[
\left\langle \beta^{-1}_{\mu(t)} (dl(\mu(t))), \delta x \oplus \frac{D}{Dt} \delta x \oplus \left( \frac{D}{Dt} \bar{\eta} + [\bar{v}, \bar{\eta}] - \tilde{B}(\dot{x}, \delta x) \right) \right\rangle.
\]
On the other hand, from (40) we know that \( \beta^{-1}_{\mu(t)} (dl(\mu(t))) = B \mathcal{L}(\mu(t)) \oplus F_{\mathcal{L}}(\mu(t)) \oplus F_{\tilde{g}}(\mu(t)) \), or with the notation of [15], i.e., writing
\[
\mathcal{B}l = \frac{\partial l}{\partial x}, \quad F_{\mathcal{L}}l = \frac{\partial l}{\partial \dot{x}} \quad \text{and} \quad F_{\tilde{g}}l = \frac{\partial l}{\partial \bar{v}},
\]
we have
\[
\beta^{-1}_{\mu(t)} (dl(\mu(t))) = \frac{\partial l}{\partial x}(\mu(t)) \oplus \frac{\partial l}{\partial \dot{x}}(\mu(t)) \oplus \frac{\partial l}{\partial \bar{v}}(\mu(t)).
\]
Accordingly, \( \langle dL(\gamma'), \kappa (\delta \gamma') \rangle \) adopts the form (omitting \( \mu \))
\[
\left\langle \frac{\partial l}{\partial x}, \delta x \right\rangle + \left\langle \frac{\partial l}{\partial \dot{x}}, \frac{D}{Dt} \delta x \right\rangle + \left\langle \frac{\partial l}{\partial \bar{v}}, \frac{D}{Dt} \bar{\eta} + [\bar{v}, \bar{\eta}] - \tilde{B}(\dot{x}, \delta x) \right\rangle;
\]
and integrating by parts (and assorting terms with \( \delta x \) and \( \bar{\eta} \))
\[
\left\langle - \frac{D}{Dt} \frac{\partial l}{\partial x} + \frac{\partial l}{\partial \dot{x}} \dot{\delta x} \right\rangle + \left\langle \frac{\partial l}{\partial \bar{v}}, \bar{\delta x} \right\rangle + \left\langle - \frac{D}{Dt} \frac{\partial l}{\partial \bar{v}} + \text{ad}_v^* \frac{\partial l}{\partial \bar{v}}, \bar{\eta} \right\rangle = 0.
\]
We can now easily prove the following result.

**Theorem 9.** Let \((L, C_K, C_V)\) be a GNHS and \(G\) a Lie group acting on \(Q\). Suppose that the system is \(G\)-invariant and that \(\pi: Q \to \mathcal{X} = Q/G\) is a principal fiber bundle with generalized nonholonomic connection \(A: TQ \to \mathfrak{g}\). Let \(\gamma: [t_1, t_2] \to Q\) be a curve on \(Q\). Then, the following statements are equivalent:

1. the curve \(\gamma\) satisfies
   \[
   \gamma'(t) \in C_K
   \]
   and
   \[
   \left\langle - \frac{D}{Dt} \mathcal{F}L(\gamma'(t)) + \mathcal{B}L(\gamma'(t)), \delta \gamma(t) \right\rangle = 0,
   \]
   for all \(\delta \gamma: [t_1, t_2] \to TQ\) such that
   \[
   \delta \gamma(t) \in C_V|_{\gamma(t)}.
   \]

19
Lemma 10. Consider a curve \( \gamma \colon [t_1, t_2] \to TX \oplus \mathfrak{g} \), given by
\[
\gamma(t) = x(t) \oplus \eta(t) = A \circ p(\gamma'(t)),
\]
satisfies
\[\gamma(t) \in \mathcal{C}_K,\]
the Horizontal Generalized Lagrange–d’Alembert–Poincaré Equations
\[
\left\langle - \frac{D}{Dt} \frac{\partial}{\partial x} (\mu(t)) + \frac{\partial}{\partial x} (\mu(t)) - \left\langle \frac{\partial}{\partial v}(\mu(t)), \dot{\gamma}(t) \right\rangle, \delta x(t) \right\rangle = 0 \quad (48)
\]
and the Vertical Generalized Lagrange–d’Alembert–Poincaré Equations
\[
\left\langle - \frac{D}{Dt} \frac{\partial}{\partial v} (\mu(t)) + \text{ad}_v^* \frac{\partial}{\partial v} (\mu(t)), \dot{\eta}(t) \right\rangle = 0, \quad (49)
\]
for all curves \( \delta x : [t_1, t_2] \to TX \) and \( \dot{\eta} : [t_1, t_2] \to \mathfrak{g} \)
satisfying [see Eqs. (35) and (36)]
\[
\delta x(t) \in \mathcal{C}^{\text{hor}}_V_{\mid x(t)} \quad \text{and} \quad \dot{\eta}(t) \in \mathcal{C}^{\text{ver}}_V_{\mid x(t)}.
\]

In order to prove the theorem, we only need the following Lemma.

Lemma 10. Consider a curve \( \gamma : [t_1, t_2] \to Q \), and its projection
\[
x : [t_1, t_2] \to X : \quad t \mapsto x(t) = \pi(\gamma(t)).
\]
Given curves \( \delta x : [t_1, t_2] \to TX \) and \( \dot{\eta} : [t_1, t_2] \to \mathfrak{g} \), we have that
\[
\delta x(t) \in \mathcal{C}^{\text{hor}}_V_{\mid x(t)} \quad \text{and} \quad \dot{\eta}(t) \in \mathcal{C}^{\text{ver}}_V_{\mid x(t)} \quad (50)
\]
if and only if there exists a curve \( \delta \gamma : [t_1, t_2] \to TQ \) satisfying
\[
\pi_* (\delta \gamma(t)) = \delta x(t) \quad \text{and} \quad A (\delta \gamma(t)) = \dot{\eta}(t), \quad (51)
\]
and such that \( \delta \gamma(t) \in \mathcal{C}_V_{\mid x(t)}. \)

Proof. Given a curve \( \delta \gamma \) such that \( \delta \gamma(t) \in \mathcal{C}_V_{\mid x(t)} \), it is clear that the curves \( \delta x(t) = \pi_* (\delta \gamma(t)) \) and \( \dot{\eta}(t) = A (\delta \gamma(t)) \) satisfy (50). Reciprocally, consider curves \( \delta x \) and \( \dot{\eta} \) fulfilling (50). Let us write
\[
\dot{\eta}(t) = [\gamma(t), \eta(t)]
\]
for some curve \( \eta : [t_1, t_2] \to \mathfrak{g} \) (which is, in fact, unique). Recall that for each \( q \in Q \), \( \pi_* \circ (T_qQ \to T_{\pi(q)}X) \) gives rise to an isomorphism when restricted to the horizontal space \( \mathcal{H}_q \). Define
\[
\delta \gamma^h(t) = \left( \pi_* (\gamma(t)) \big|_{\mathcal{H}_q(t)} \right)^{-1} (\delta x(t)), \quad \delta \gamma^v(t) = \frac{d}{ds}_{s=0} \rho \left( \exp (s \eta(t)), \gamma(t) \right)
\]
and \[ \delta \gamma (t) = \delta \gamma^h (t) + \delta \gamma^v (t). \]

Note that \( \delta \gamma^h \) is the horizontal lift of \( \delta x \) along the curve \( \gamma \), and \( \delta \gamma^v (t) \) is the fundamental vector field corresponding to \( \eta (t) \in \mathfrak{g} \) at the point \( \gamma (t) \). Accordingly,
\[ \pi_* (\delta \gamma (t)) = \pi_* (\delta \gamma^h (t)) = \delta x (t) \]
and, since
\[ A (\delta \gamma^v (t)) = A \left( \frac{d}{ds} \bigg|_{s=0} \rho (\exp (s \eta (t)) \cdot \gamma (t)) \right) = \eta (t) \]
by definition of \( A \),
\[ a (\delta \gamma (t)) = [\gamma (t) \cdot A (\delta \gamma (t))] = [\gamma (t) \cdot A (\delta \gamma^v (t))] = [\gamma (t) \cdot \eta (t)] = \bar{\eta} (t). \]

We shall show that
\[ \delta \gamma^h (t) \in \mathcal{C}_V \cap \mathcal{H} \quad \text{and} \quad \delta \gamma^v (t) \in \mathcal{C}_V \cap \mathcal{V}, \]
what implies that \( \delta \gamma (t) \in \mathcal{C}_V \). Since
\[ \pi_{*, \gamma(t)} (\delta \gamma^h (t)) = \delta x (t) \in \mathcal{C}_V \bigg|_{x(t)} = \pi_{*, \gamma(t)} \left( \mathcal{C}_V \big|_{\gamma(t)} \right), \]
we have that
\[ \delta \gamma^h (t) \in \mathcal{C}_V \big|_{\gamma(t)} + \ker (\pi_{*, \gamma(t)}) = (\mathcal{C}_V + \mathcal{V}) \big|_{\gamma(t)} . \]

But \( \delta \gamma^h (t) \in \mathcal{H} \big|_{\gamma(t)} \), so
\[ \delta \gamma^h (t) \in (\mathcal{C}_V + \mathcal{V}) \cap \mathcal{H} = \mathcal{C}_V \cap \mathcal{H}. \]

In the same way (replacing \( \pi_* \) by \( a \)) it can be shown that
\[ \delta \gamma^v (t) \in (\mathcal{C}_V + \mathcal{H}) \cap \mathcal{V} = \mathcal{C}_V \cap \mathcal{V}, \]
finishing our proof. \( \square \)

4 Rolling ball on a moving platform: a stabilization problem

Let us consider a mechanical system consisting of two rigid bodies: a homogeneous ball of mass \( m \) and radius \( r \), and a platform with momentum of inertia \( I \); both subjected to the acceleration of gravity \( g \). Suppose that the center of mass of the platform is fixed (w.r.t. an inertial reference system) and that the platform can rotate freely around it. Suppose, in addition, that the ball rolls on the platform without sliding. A classical control problem consists in making
the ball converge to the center of mass of the platform, moving the platform by applying an adequate torque. Following with the applications of GNHSs to the control of servomechanisms, this stabilization can be achieved by means of an appropriate set of constraints, the constraint force playing the role of a control signal. The explicit form of these constraints will appear in a forthcoming paper. Since the described mechanical system is SO (3)-invariant, the stabilization problem can be solved by studying the related reduced equations, which in this case will be derived with the purpose of illustrating the general procedure presented in the previous section. We shall develop in detail all those steps that are strictly related to GNHSs, such as the construction of the generalized nonholonomic connection and the reduced variational constraints, omitting, for instance, the calculation of covariant derivatives and curvature tensor, which also appear in the reduction of standard Lagrangian systems [15].

4.1 Description of the system

We shall first describe the system formed by the ball and the platform (with the platform with its center of mass fixed and the ball rolling on the platform without sliding), and then add progressively the elements that will transform this system into a GNHS.

In order to construct the configuration space, let us take a spatial system $S$ of orthogonal axes and another one $S'$ fixed to the platform, both having the origin at the center of mass of the platform. Let us call $e_1, e_2, e_3$ and $E_1, E_2, E_3$ the vectors $(1, 0, 0)$, $(0, 1, 0)$ and $(0, 0, 1)$ corresponding to such orthogonal axes, respectively. We choose these axes so that $g$ is parallel to $e_3$, i.e., $g = g e_3$, and $E_3$ is perpendicular to the platform. It is clear that the position of the platform will be described by an element $C \in$ SO (3) representing its rotation w.r.t. $S$. The position of the ball can be described by the manifold $\mathbb{R}^2 \times$ SO (3), together with the position of the platform. The elements $a \in \mathbb{R}^2$ give the position of the point of contact between the ball and the platform, and the points $R \in$ SO (3) describe the rotation of the ball, both w.r.t. $S'$. Summing up, the configuration space is $Q = \mathbb{R}^2 \times$ SO (3) $\times$ SO (3), and its points will be indicated by $(a, C, R)$. In these terms, the no sliding condition reads

$$\dot{a} = r \dot{R} R^{-1} E_3.$$  \hspace{1cm} (52)

The Lagrangian of the system is

$$L = \frac{1}{2} \Omega \dot{\Omega} + \frac{1}{2} I \dot{\varpi}^2 + \frac{1}{2} m \dot{x}^2 - m \langle g, x \rangle,$$

$\Omega$ being the angular velocity of the platform expressed w.r.t. $S'$, $I = 2mr^2/5$ the momentum of inertia of the ball (which is homogeneous), $\varpi$ its angular velocity and $x$ the position of its center of mass. (By $\langle \cdot, \cdot \rangle$ we are denoting the canonical scalar product of $\mathbb{R}^3$.) The quantities $\varpi$ and $x$ are expressed in the axes $S$. 
In order to write down $\Omega$, $\varpi$ and $x$ as functions of $(a, C, R)$ and their derivatives $\left(\dot{a}, \dot{C}, \dot{R}\right)$, and arrive in this way to an expression of $L$ as a function of $(a, C, R, \dot{a}, \dot{C}, \dot{R}) \in TQ$, it will be useful to use the identification $\hat{\cdot} : \mathbb{R}^3 \to \mathfrak{so}(3)$ given by

$$x = (x^1, x^2, x^3) \mapsto \hat{x} = \begin{bmatrix} 0 & -x^3 & x^2 \\ x^3 & 0 & -x^1 \\ -x^2 & x^1 & 0 \end{bmatrix}.$$

It can be shown that

$$\hat{x} \times \hat{y} = [\hat{x}, \hat{y}], \quad \hat{x} y = x \times y, \quad \hat{R} x = R \hat{x} R^{-1}, \quad \langle x, y \rangle = -\frac{1}{2} \text{tr}(\hat{x} \hat{y}),$$

$\forall x, y \in \mathbb{R}^3$, $R \in \text{SO}(3)$. (The symbol $\times$ indicates the cross product in $\mathbb{R}^3$.) In these terms

$$\hat{\Omega} = C^{-1} \dot{C}, \quad \hat{\varpi} = \left(\dot{C} R + C \dot{R}\right) (CR)^{-1}, \quad x = C (a + r E_3),$$

and

$$\dot{x} = \dot{C} (a + r E_3) + C \dot{a}.$$

Then,

$$L \left( a, C, R, \dot{a}, \dot{C}, \dot{R} \right) = -\frac{1}{4} \text{tr} \left( I^{-1} C^{-1} \dot{C} I C^{-1} \dot{C} \right) - \frac{1}{4} \text{tr} \left( \left[ \left( \dot{C} R + C \dot{R}\right) (CR)^{-1} \right]^2 \right) + \frac{1}{2} m \left( \dot{C} (a + r E_3) + C \dot{a} \right)^2 - m \langle g, C (a + r E_3) \rangle.$$

By now, we have a Lagrangian system subject to nonholonomic constraints, given by (52). It is clear that $L$ and the constraints are $\text{SO}(3)$-invariant with respect to the action

$$\rho : Q \times \text{SO}(3) \to Q : ((a, C, R), A) \mapsto (a, C, RA).$$

Note that $Q$ is a trivial principal bundle with base $\mathbb{R}^2 \times \text{SO}(3)$ and structure group $\text{SO}(3)$, i.e.,

$$Q = \mathcal{X} \times \text{SO}(3) \quad \text{with} \quad \mathcal{X} = \mathbb{R}^2 \times \text{SO}(3).$$

As we have mentioned before, the idea is to impose other constraints on the system that ensure the desired stabilization. These constraints, which will be implemented by torques on the platform, can be determined later. Of course,
they must also be SO(3)-invariant. In other words, beside the non sliding condition we will have three constraints of the form
\[ H_i \left( a, C, \dot{a}, \dot{C}, \omega \right) = 0, \]
where
\[ \dot{\omega} = \dot{RR}^{-1}, \]
giving rise to the constraint submanifold
\[ C_K = \left\{ \left( a, C, R, \dot{a}, \dot{C}, \dot{R} \right) : \dot{a} = r\dot{RR}^{-1}E_3, h_i \left( a, C, \dot{a}, \dot{C}, \omega \right) = 0 \right\}. \] (54)

In order to specify \( C_V \), it is convenient to analyze the constraint forces. These live on a codistribution \( F_V \subset T^*Q = T^*\mathbb{R}^2 \times T^*SO(3) \times T^*SO(3) \) which is the sum of the codistribution \( F_{nsc}^V \) corresponding to the non sliding condition, and that corresponding to the torques on the platform
\[ F_{control}^V = \left\{ (a, C, R, \delta a, \delta C, \delta R) \in T^*Q : f_a = f_R = 0 \right\} = 0 \times T^*SO(3) \times 0. \] (55)

Since the forces \( F_{nsc}^V \) satisfy d’Alembert principle, its annihilator is
\[ (F_{nsc}^V)^o = \left\{ (a, C, R, \delta a, \delta C, \delta R) : \delta a = r\delta RR^{-1}E_3 \right\} \] (56)
[see (52)]. Then, because
\[ C_V = F_V^o = (F_{nsc}^V + F_{control}^V)^o = (F_{nsc}^V)^o \cap (F_{control}^V)^o, \]
it follows from (55) and (56) that
\[ C_V = \left\{ (a, C, R, \delta a, \delta C, \delta R) : \delta a = r\delta RR^{-1}E_3, \delta C = 0 \right\}. \] (57)

Let us see that \( C_V \) can also be described as
\[ C_V = \left\{ \left( a, C, R, \delta a, 0, \frac{1}{r} \left[ \hat{E}_3, \delta a \right] R + \lambda \hat{E}_3 R \right) : \lambda \in \mathbb{R} \right\}, \] (58)
which will be useful later. If we call \( \alpha \) the vector of \( \mathbb{R}^3 \) related to the antisymmetric matrix \( \delta RR^{-1} \) [via the inverse of \( \hat{\cdot} : \mathbb{R}^3 \rightarrow \mathfrak{so}(3) \)], the condition involving \( \delta RR^{-1} \) translates into
\[ \delta a = r\alpha \times E_3. \]
Thus, using basic properties of the cross product, we have
\[ \alpha = \frac{1}{r} E_3 \times \delta a + \lambda E_3, \]
for some \( \lambda \in \mathbb{R} \). This justifies the description of \( C_V \) given in (58). The same calculations transform \( C_K \), given in (54), into
\[ C_K = \left\{ \left( a, C, R, \dot{a}, \dot{C}, \frac{1}{r} \left[ \hat{E}_3, \dot{a} \right] R + \lambda \hat{E}_3 R \right) : H_i \left( a, C, \dot{a}, \dot{C}, \omega \right) = 0 \right\}. \] (59)

The so constructed GNHS \((L, C_K, C_V)\) [see (53), (54) and (57)] is clearly SO(3)-invariant with respect to the action \( \rho \) given above.
4.2 The related generalized nonholonomic connection

In order to obtain the Generalized Lagrange–d’Alembert–Poincaré Equations related to our triple \((L, C_K, C_V)\), we must follow the steps indicated in Section 3.1. Since

\[
\mathcal{V} = \{(a, C, R, 0, 0, \delta R)\},
\]

the distribution \(S = C_V \cap \mathcal{V}\) results

\[
S = \{(a, C, R, 0, 0, \delta R) : \delta R R^{-1} E_3 = 0\},
\]

using the expression of \(C_V\) given in (57), or equivalently

\[
S = \{(a, C, R, 0, 0, \lambda \hat{E}_3 R)\},
\]

using that given in (58). Let us consider the usual metric on

\[
TQ = T\mathbb{R}^2 \times TS\mathbb{O}(3) \times TS\mathbb{O}(3),
\]

i.e., that defined [on a base point \((a, C, R)\)] as

\[
\langle \delta a, \delta C, \delta R \rangle, \langle \dot{a}, \dot{C}, \dot{R} \rangle = \langle \delta a, \dot{a} \rangle - \frac{1}{2} \text{tr} \left( \delta C^{-1} \dot{C} C^{-1} \right) - \frac{1}{2} \text{tr} \left( \delta R R^{-1} \dot{R} R^{-1} \right).
\]

With respect to this metric, the orthogonal complement of \(S\) inside \(C_V\) is

\[
T = \{(a, C, R, \delta a, 0, \delta R) : \delta a = r \delta R R^{-1} E_3, \; \text{tr} \left( \delta R R^{-1} \hat{E}_3 \right) = 0\},
\]

or equivalently

\[
T = \left\{(a, C, R, \delta a, 0, \frac{1}{r} \left[ \hat{E}_3, \delta \hat{a} \right] R)\right\},
\]

Consider now the sum \(C_V + \mathcal{V}\), which is given by

\[
C_V + \mathcal{V} = T\mathbb{R}^2 \times 0 \times TS\mathbb{O}(3).
\]

It is clear that its orthogonal complement is

\[
\mathcal{R} = 0 \times TS\mathbb{O}(3) \times 0.
\]

Then, the horizontal space that defines the generalized nonholonomic connection \(A\) is

\[
\mathcal{H} = \mathcal{R} \oplus T = \left\{(a, C, R, \delta a, \delta C, \frac{1}{r} \left[ \hat{E}_3, \delta \hat{a} \right] R)\right\}.
\]

This means that connection 1-form \(A : TQ \to so(3)\) is given by the formula

\[
A(a, C, R, \delta a, \delta C, \delta R) = R^{-1} \delta R - R^{-1} \frac{1}{r} \left[ \hat{E}_3, \delta \hat{a} \right] R.
\]
4.3 Reduced data

Now, let us construct the reduced Lagrangian $l$ and the reduced constraints $E_K, E_V$.

Since $Q$ is a trivial principal bundle with base $\mathbb{R}^2 \times \text{SO}(3)$ and structure group $\text{SO}(3)$, we have that

$$Q/\text{SO}(3) = \mathcal{X} = \mathbb{R}^2 \times \text{SO}(3),$$

and $\pi: Q \to \mathcal{X}$ is simply given by projection onto the first two factors. Note that the adjoint bundle is also trivial, i.e.

$$\tilde{\mathfrak{so}}(3) = \mathcal{X} \times \mathfrak{so}(3) = \mathbb{R}^2 \times \text{SO}(3) \times \mathfrak{so}(3).$$

Then, $\pi_*: TQ \to \mathcal{X}$ and $a: TQ \to \tilde{\mathfrak{so}}(3)$ are defined by the formulae

$$\pi_* \left( a, C, R, \dot{a}, \dot{C}, \dot{R} \right) = \left( a, C, \dot{a}, \dot{C} \right)$$

and [recall (33)]

$$a \left( a, C, R, \dot{a}, \dot{C}, \dot{R} \right) = \left( a, C, R^{-1} \dot{R} - R^{-1} \frac{1}{r} \left[ \mathbf{E}_3, \dot{a} \right] R \right),$$

respectively. Consequently, the isomorphism

$$\alpha_A: TQ/\text{SO}(3) \to T\mathcal{X} \oplus (\mathcal{X} \times \mathfrak{so}(3))$$

is given by [see (34)]

$$\alpha_A \circ p \left( a, C, R, \dot{a}, \dot{C}, \dot{R} \right) = \left( a, C, \dot{a}, \dot{C}, \dot{R} R^{-1} - \frac{1}{r} \left[ E_3, \dot{a} \right] \right),$$

$p$ being the submersion $p: TQ \to TQ/\text{SO}(3)$ related to the lifted action $\rho_*$. We are denoting the elements of $T\mathcal{X} \oplus (\mathcal{X} \times \mathfrak{so}(3))$ by $\left( a, C, \dot{a}, \dot{C}, \xi \right)$, with $\left( a, C, \dot{a}, \dot{C} \right) \in T\mathcal{X}$ and $\xi \in \mathfrak{so}(3)$. Note that

$$\xi = \dot{R} R^{-1} - \frac{1}{r} \left[ E_3, \dot{a} \right] = \tilde{\omega} - \frac{1}{r} \left[ \mathbf{E}_3, \dot{a} \right],$$

and in vector notation

$$\xi = \omega - \frac{1}{r} \mathbf{E}_3 \times \dot{a}.$$

From (53) it follows that the reduced Lagrangian

$$l: T\mathcal{X} \oplus (\mathcal{X} \times \mathfrak{so}(3)) \to \mathbb{R},$$

26
which is defined by equation $l \circ \alpha_A \circ p = L$, turns out to be

$$
\begin{align*}
l(a, C, \dot{a}, \dot{C}, \xi) &= -\frac{1}{4} \text{tr} \left( (\dot{C})^{-1} \dot{C} \ddot{C} \right) \\
    &\quad - \frac{1}{4} I \text{tr} \left( \left[ \left( \dot{C} + C \left( \frac{1}{2} [\hat{E}_3, \dot{a}] + \xi \right) \right) C^{-1} \right]^2 \right) \\
    &\quad + \frac{1}{2} m \left( C(a + r \hat{E}_3) + C \dot{a} \right)^2 - m \langle g, C(a + r \hat{E}_3) \rangle.
\end{align*}
$$

The reduced constraints $\mathcal{C}_K, \mathcal{C}_V \subset T \mathcal{X} \oplus (\mathcal{X} \times so(3))$, with

$$
\mathcal{C}_K = \alpha_A \circ p(C_K) \quad \text{and} \quad \mathcal{C}_V = \alpha_A \circ p(C_V),
$$

are given by [remember (58) and (59)]

$$
\mathcal{C}_K = \left\{ (a, C, \dot{a}, \dot{C}, \lambda \hat{E}_3) : h_i (a, C, \dot{a}, \dot{C}, \lambda) = 0 \ (i = 1, 2, 3) \right\},
$$

with the $h_i$'s to be determined, and

$$
\mathcal{C}_V = \left\{ (a, C, \delta a, 0, \lambda \hat{E}_3) \right\}.
$$

Clearly,

$$
\mathcal{C}^\text{hor}_V = \{(a, C, \delta a, 0, 0)\} \quad \text{and} \quad \mathcal{C}^\text{ver}_V = \{(a, C, 0, 0, \lambda \hat{E}_3)\}.
$$

### 4.4 Reduced equations

We are now in conditions to write down the Generalized Lagrange–d’Alembert–Poincaré Equations (48) and (49) for our system. This requires, among other things, to calculate the covariant derivatives $\partial/\partial x$ and $D/Dt$, and the reduced curvature tensor $\tilde{B} : T \mathcal{X} \times T \mathcal{X} \rightarrow \mathcal{X} \times so(3)$. To this end we must proceed in the same way as for a standard Lagrangian system. So, we restrict ourselves to list the results, omitting the related calculations. It can be shown that

$$
\tilde{B}(a, C) \left( \dot{a}, \dot{C}, \delta a, \delta C \right) = \left( a, C, \frac{1}{r^2} [\hat{a}, \dot{a}] \right).
$$

Since we are dealing with a trivial principal bundle, the above mentioned covariant derivatives can be written in terms of partial derivatives and ordinary time derivatives. Once this is done, the horizontal and vertical equations turn out to be

$$
\begin{align*}
\left\langle - \frac{d}{dt} \frac{\partial l}{\partial \dot{a}} + \frac{\partial l}{\partial a}, \delta a \right\rangle + \left\langle - \frac{d}{dt} \frac{\partial l}{\partial \dot{C}} + \frac{\partial l}{\partial C}, \delta C \right\rangle &= \left\langle \frac{\delta l}{\partial \xi}, \frac{1}{r^2} \left[ \hat{a}, \dot{a} \right] \right\rangle \\
    &\quad - \left\langle \frac{\partial l}{\partial \xi}, \left[ \frac{1}{r} [\hat{E}_3, \delta \hat{a}] \right], \xi \right\rangle.
\end{align*}
$$

27
\[ \delta a \text{ being arbitrary, } \delta C = 0 \text{ and } \eta = \lambda \hat{E}_3, \text{ with } \lambda \text{ also arbitrary. Therefore, these} \]
\[ \text{equations take the form} \]
\[ -d \frac{\partial l}{\partial \dot{a}} + \frac{\partial l}{\partial a} = \left\langle \frac{\partial l}{\partial \xi}, \frac{1}{r^2} \left[ \hat{a}, \cdot \right] \right\rangle - \left\langle \frac{\partial l}{\partial \xi}, \left[ \frac{1}{r} \left[ \hat{E}_3, \cdot \right], \xi \right] \right\rangle \]
\[ \text{and} \]
\[ \left\langle d \frac{\partial l}{\partial \xi} + \frac{\partial l}{\partial a} \right\rangle = 0. \]

In vector notation,
\[ -d \frac{\partial l}{\partial \dot{a}} + \frac{\partial l}{\partial a} = p \left[ \frac{1}{r^2} \frac{\partial l}{\partial \xi} \times \hat{a} - \frac{1}{r} \left( \xi \times \frac{\partial l}{\partial \xi} \right) \times E_3 \right] \quad (60) \]
and
\[ \left\langle d \frac{\partial l}{\partial \xi} + \frac{\partial l}{\partial a} \times \xi, E_3 \right\rangle + \frac{1}{r} \left\langle \frac{\partial l}{\partial \xi}, \hat{a} \right\rangle = 0. \quad (61) \]

We are seeing $\partial l/\partial \dot{a}$ and $\partial l/\partial a$ as vectors in $\mathbb{R}^2$, $\partial l/\partial \xi$ as one in $\mathbb{R}^3$, and we are denoting by $p$ the projection of $\mathbb{R}^3$ onto $\mathbb{R}^2$ given by the projection onto the first two factors.

In the following, we shall derive more concrete expressions for these equations.

The relevant partial derivatives of $l$ are
\[ \frac{\partial l}{\partial a} = p \left[ -m \left( \Omega \times (\Omega \times (a + r E_3)) + \hat{a} \right) + C^{-1} g \right], \]
\[ \frac{\partial l}{\partial \dot{a}} = p \left[ m (\Omega \times (a + r E_3)) + \hat{a} + \frac{I}{r} (\Omega + \xi) \times E_3 + \frac{1}{r} \hat{a} \right], \]
\[ \frac{\partial l}{\partial \xi} = l \left( \Omega + \xi + \frac{1}{r} E_3 \times \hat{a} \right). \]

Replacing this in (60) and (61) we arrive at the equations
\[ p \left[ (1 + \alpha) \hat{a} + \hat{\Omega} \times (a + r (1 + \alpha) E_3) + r \hat{\xi} \times E_3 + (2 + \alpha) \Omega \times \hat{a} \right] \]
\[ + p \left[ \Omega \times (\Omega \times (a + r E_3)) + r \alpha (\Omega \times \xi) \times E_3 + C^{-1} g \right] = 0 \]
and
\[ \left\langle \dot{\hat{\Omega}} + \dot{\hat{\xi}} + \Omega \times \xi, E_3 \right\rangle + \frac{1}{r} \left\langle \Omega, \hat{a} \right\rangle = 0 \]
with $\alpha = I/mr^2$. Using one of the kinematic constraints
\[ \xi = \lambda E_3, \]

28
where we are introducing a new variable $\lambda$, we have

$$p \left[ (1 + \alpha) \ddot{a} + \dot{\Omega} \times (a + r (1 + \alpha) E_3) + (2 + \alpha) \Omega \times \dot{a} \right]$$

$$+ p \left[ \dot{\Omega} \times (\Omega \times (a + r E_3)) - r \alpha \lambda \Omega + C^{-1} g \right] = 0$$

and

$$\left\langle \dot{\Omega}, E_3 \right\rangle + \dot{\lambda} + \frac{1}{r} \left\langle \Omega, \dot{a} \right\rangle = 0.$$ 

In components,

$$(1 + \alpha) \dot{a}_1 + \dot{\Omega}_3 a_2 + r (1 + \alpha) \dot{\Omega}_2 + (2 + \alpha) \Omega_3 \dot{a}_2$$

$$+ \Omega_1 (\Omega_1 a_1 + \Omega_2 a_2 + r \Omega_3) - a_1 (\Omega^2_1 + \Omega^2_2 + \Omega^2_3) - r \alpha \lambda \Omega + C_{31} g = 0,$$

and

$$(1 + \alpha) \dot{a}_2 - \dot{\Omega}_3 a_1 - r (1 + \alpha) \dot{\Omega}_1 - (2 + \alpha) \Omega_3 \dot{a}_1$$

$$+ \Omega_2 (\Omega_1 a_1 + \Omega_2 a_2 + r \Omega_3) - a_2 (\Omega^2_1 + \Omega^2_2 + \Omega^2_3) - r \alpha \lambda \Omega + C_{32} g = 0,$$

and

$$\dot{\Omega}_3 + \dot{\lambda} + \frac{1}{r} (\Omega_1 \dot{a}_1 + \Omega_2 \dot{a}_2) = 0.$$ 

We must add the remaining kinematic constraints $h_i = 0$ of $\mathcal{C}_K$ in order to have all of the equations of motion of the system. We want to emphasize again that we have not needed to specify the kinematic constraints in order to write the reduced equations. As a consequence, if we want to design a constraint to stabilize a given Lagrangian system with symmetry, we can analyze the reduced equations instead of the original ones, which can imply a significant simplification of the problem. This is what we shall do in a forthcoming paper.

Acknowledgements

S. Ferraro wants to thank Project SIMUMAT S-0505/ESP/0158 of the CAM for a Research contract.

References


